Circuits and Systems
Mekelweg 4,
2628 CD Delft
The Netherlands
http://cas.tudelft.nl/

M.Sc. Thesis

Analyzing dynamic functional connectivity using state-space models on mice fUS data

Ruben Wijnands

Analyzing dynamic functional connectivity using state-space models on mice fUS data

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by

Ruben Wijnands
born in Katwijk, The Netherlands

This work was performed in:

Circuits and Systems Group
Department of Microelectronics & Computer Engineering
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Delft University of Technology

Delft University of Technology
Department of
Microelectronics & Computer Engineering

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled **"Analyzing dynamic functional connectivity using state-space models on mice fUS data"** by **Ruben Wijnands** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: 30/06/2022

Chairman: _____
Prof. Dr. Ir. J. Dauwels

Advisors: _____
Prof. Dr. Ir. J. Dauwels

_____
Prof. Dr. B. Hunyadi

Committee Members: _____
Prof. Dr. A. Badura

Abstract

In recent years, the increase in brain research led to the development of large-scale brain imaging techniques. With large-scale brain imaging techniques, such as functional magnetic resonance imaging (fMRI), functional connectivity analyses have shown altered connectivity patterns in humans and mice with neurobiological disorders, such as autism spectrum disorder (ASD). To further investigate different mutations that contribute to ASD, a behavioral neuroscientific experiment has been performed at the neuroscientific department of Erasmus MC. During the trial, brain activity in groups of wild-type and homozygous mice is measured using functional ultrasound (fUS) while allowing for spontaneous behaviors of mice. The homozygous mice lack the SHANK2 protein, resulting in hyperactivity and autistic-like behavioral alterations associated with ASD in humans. Understanding the origin of ASD is key to providing effective treatment. However, with the introduction of novel large-scale brain imaging techniques such as fUS, new methods have to be developed that enable functional connectivity analyses. Furthermore, new insights might also be acquired from dynamic functional connectivity analyses, in which changes in functional connectivity over time are evaluated. Subsequently, the question remains if it is possible to unravel differences in brain dynamics between wild-type and homozygous mice using a dynamic functional connectivity analysis.

First, an fUS data model is developed to model how fUS signals arise from a generative perspective. This model comprises a combination of a convolutive and a state-space model. Subsequently, inference of functional networks and their temporal dynamics can be performed. Also, a pre-processing pipeline for experimental fUS data is designed to reduce problem complexity and data cleaning. The performance of the developed methods is evaluated on the experimental data set, where a difference in brain dynamics between wild-type and homozygous mice is investigated.

It is found that a deconvolution procedure using the non-negative least absolute shrinkage and selection operator (NNLASSO) is necessary to reconstruct the underlying activity of neural populations. After that, using the hidden Markov model (HMM) as a state-space model, it is found that functional networks and their temporal dynamics can be learned from fUS data using expectation maximization (EM). It has been discovered that the developed methods consistently decompose reconstructed neural activity into biologically plausible functional networks from experimental fUS data. Also, with 96% certainty, a difference in brain dynamics between wild-type and homozygous mice is found using this method.

In summary, in this thesis, novel methods are developed to perform a dynamic functional connectivity analysis on experimental fUS data. Also, by performing such dynamic functional connectivity analysis for the first time on fUS data, a consistent decomposition of reconstructed neural activity into biologically plausible functional networks and a possible difference in brain dynamics between wild-type and homozygous mice are found. This research highlights the potential of fUS as a large-scale brain imaging technique in the quest to understand the origin of ASD and other neurobiological disorders.

Acknowledgments

I would like to start by expressing my deepest gratitude to my supervisors, Prof. Dr. Ir. Justin Dauwels and Prof. Dr. Borbála Hunyadi, for their valuable feedback, interesting discussions, and trust. Also, I am honored to have collaborated with Prof. Dr. Aleksandra Badura, Dr. Ines da Silva Serra, and MSc. Saffira Tjon. I am thankful for their support and enthusiasm. Never would I have thought to learn and be challenged this much within a year. I am grateful to bring a contribution to this field of research together with the five of you, and I have the utmost respect for your work and expertise.

I also want to thank my family, especially my parents Martijn and Nel Wijnands and girlfriend Connie Brouwer, for their help and support during difficult times. These people ensured I brought the best out of myself and helped me grow personally.

Finally, I am also grateful to all my study mates, colleagues at CAS, and friends at home for being part of this successful and enjoyable journey. I am happy to have shared many beautiful moments together.

Ruben Wijnands
Delft, The Netherlands
30/06/2022

Contents

List of Figures

List of Tables

List of Abbreviations

ASD	autism spectrum disorder
BOLD	blood oxygen level dependent
CAP	co-activation pattern
CBV	cerebral blood volume
CUBE	Center for Ultrasound and Brain imaging at Erasmus MC
EEG	electroencephalography
EM	expectation maximization
ERQ	engineering research question
fMRI	functional magnetic resonance imaging
fUS	functional ultrasound
HMM	hidden Markov model
HOM	homozygous
HRF	hemodynamic response function
ICA	independent component analysis
iCAP	innovation-driven co-activation pattern
MEG	magnetoencephalography
MSE	mean squared error
NNLASSO	non-negative least absolute shrinkage and selection operator
NNLS	non-negative least-squares
NRQ	neuroscientific research question
NVC	neurovascular coupling
PCA	principal component analysis
PDI	power Doppler image
PET	positron emission tomography
PFM	paradigm free mapping
PPA	point process analysis
ROI	region of interest
sICA	spatial independent component analysis
SNR	signal-to-noise ratio
SVD	singular value decomposition
TA	total activation
TFM	temporal functional mode
tICA	temporal independent component analysis
WT	wild-type

1 Introduction

It is perhaps the most undiscovered and complex system known on earth. Recently, the quest to understand the human brain has led to an increase in brain research and the development of brain activity recording techniques. As technology advances, more precise measurement equipment becomes available, providing the capability to gain more knowledge about the brain. Whereas a few decades ago, scientists were only able to measure the activity of a few brain cells, nowadays, thousands of neurons can be measured simultaneously [1]. Combined with advanced signal processing techniques, these technological developments contribute to visualizing the brain structure and large-scale brain function at higher spatiotemporal resolution than ever before. Subsequently, a better understanding of the complex organization of the brain is desired.

In neuroimaging, two imaging categories can be considered: structural and functional brain imaging. By using structural imaging methods, anatomical properties of the brain are visualized. However, in structural imaging, brain function is not measured. Functional imaging focuses on revealing physiologic activities and the underlying processes within the brain. With the use of functional imaging techniques, the notion of functional connectivity has been developed. Functional connectivity can be defined as the connectivity between anatomically distinct brain regions, measured by the statistical relationship between the time courses measured at each region of interest. Multiple regions that are functionally connected comprise a so-called functional network. Until recently, the stationarity of these networks was assumed. However, it is found that the structure of these networks is changing dynamically in time. This discovery led to the development of dynamic functional connectivity, in which changes in functional connectivity of networks are measured. Measures of (dynamic) functional connectivity have proven useful in characterizing abnormalities in brain connectivity. For example, altered connectivity patterns in disorders such as autism spectrum disorder (ASD) have been identified in humans and mice [2, 3]. Identifying disorder-specific functional connectivity alterations can help understand the origin of such disorders to improve effective treatment.

With the introduction of a relatively new brain imaging technique called functional ultrasound (fUS), the brain can be imaged at a higher spatiotemporal resolution than before. Currently, fUS is mainly a 2D imaging technique, although some 3D implementations are deployed. The fUS imaging technique measures the cerebral blood volume (CBV) in the brain. It thereby is an indirect method of recording the activity of neural populations, as the underlying activity of neural populations is coupled with the CBV through the neurovascular coupling (NVC).

With the development of new neuroimaging techniques, also issues arise regarding the reliability and reproducibility of results. For example, a popular, well-established neuroimaging technique called functional magnetic resonance imaging (fMRI) requires extensive workflows, including many parameters. Recent research discovered that this

analytic flexibility manifests itself in significantly varying outcomes among different researchers [4], and severe artifacts [5]. Thus, with the introduction of the novel large-scale fUS recording technique, reliable signal processing methods have to be developed that can infer networks of functional connectivity and their temporal dynamics.

In this thesis, a signal model for fUS data is developed, and a subsequent state-space inference approach is leveraged to analyze dynamic functional connectivity in mice fUS data. The developed technique is applied to a specific neuroscientific experiment, explained in Sec. 1.1. Subsequently, the problem statement that arises from the experiment is treated in Sec. 1.2. Next, the outline of the thesis is presented in Sec. 1.3. This thesis is performed in collaboration with the Center for Ultrasound and Brain imaging at Erasmus MC (CUBE) and the Department of Neuroscience at Erasmus MC in Rotterdam, Netherlands.

1.1 Experiment

The behavioral neuroscientific experiment at hand consists of two groups of mice. The first group is the control group composed of four wild-type (WT) mice, and the other group is the mutated group consisting of four homozygous (HOM) mice. The mutated mice lack the SHANK2 protein, resulting in hyperactivity and autistic-like behavioral alterations. At the CUBE, 2D recordings consisting of power Doppler images (PDIs) of a slice of the mouse brain have been made at a sampling rate of 4 Hz. Each recording lasts 12 minutes, resulting in 2880 successive images per recording. Specifically, the brain section containing the thalamus and motor cortex has been imaged. This section, related to the movement behavior of mice [2], is of particular interest considering the measurement setup and the goal of the experiment. The mice are placed on top of a squirrel cage with their heads fixed by a metal plate, such that their position is fixed and reliable imaging can be performed. Also, the fUS recording is performed in a dark room, so no visual impulses are present during the experiment. The data set is a semi-resting-state data set since the mice are free to show behavioral activity, such as, e.g. movement on top of the squirrel cage, grooming, sniffing, or whisker movement. The fUS data contains approximately similar amounts of sustained movement periods for each mouse. Extra recordings were made if a mouse did not move much during a single recording. This process repeated until a mouse had walked for a significant period, typically resulting in double and triple recording sessions. Also, simultaneously with the fUS data acquisition, the movement of the mouse is recorded by video.

1.2 Problem statement

Concerning the problem statement, a differentiation is made between engineering research questions (ERQs) and neuroscientific research questions (NRQs). By starting with the ERQs, the first question is as follows:

ERQ 1. How can functional networks and their temporal dynamics be learned from fUS data?

Once a method is found that can learn those functional networks, also called brain states, and their temporal dynamics, the specific method can be applied to the experimental fUS data set. Before applying these techniques, the experimental data requires a pre-processing stage, in which data complexity must be reduced while maintaining a certain desired level of detail. This results in ERQ 2:

ERQ 2. What pre-processing techniques are necessary to reduce problem complexity?

Furthermore, it is of interest to analyze the influence of noise on the inference methods, as in real-world systems, noise will undoubtedly be present. This interest leads to posing ERQ 3:

ERQ 3. How do different noise intensities influence the inference of temporal dynamics?

In order to answer this question, synthetic data is generated, and different noise intensities are imposed on the clean synthetic data to analyze the influence on the inference of temporal dynamics. After artificial data analysis, the experimental data is pre-processed, and subsequently, functional networks and their dynamics can be learned from fUS data using the developed methods of ERQ 1. However, having inferred the parameters of interest, it is unknown whether the learned functional networks and their dynamics are reliably and consistently estimated, leading to ERQ 4.

ERQ 4. How consistent are the inferred functional networks and their temporal dynamics?

Now, transitioning from the engineering research questions to the neuroscientific research questions, the first neuroscientific research question, NRQ 1, concerns the biological plausibility of inferred functional networks:

NRQ 1. Are the networks of functional connectivity meaningful from a neuroscientific perspective?

Finally, the last NRQ concerns the ability to reveal any differences between WT and HOM mice, based on dynamic functional connectivity

NRQ 2. Is it possible to differentiate normal and mutated mice based on the temporal dynamics of functional networks?

1.3 Outline

First of all, the preliminaries, state-of-the-art methods for dynamic functional connectivity analysis, and the novel contribution of this thesis are treated in Ch. 2. Then, a signal model for fUS data is constructed in Ch. 3, to understand how fUS signals arise from a generative perspective. Having a signal model, the methods required to learn the functional networks and their dynamics from experimental data are outlined in Ch. 4. After that, in Ch. 5 a simulation is created to analyze the performance of the methods on synthetic data. Next, the methods are applied to the experimental data set, and the results are illustrated in Ch. 6. Finally, the experimental data analysis results are discussed in Ch. 7, and conclusions are drawn in Ch. 8.

2 Literature review

This chapter reviews relevant basic ideas and presents an overview of state-of-the-art methods for analyzing dynamic functional connectivity. The literature review has been performed as part of the extra project course ET4399. First of all, the physiological principles used to measure brain activity are treated in Sec. 2.1. Then, the functional ultrasound (fUS) imaging technique is explained in Sec. 2.2. By having the capability of measuring large-scale brain activity using fUS imaging, methods often used for analyzing dynamic functional connectivity in similar brain imaging techniques are reviewed in Sec. 2.3. Finally, in Sec. 2.4, the contribution of this thesis to the field of brain research, and to be specific, dynamic functional connectivity, is discussed.

2.1 Physiological principles

In this literature review, two fundamental physiological principles, which enable brain activity measurement, are evaluated. Brain activity can directly be measured by recording the electrical activity generated by neural interactions or indirectly by measuring the hemodynamic response due to the neurovascular coupling (NVC) [6]. Another method for measuring brain activity is via metabolism, performed in positron emission tomography (PET). This method requires the injection of a foreign substance that is subsequently detected. Due to relatively slow image acquisition and thus not being suitable for dynamic functional connectivity analysis, this method is left unexplored.

2.1.1 Electrical activity

According to literature, Richard Caton was the first person to measure electrical activity in rabbit and monkey brains using a galvanometer in 1874. Electrical activity in the human brain was first measured by German scientist Hans Berger using electroencephalography (EEG), published in 1929 [7]. More than half a decade later, theories on the physiology of this phenomenon were developed [8]. Electrical activity in the brain is exploited by techniques such as EEG and magnetoencephalography (MEG). However, these techniques are not used for measuring brain activity in this thesis and are thus not explained in detail.

2.1.2 Neurovascular coupling (NVC)

The neurovascular coupling (NVC), also known as functional hyperemia, refers to the mechanism that serves the increase in blood flow and volume in regions of heightened neural activity. For the past 25 years, there has been extensive research into the underlying mechanism of the NVC. These efforts are, amongst others, due to the clinical use of blood oxygen level dependent (BOLD) functional magnetic resonance imaging

(fMRI), which is built on the local hemodynamic response caused by the activation of the NVC. While the exact mechanism of the NVC is partly unknown, it is known to consist of a complex interplay between cellular and metabolic processes. These processes contribute to vasodilation and vasoconstriction, which regulate the increase in blood flow and volume in regions of heightened neural activity.

It is believed that due to a lack of energy, oxygen and glucose are delivered to activated brain areas by an increased blood flow. Blood support can sustain neural activity and clear toxic by-products from metabolic processes. Thereby, the homeostasis of the cerebral micro-environment can be maintained [9].

The coupling between neural activity and the corresponding change in blood flow is governed by the interaction of neurons and astrocytes with smooth muscle cells in the wall of arteries and arterioles. Astrocytes are star-shaped glial cells, collectively named astroglia. Neurons can directly contribute to a hemodynamic response by releasing two potent vasodilators: nitric oxide (NO) and prostaglandin (PG). During a stimulus, also astrocytes can contribute to vasodilation via one or a combination of NO, PG, epoxyeicosatrienoic acids (EET), and potassium (K^+). Also, astrocytes can produce arachidonic acid (AA), which can contribute to vasoconstriction [10]. Besides interaction with smooth muscle cells, also pericytes are believed to contribute to vasodilation in capillaries, which are the smallest blood vessels in the body, conveying blood from arterioles to the venules. The exact mechanism and contribution of capillary pericytes to regional blood flow changes remains controversial [10]. However, it is shown that functional hyperemia starts in capillaries due to oxygen depletion [11].

2.2 Functional ultrasound (fUS)

The functional ultrasound (fUS) imaging method is a relatively new technique that measures hemodynamic changes proportional to the cerebral blood volume (CBV) [12]. Using the μDoppler sequence, fUS is able to image the brain at a high spatiotemporal resolution and high SNR, without the use of contrast agents [6]. Now, 2D fUS imaging techniques are primarily used to image rodent or human brains. For rodent functional imaging, often, the skull is removed or thinned with a drill to ensure the penetration of transmitted ultrasound waves. Human brain imaging is less common and is performed intracranially. In contrast, imaging in infants can be performed directly through the fontanel.

2.2.1 Signal acquisition

In conventional ultrasound imaging, focused beam scanning is used. However, this technique is too slow to acquire brain images at the rate necessary for a functional evaluation. Therefore, the μDoppler sequence has been developed to improve the spatiotemporal resolution of ultrasound imaging. Tilted plane-wave illumination is used as part of the μDoppler sequence, as plane-wave illumination is faster than focused beam scanning. Also, higher resolution and lower noise characteristics are obtained by summing the resulting images of back-scattered tilted plane waves, resulting in a compound image. The number of tilted plane waves used to construct a single compound

image is restricted by the attenuation time of a previously emitted plane wave [13].

In Doppler imaging, blood flow is measured by comparing variations in ultrasonic energy back-scattered by red blood cells. After rejecting tissue signals by the use of a high-pass filter, the filtered signal $s_F(u, v, t_i)$ at pixel location (u, v) with time $t_i = 1, ..., N_s$ is obtained. From the high-pass filtered signal, two parameters can be computed. First, the velocity of blood flow can be calculated. In fUS, however, the power Doppler signal is computed, which is proportional to the CBV and provides no information on the blood flow velocity. The power Doppler signal at time t is the average squared pixel intensity $I(u, v)$ at pixel location (u, v) over N_s samples, calculated according to Eq. 2.1 [13]:

$$I(u, v) = \frac{1}{N_s} \sum_{i=1}^{N_s} |s_F(u, v, t_i)|^2, \tag{2.1}$$

To reduce the influence of noise, N_s must be increased, while accounting for the underlying dynamics of the measured signal. Another important parameter is the Doppler frequency f_D, given in Eq. 2.2:

$$f_D = \frac{2 f_{us} w_v}{c}, \tag{2.2}$$

where f_{us} is the emitted ultrasound frequency, w_v is the velocity of red blood cells along the beam axis, and c is the speed of sound. To prevent aliasing, ultrasound systems sample at twice the Doppler frequency f_D [6].

2.2.2 Performance

Currently, ultrafast ultrasound imaging systems are available with probes operating in the order of MHz, resulting in spatial resolutions in the order of 100 μm $\times$ 100 μm. As there are summation and averaging steps involved in the process of obtaining a single power Doppler image (PDI), the resulting whole-brain imaging rate is in the order of a few Hz [13]. As previously mentioned, primarily 2D imaging techniques are available. However, recently measurements have been performed with 3D fUS imaging systems [14], using a matrix probe, which could open up a wide range of possibilities and new insights into dynamic functional connectivity.

2.3 Methods for dynamic functional connectivity analysis

The assumption of stationarity of brain networks and subsequent static functional connectivity analyses provided insight into the functional connectivity of large-scale brain networks. However, methods analyzing static functional connectivity do not account for changes in functional connectivity over time, referred to as dynamic functional connectivity. Previously found static functional networks can change both strength and directionality, being a change in polarity of correlation of functional connections over time [15, 16]. Also, functional networks can be spatially overlapping, where the time series of a single node or region in a functional network can partially be correlated

with the time series of multiple networks [15]. Or seen from another perspective, nodes or regions of a large-scale static network can become uncorrelated in specific periods. This work aims to quantify changes in functional connectivity metrics over time, as functional connectivity changes according to external and internal processes affecting the brain [15, 17].

In this section, methods for resolving dynamic functional connectivity networks are discussed. Although not explicitly mentioned, many of these techniques are applied in the well-researched field of fMRI. The first method discussed in Sec. 2.3.1 is the sliding window method, in which measurement data is windowed for subsequent analysis. Also, methods such as temporal independent component analysis (tICA) and co-activation patterns (CAPs), that evaluate functional connectivity dynamics at a finer temporal scale, are reviewed in Sec. 2.3.2 and Sec. 2.3.3. Dynamic functional network interactions and transitory roles are typically investigated by modeling latent brain states, discussed in Sec. 2.3.4.

2.3.1 Sliding window analysis

Sliding window analysis is often used to capture temporal changes in functional connectivity. Here, functional connectivity metrics are computed by choosing a fixed-length time window and computing the desired metric in this window. Subsequently, the time window is shifted with some overlap to capture temporal variations in functional connectivity [15]. By correlating the windowed time series of regions of interest (ROIs), a correlation or covariance matrix can be established, measuring the functional connectivity between the selected ROIs. Correlation matrices originating from applying multiple successive sliding windows can be grouped using clustering methods, identifying repeating transient patterns of functional connectivity. A cluster containing multiple correlation matrices with similar characteristics is often called a brain 'state' or 'functional network'.

There are, however, some issues regarding sliding window analysis. First of all, the window length is set more or less arbitrary and is not necessarily suitable for analyzing the data, as relatively long or short changes in functional connectivity could be ignored [18]. Also, a small window size can lead to noisy estimates. It is thus essential to note that the notion of 'functional networks' in sliding window analysis is elusive, as it is dependent on the time scale of evaluation.

Another option is time-frequency analysis with wavelets, similar to applying multiple sliding windows with varying lengths, and thus captures changes at different time scales [19]. Higher frequencies are captured with relatively shorter time windows, whereas low frequencies are captured with relatively longer windows. However, such a time-frequency analysis leads to an explosion in the amount of data, and a selection of suitable parameters should be performed [15].

2.3.2 Temporal independent component analysis (tICA)

For dynamic functional connectivity analysis, also temporal independent component analysis (tICA) can be applied to single-subject data, which assumes temporal independence, in contrast to the spatial independence encouraging variant of ICA, named

spatial independent component analysis (sICA). Thus, tICA produces spatially overlapping sources that are independent in time, also called temporal functional modes (TFMs) [20]. Therefore, tICA can detect smaller functional networks that are part of larger static functional networks and analyze the dynamics of these sub-networks within the larger functional network found in static functional connectivity analyses. Sometimes, also sICA can be applied prior to tICA to find a set of ROIs and subsequently analyze the temporal characteristics of the regions [21].

However, tICA is likely to suffer from noise in fMRI applications, as there are a relatively small number of time points. Also, similar to sICA, model order selection affects the parcellation of networks. Therefore, tICA is less commonly used for dynamic functional connectivity analysis.

2.3.3 Co-activation patterns (CAPs)

Another view on brain function is that monitored activity primarily originates from separate spontaneous events. These events all produce a hemodynamic response that can be measured by techniques such as fMRI and fUS. Perhaps the most basic method to detect an event is thresholding on the amplitude of a voxel or ROI, also known as point process analysis (PPA).

Another approach for detecting brain activity is to deconvolve a measured time series with a hemodynamic response function. The deconvolution procedure reconstructs the underlying events that trigger a hemodynamic response [22]. For deconvolution approaches, knowledge of the hemodynamic response function (HRF) is required. Since the deconvolution procedure is sensitive to data noise, sparse regularizers are often deployed to obtain a better solution. It is common to enforce sparsity on the activity-inducing signal directly or on the first-order derivative of the activity-inducing signal, often referred to as paradigm free mapping (PFM) and total activation (TA), respectively. These forms of regularization lead to a sparse and a more piece-wise constant activity-inducing signal for PFM and TA, respectively. The complexity of event analysis approaches drastically increases for multiple ROIs or whole-brain imaging. Therefore, spatial configurations of simultaneously activating ROIs, called co-activation patterns (CAPs), can be detected based on PFM and TA activity-inducing signals [16].

The technique of CAPs reveals dynamic functional networks by clustering the co-activation and co-deactivation patterns of regions that behave similarly. The work of [16] demonstrates that seed-based correlated or sICA-derived static resting-state networks are a temporal average of CAPs, showing that functional connectivity networks are not static but varying in time. This method thus emphasizes the non-stationary nature of the brain, i.e., the variability in interactions between brain regions. To detect CAPs, the previously discussed PPA on seeds is applied in [16]. In the study of [23], this method is extended beyond a seed-based approach to a data-driven approach, using k-means clustering of time points with similar spatial distributions of activity. In a more recent study using TA, the first-order derivative of deconvolved fMRI data, called the innovation signal, was used to cluster and recover spatial activity patterns, called innovation-driven co-activation patterns (iCAPs) [20]. Thus, similar to TFMs, CAPs temporally segregate brain data.

2.3.4 State-space approaches

A brain 'state' can be defined as a relatively stable spatial activity pattern in time [19]. A difficulty encountered in modeling states in the brain is the choice of metric. Brain states can be clustered based on the amplitude of ROIs such as in CAPs, or based on correlation or coherence of the time courses of ROIs or sICA-derived time courses with the use of sliding windows. Also, tICA could be applied to evaluate the temporal variations and reveal temporally independent functional networks or states.

Finally, state-space models can be used, such as hidden Markov models (HMMs), which incorporate first- and second-order moments and provide a spatiotemporal insight into the recurring activity of and transitional roles between dynamic functional brain networks. These so-called state-space models alleviate the issue regarding choosing sliding window size for correlation of ROIs or sICA-derived time courses and the need for many time samples in tICA. Also, from a neuroscientific point of view, a state-space-based approach is feasible, as the spatial distribution of neural activation in the mice brain varies during vision, choice, action and behavioral engagement [24]. In the work of [24], it is also found that neural activity is globally distributed through the brain during action engagement, following [25]. In contrast, visual and choice engagement signals only concern a subset of brain regions.

In literature, HMMs have been applied to fMRI data to find relatively stable spatial activity patterns [26, 27, 28]. The samples belonging to a certain state are often modeled to be generated from a multivariate Gaussian distribution. This state-specific multivariate Gaussian distribution is described by a mean, encoding the average amplitude of the observed activity, and a covariance matrix, containing the covariance between ROIs. This method could thus be described as an automated windowing method in which the window width is adapted based on data. However, a changing correlation pattern between time courses of fixed mean activation is hard to identify. Essentially, this requires the separation of two multivariate Gaussian distributions with the same mean activation.

2.4 Contribution

This thesis brings a novel contribution to the field of brain research, and specifically dynamic functional connectivity, by investigating the combination of a deconvolution procedure and an HMM on fUS data. By introducing a deconvolution procedure, the aim is to recover the underlying activity of neural populations of ROIs and subsequently learn networks of functional connectivity and the temporal dynamics of the underlying activity from data. In this way, it is possible to investigate brain activity at a finer temporal scale, overcoming the issue of changing correlation patterns at longer time scales. Furthermore, to the best of the authors' knowledge, this is the first time that dynamic functional connectivity has been analyzed on fUS data. Finally, a neuroscientific contribution is made, as biologically plausible networks of functional connectivity are revealed, which play a role in the spontaneous motor behavior of mice. Also, a potential difference in brain dynamics between wild-type and homozygous mice is revealed.

3 Signal model

In this chapter, the fUS signal model is discussed. The signal model is constructed to create or use methods that can answer ERQ 1, which is as follows:

ERQ 1. How can functional networks and their temporal dynamics be learned from fUS data?

By constructing a signal model, insight can be given into how the fUS signals arise from a generative perspective. Although it is unknown how the signals are exactly generated, the signal model might still be useful for inference of networks of functional connectivity and their temporal dynamics.

The signal model consists of two components. First, a convolutive model, described in Sec. 3.1, is constructed for an fUS time series of a region of interest (ROI). The convolutive model relates the underlying activity of neural populations to the measured fUS signals by simulating the effects of the neurovascular coupling (NVC). Secondly, to investigate dynamic functional connectivity, state-space models are deployed. In particular, the hidden Markov model (HMM) is chosen as a generative model for the underlying activity signals produced by certain subsets of neural populations. The signal model of the HMM is explained in Sec. 3.2.

3.1 fUS data model

Let the fUS time course of an ROI be a vector $\mathbf{f}_m \in \mathbb{R}^{N\times 1}$, where $m \in \{1, ..., M\}$ is the region index. Each discrete time point is denoted by $f_{n,m}$, where $n \in \{1, ..., N\}$ is the discrete time index. This signal is modeled as the linear convolution between the hemodynamic response function (HRF) $\mathbf{h}_m$ and the activity inducing signal $\mathbf{y}_m$. This model is chosen with the support of [29], as it is found that a measured fUS signal can be related to neural activation via the hemodynamic response function:

$$f_{n,m} = \sum_{k=0}^{L-1} h_{k,m} y_{n-k,m}, \tag{3.1}$$

or in a more compact form, using matrix-vector multiplication:

$$\mathbf{f}_m = \mathbf{H}_m \mathbf{y}_m$$

$$\begin{bmatrix} f_{1,m} \\ f_{2,m} \\ \vdots \\ f_{N,m} \end{bmatrix} = \begin{bmatrix} h_{L-1,m} & \dots & h_{0,m} & 0 & \dots & 0 \\ 0 & h_{L-1,m} & \dots & h_{0,m} & \dots & 0 \\ \vdots & 0 & \ddots & & \ddots & 0 \\ 0 & \dots & 0 & h_{L-1,m} & \dots & h_{0,m} \end{bmatrix} \begin{bmatrix} y_{-L+2,m} \\ \vdots \\ y_{1,m} \\ \vdots \\ y_{N,m} \end{bmatrix}, \tag{3.2}$$

where L is the length of the HRF. As the behavior and brain activity of mice are recorded simultaneously, only recovering the $y_{n,m}$ with $n \in \{1, ..., N\}$ is of interest, coinciding with video. Also, it should be noted that the convolution matrix $\mathbf{H}_m$ is truncated and thus wide, as the observed data is only a portion of a longer continuous stochastic process. Now, extending Eq. 3.2 to cover M selected ROIs, the matrix $\mathbf{Y} \in \mathcal{R}^{N \times M}$ is created:

$$\mathbf{Y} = [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_M]\,. \tag{3.3}$$

Assuming the same HRF for each ROI m, a generic convolution matrix $\mathbf{H}$ can be constructed. This results in the following multivariate signal model:

$$\mathbf{F} = \mathbf{H}\mathbf{Y}, \tag{3.4}$$

where the fUS ROI time series $\mathbf{f}_m$ are also stacked column-wise in $\mathbf{F}$. By adding a noise matrix with independent columns, each following an i.i.d. Gaussian white noise process

$$\mathbf{N} = [\boldsymbol{\epsilon}_1 \quad \boldsymbol{\epsilon}_2 \quad \cdots \quad \boldsymbol{\epsilon}_M] \quad \text{with } \epsilon_{n,m} \sim \mathcal{N}(0, \sigma), \tag{3.5}$$

the resulting convolutive signal model is equal to

$$\mathbf{F} = \mathbf{H}\mathbf{Y} + \mathbf{N}. \tag{3.6}$$

3.2 State-space model

Now, by transposing matrix $\mathbf{Y}$, each column $\mathbf{y}_n \in \mathcal{R}^M$ of $\mathbf{Y}^T$ contains an activity-inducing signal at discrete time n for each of the M ROIs. Then, $\mathbf{y}_n$ can be considered a multivariate random variable, and it can be noticed that matrix $\mathbf{Y}^T$ actually contains the realization of a stochastic process Y. The distribution, from which each 'observation' $\mathbf{y}_n$ is generated, can change across time, leading to a non-stationary process. It is reasonable to assume non-stationarity for $\mathbf{y}_n$, since certain brain regions can be activated or not. Regions activating approximately simultaneously, i.e., within a time window of 250 ms determined by the sampling rate, are said to be functionally connected. It is thus assumed that the process is stationary within the duration of one power Doppler image (PDI) and that the underlying activity originates from directly or indirectly interacting neural populations between these regions. Now, the latent sequence Z, which indicates the presence of a certain activity pattern, can be modeled as a stochastic process as well

$$Z = \{z_n : n \in \mathbb{N}\},$$

where z_n is a realization of Z at time instant n. The observed activity patterns $\mathbf{y}_n$ with $n = \{1, ..., N\}$, can be assigned to a set of K discrete hidden states $\mathcal{S} = \{1, ..., K\}$, called the state space, governing and labeling different brain activation patterns. As only one state $z_n \in \mathcal{S}$ can be active at a time, and the state transition probability

does only depend on the present state, this process can be recognized as a first-order discrete-time Markov chain. The state transitions are governed by a state transition probability matrix $\mathbf{A}$ containing elements a_{ij}:

$$a_{ij} = P(z_{n+1} = j | z_n = i). \tag{3.7}$$

With this discrete-time Markov chain, the hidden Markov model (HMM) framework establishes itself by relating the random vector $\mathbf{y}_n$ with observed output variables to the hidden process Z that must be learned from data. The HMM model framework is illustrated in Fig. 3.1.

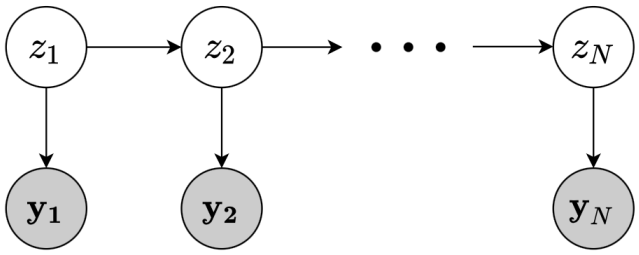


Figure 3.1: Graphical model of the HMM, where white and grey circles are the latent and observed stochastic process, respectively.

For the Gaussian HMM, i.e. an HMM where each state generates emissions from a state-specific multivariate Gaussian distribution, simultaneous observations $\mathbf{y}_n$ are conditionally independent of the past states given z_n:

$$p(\mathbf{y}_n | z_n = i) \sim \mathcal{N}(\boldsymbol{\mu}_{z_n}, \boldsymbol{\Sigma}_{z_n}), \tag{3.8}$$

where $\boldsymbol{\mu}_{z_n} \in \mathcal{R}^M$ and $\boldsymbol{\Sigma}_{z_n} \in \mathcal{R}^{M \times M}$. This means that the functional connectivity pattern is captured by the mean $\boldsymbol{\mu}_{z_n}$ of the multivariate Gaussian distribution. Thus, this finally results in a mixture of multivariate Gaussian distributions for all possible states in $\mathcal{S}$, and only one state and thus one multivariate Gaussian distribution can be active at a time. All state-specific emission distributions are collected in the emission distribution set $\mathbf{B} = \{B_i : i \in \mathcal{S}\}$ with $B_i = \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$. Finally, the last parameter of the HMM is the initial state probability distribution $\boldsymbol{\pi} = \{\pi_1, ..., \pi_K\}$. The HMM model parameters are collected in $\boldsymbol{\theta} = \{\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}\}$.

4 Methods

This chapter treats the methods to learn the functional networks and corresponding temporal dynamics. First, a pre-processing pipeline for the experimental fUS data is developed in Sec. 4.1. Following the previously constructed signal model, it is evident that subsequent deconvolution and HMM inference procedures are required to learn the functional networks and corresponding temporal dynamics from the pre-processed experimental data, treated in Sec. 4.2 and Sec. 4.3, respectively. The deconvolution procedure aims to reconstruct the underlying activity of neural populations from the fUS data. Subsequently, an HMM inference procedure is applied to the reconstructed underlying activity of neural populations to learn the activity patterns in the mice brain from the data.

4.1 Pre-processing

In this section, the pre-processing techniques applied to the fUS data are explained, following the structure of Fig. 4.1. This chapter aims to answer ERQ 2:

ERQ 2. What pre-processing techniques are necessary to reduce problem complexity?

Since the dimensionality of fUS data is high, the data complexity must be reduced. A complexity reduction step is necessary for ease of data interpretation and functionality of the methods, to be discussed in Sec. 4.2 and Sec. 4.3. Furthermore, there is significant variability in the blood vessel structure of mice. Thus, a reliable dynamic functional connectivity analysis on relatively small neural populations is not considered possible. Therefore, the activity of neural populations is evaluated at a somewhat larger scale, namely on the level of anatomical regions. An analysis of the activity in anatomical regions enables a reliable comparison between mice and still maintains enough detail to perform a meaningful study.

However, since fUS measures blood volume, no anatomical regions are apparent in the data. Therefore, the anatomical areas have to be identified. Identification can be done by applying a blind source separation technique called spatial independent component analysis (sICA), explained in Sec. 4.1.1. Since the decomposition of spatially independent sources in sICA is dependent on the model order, there can be different anatomical regions overlapping, or a single anatomical region can decompose. Therefore, the Paxinos brain atlas is warped upon the mouse brain using landmarks. The warping procedure is performed by the Department of Neuroscience at Erasmus MC. Then, both the sICA extracted regions and the Paxinos brain warp are combined to identify the anatomical regions, referred to as regions of interest (ROIs), as described in Sec. 4.1.2. Furthermore, global noise reduction steps are required due to scattering artifacts on the data, as described in Sec. 4.1.3. Also, dimensionality reduction is ap-

plied as outlined in Sec. 4.1.4. Finally, drift correction and offset removal are applied, as treated in Sec. 4.1.5 and Sec. 4.1.6, respectively. In the remainder of this section, a recording labeled with the number 890 and section number 3 will be used for visual illustrations.

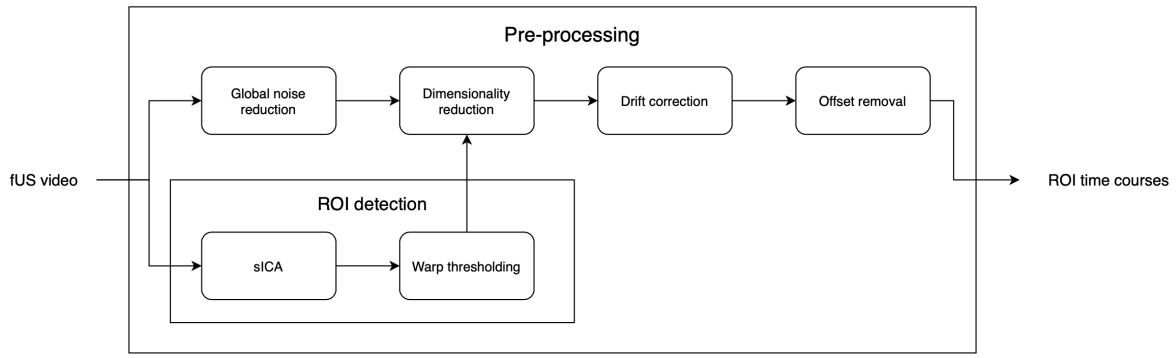


Figure 4.1: Pre-processing pipeline

4.1.1 Spatial independent components analysis (sICA)

A blind source separation technique called spatial independent component analysis (sICA) is used, which uses higher-order statistics to separate spatial sources in data. For this sICA procedure, the SKLEARN.DECOMPOSITION.FASTICA package is used to perform the sICA procedure in Python.

Now, let the fUS data be a 3D matrix $\mathbf{I}$, where an fUS data point is denoted by $\mathbf{I}_n(u, v)$, with $n \in \{1, ..., N\}$ being the discrete time index and (u, v) the pixel location. An image at discrete time n is indicated by $\mathbf{I}_n$, dropping the pixel location. A pixel time course is indicated by $\mathbf{I}(u, v)$, dropping the discrete time index.

Before applying sICA the data is log-transformed and each pixel time course $\mathbf{I}(u, v)$ in the fUS video is standardized temporally. Subsequently, the images $\mathbf{I}_n$ are centered by subtracting the spatial mean. By vectorizing each image $\mathbf{I}_n$, a matrix $\mathbf{X} \in \mathbb{R}^{N \times W}$ can be created, with N discrete time samples and $W = UV$ voxels. In sICA, this matrix is decomposed into

$$\mathbf{X} = \begin{bmatrix} - & \text{vec}(\mathbf{I}_1) & - \\ - & \text{vec}(\mathbf{I}_2) & - \\ & \vdots & \\ - & \text{vec}(\mathbf{I}_N) & - \end{bmatrix} = \begin{bmatrix} d_{1,1} & & d_{1,C} \\ \vdots & \cdots & \vdots \\ d_{N,1} & & d_{N,C} \end{bmatrix} \begin{bmatrix} - & \text{vec}(\mathbf{S}_1) & - \\ - & \text{vec}(\mathbf{S}_2) & - \\ & \vdots & \\ - & \text{vec}(\mathbf{S}_C) & - \end{bmatrix} = \mathbf{DS}, \tag{4.1}$$

where $\mathbf{D} \in \mathbb{R}^{N \times C}$ is called the mixing matrix, containing C time courses of the independent spatial components located in the rows of matrix $\mathbf{S}$, where $\mathbf{S}_c \in \mathbb{R}^{C \times W}$ denotes each spatial component with $c \in \{1, ..., C\}$. In order to simplify the problem of solving the decomposition of Eq. 4.1, principal component analysis (PCA) is often used as a pre-processing step prior to sICA [30]. In PCA, the data is whitened and data dimensionality is reduced. PCA can be performed via the singular value decomposition

(SVD)

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{4.2}$$

where $\mathbf{U} \in \mathbb{R}^{N\times N}$ is a matrix where each column is a left eigenvector of $\mathbf{X}$. Also, $\mathbf{\Sigma} \in \mathbb{R}^{N\times N}$ is a diagonal matrix containing the singular values ordered from largest to smallest, corresponding to the eigenvectors in $\mathbf{U}$. Finally, $\mathbf{V}^T \in \mathbb{R}^{N\times W}$ is a matrix with orthogonal spatial components of each eigenvector in $\mathbf{U}$, which are temporal components. Data dimensionality reduction is performed by setting all singular values after the first C singular values to zero, leading to reduced matrices $\mathbf{U}_r \in \mathbb{R}^{N\times C}$, $\mathbf{\Sigma}_r \in \mathbb{R}^{C\times C}$, and $\mathbf{V}_r^T \in \mathbb{R}^{C\times W}$.

The data can be decorrelated and subsequently whitened by transforming the data via rotation $\mathbf{U}_r^T$ and scaling $\sqrt{N-1}\mathbf{\Sigma}_r^{-1}$, respectively:

$$\mathbf{X}_w = \sqrt{N-1}\mathbf{\Sigma}_r^{-1}\mathbf{U}_r^T\mathbf{X} = \sqrt{N-1}\mathbf{V}_r^T. \tag{4.3}$$

By computing the covariance of $\mathbf{X}_w$, it can be shown that the transformed data is now whitened:

$$\frac{1}{N-1}\mathbf{X}_w\mathbf{X}_w^T = \mathbf{V}_r^T\mathbf{V}_r = \mathbf{\Sigma}_r^{-1}\mathbf{U}_r^T\mathbf{X}\mathbf{X}^T\mathbf{U}_r\mathbf{\Sigma}_r^{-1} = \mathbf{I}. \tag{4.4}$$

Thus, sICA is performed on whitened data $\mathbf{X}_w$, resulting in the following decomposition:

$$\mathbf{X}_w = \mathbf{D}_w\mathbf{S}_w. \tag{4.5}$$

By looking at the original problem

$$\underbrace{\sqrt{N-1}\mathbf{\Sigma}_r^{-1}\mathbf{U}_r^T\mathbf{X}}_{\mathbf{X}_w} = \underbrace{\mathbf{\Sigma}_r^{-1}\mathbf{U}_r^T\mathbf{D}}_{\mathbf{D}_w}\underbrace{\sqrt{N-1}\mathbf{S}}_{\mathbf{S}_w}, \tag{4.6}$$

it appears that an approximation of matrix $\mathbf{D}$ and $\mathbf{S}$ can be found back:

$$\tilde{\mathbf{D}} = \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{D}_w \tag{4.7}$$

$$\tilde{\mathbf{S}} = \frac{1}{\sqrt{N-1}}\mathbf{S}_w. \tag{4.8}$$

Now, biologically plausible spatial configurations can be manually identified from $\tilde{\mathbf{S}}$. For handling the manual labor and interpretability of spatial components, the number of components is set to $C = 25$.

4.1.2 Warp thresholding

As mentioned before, solely trusting sICA extracted regions is not possible due to the possibility of having overlapping or separating spatial sources. Also, larger anatomical areas have to be considered to compare the brain dynamics between the two groups of mice. For this purpose, sICA extracted regions and the atlas warp are used together to extract ROIs.

On the left of Fig. 4.2, the Paxinos atlas is warped upon a logarithmic mean image of the brain over time. Since the Paxinos atlas contains many subdivisions of an anatomical region, a tool is created to connect all small sub-regions within a larger anatomical area. Specifically, a flood-filling algorithm labels all pixels in specific sub-regions belonging to a larger anatomical region. Then, the morphological operations dilation and erosion are applied successively, using a structuring element of 3 by 3 pixels, connecting all sub-regions with the same label, resulting in one complete anatomical region. The result of applying the tool to all anatomical ROIs is visualized on the right of Fig. 4.2. Note that the thalamus area is not selected, as this anatomical region could only be identified in half of the mice. Hence, the thalamus is left out of the analysis.

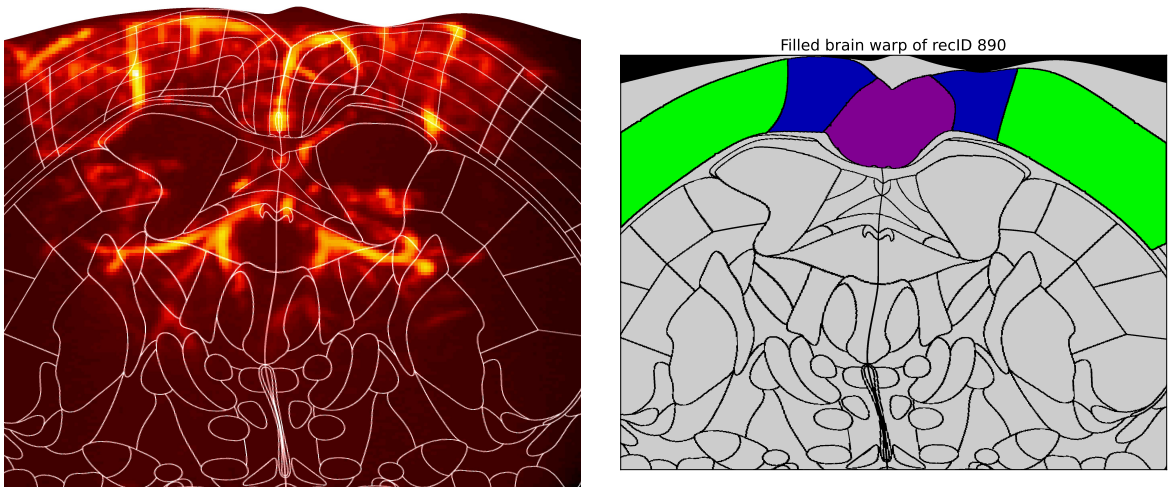


Figure 4.2: **Left:** Paxinos atlas warped upon a logarithmic mean image of the brain over time. **Right:** Masks of anatomical ROIs: anterior cingulate area (purple), motor area (blue), somatosensory area (green).

As a next step, a part of the sICA components is assigned to a specific region in the atlas warp, illustrated on the left of Fig. 4.3. Then, the sICA component pixels are threshold within the particular anatomical area, visualized on the right of Fig. 4.3, to obtain pixels that are part of the anatomical ROI. The threshold is set to 50% between the maximum pixel intensity in the sICA component and the mean background intensity. The background is found by a procedure described in Sec. 4.1.3.

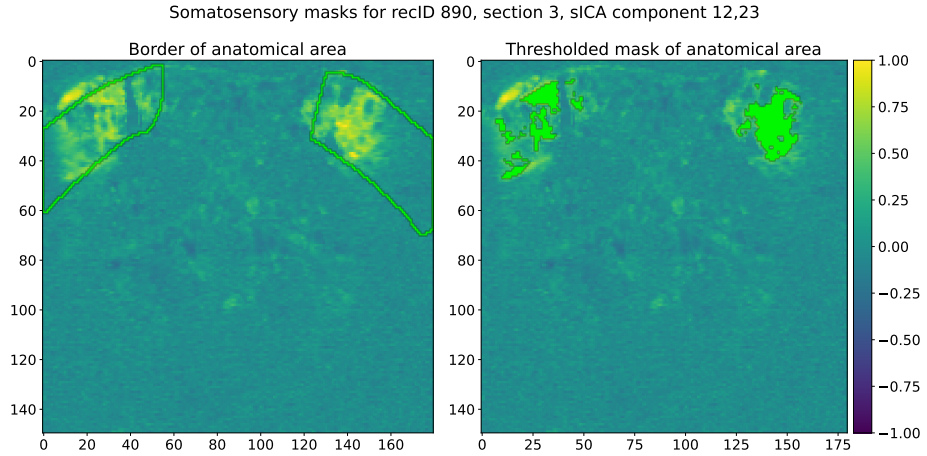


Figure 4.3: **Left:** Border of the regional mask. **Right:** Thresholded pixels within the regional mask.

It should be noted that the sign of sICA components is ambiguous. Thus, the sign of each component is manually inverted or not, such that regions that are believed to be active in the sICA component have a positive sign. Next, an sICA component is rescaled to a maximum value of 1, and the background, containing no blood vessels, is set to a value of 0. Values below -1 are clipped if there are such low values to ensure the same color space for visual comparison. The ROI detection of the other two anatomical areas can be found in Sec. B.3.

4.1.3 Global noise reduction

Furthermore, in the fUS data, a shifting mean of the image in regions without any blood vessels can be recognized. This artifact is best described as a scattering artifact, thus also influencing pixels in ROIs. The scattering artifact could be originating from movement artifacts or earlier pre-processing steps in the fUS acquisition pipeline. Tracing the exact origin of this artifact is viewed outside the scope of this thesis. Fortunately, there are methods to cope with the artifact in an ad hoc manner. In this case, the blood vessels are masked based on the pixel variance over time, as visualized in Fig. 4.4. Then, the spatial mean of the pixels that fall outside the spatial mask can be computed and subtracted from each image. In this way, the brain areas containing no blood vessels are corrected to the same (noisy) baseline. To detect the background, a percentage threshold between the maximum and minimum logarithmic pixel variance is heuristically determined for each recording. Generalization of this threshold is not possible due to the scattering artifact being not homogeneous towards the edges of the images in some recordings, also slightly visible in Fig. 4.4. Consequently, choosing a single percentage threshold leads to mixed quality of background detection.

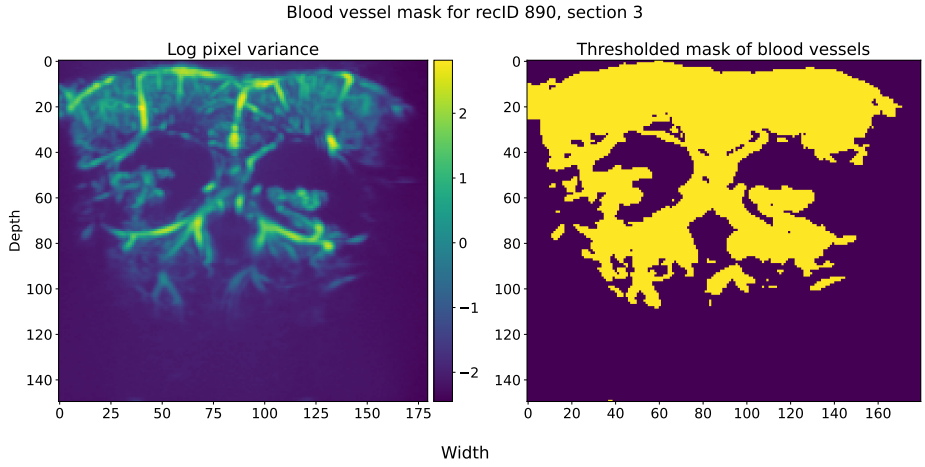


Figure 4.4: **Left:** Log pixel variance over time. **Right:** Thresholded mask

4.1.4 Dimensionality reduction

Since the fUS data is high-dimensional, data dimensionality reduction is considered. An effective method is to compute the mean time series $\mathbf{f}_m$ of the set of pixels in an anatomical ROI, denoted by R_m. The mean fUS signal at a discrete-time point n, indicated by $f_{n,m}$, is computed by:

$$f_{n,m} = \frac{1}{|R_m|} \sum_{\{u,v\} \in R_m} \mathbf{I}_n(u, v). \tag{4.9}$$

where $m \in \{1, ..., M\}$ is the ROI index. The choice to take the mean of an anatomical ROI R_m is further supported by the work of [29], where it is shown that the underlying neural populations largely activate synchronously in an anatomical ROI.

4.1.5 Baseline drift

After the dimensionality reduction step, it is noted that a relatively slow drift is present in some time courses. As an example, the original fUS mean time course of the motor area is depicted in Fig. 4.5. Removing this drift is essential for subsequent processing steps, as will become clear later in Sec. 4.2.

To estimate the signal drift present in the fUS mean time course, linear regression is performed on the time course. However, before applying linear regression, the baseline in the time course has to be detected. For this purpose, movement information is used. Each short period of movement is extended with the approximate duration of the HRF, that is 8 seconds [29]. However, only using movement information is insufficient since other behaviors and internal processes in the brain could also invoke a hemodynamic response, leading to a wrong drift estimation. Therefore, a sliding window with a heuristically determined length of 15 seconds is moved over the time course. In each window, the variance is calculated. Then, a mask is created by thresholding based on a percentage difference between the maximum and minimum variance within the sliding window over the entire time course. The selected percentages per ROI can be found

in Sec. B.2. Next, a brain activity mask is created by combining the variance-based mask and mice movement information. This activity mask, as depicted in black in Fig. 4.5, thus encapsulates more data points than the movement mask, created from video extracted movement behavior of the mouse, solely. Then, linear regression is applied only on data points that fall outside the activity mask, depicted in white in Fig. 4.5, which are data points belonging to the baseline.

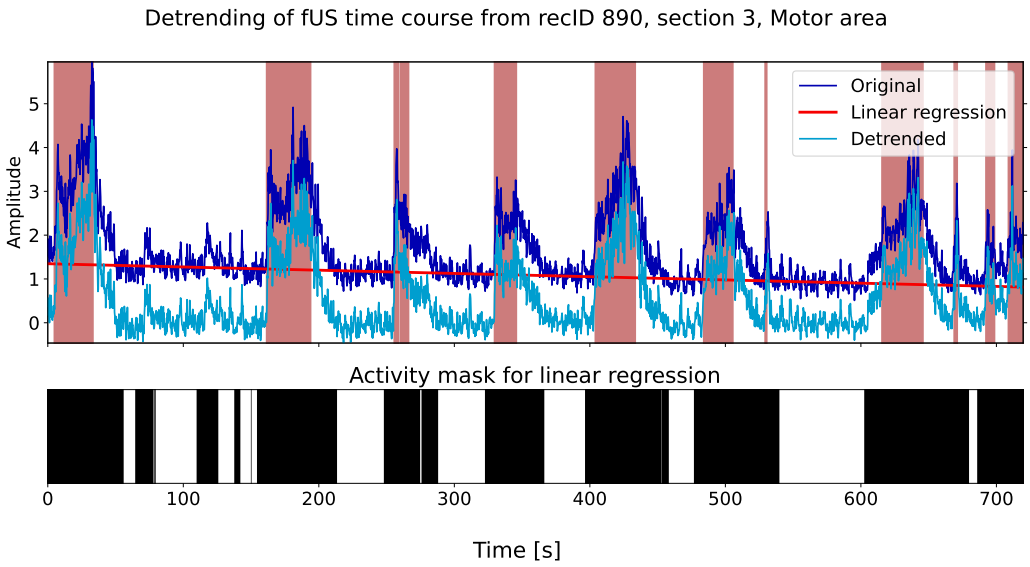


Figure 4.5: Detrending of the motor area time course. Movement of the mouse is depicted in salmon color. Detected hemodynamic activity is depicted in black in the activity mask.

In order to ensure reliable drift estimation, it is required that there are at least 30 data points on which linear regression is applied with a minimum separation of 800 samples, being 200 seconds. In case of too few data points, no drift correction is applied. In Sec. B.4, two cases can be found where no baseline could be detected. However, the effect of not detrending these time courses is considered small, as (almost) no drift correction is necessary in these cases. Also, changing the chosen percentage thresholds would result in more unreliability in baseline detection in time courses of other fUS recordings.

4.1.6 Offset removal

As a final step, the detrended and non-detrended time course baselines have to be shifted to a value of 0 for the method described in Sec. 4.2. Since the activity mask does not perfectly cover some data points corresponding to intrinsic activity, an offset removal step must be performed on the detrended time courses as well. In order to remove the offset, the same sliding window of a heuristically determined size of 15 seconds is applied to the detrended and non-detrended time courses. Then, a region-specific time course is shifted by the smallest mean value within the sliding window. Removing the offset results in the pre-processed time courses illustrated in Fig. 4.6.

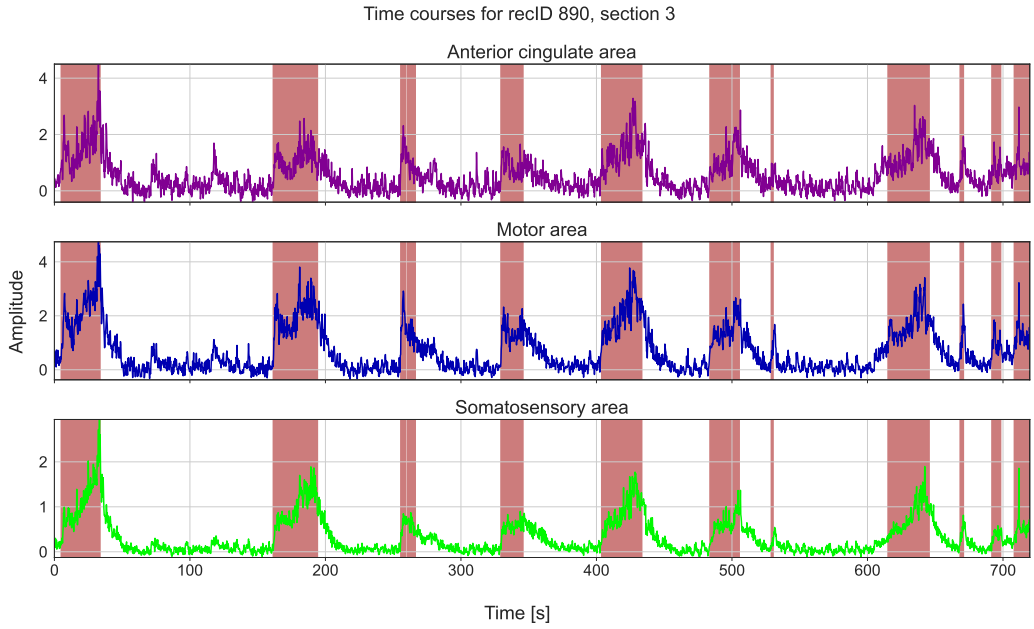


Figure 4.6: Final time courses after the pre-processing stage. The movement of the mouse is depicted in salmon color.

The set of finally obtained time series $\{\mathbf{f}_m : m = 1, ..., M\}$, stacked column-wise in matrix $\mathbf{F}$, is used for subsequent signal processing techniques, discussed below.

4.2 Deconvolution

For convenience, the signal model of Eq. 3.6 is repeated:

$$\mathbf{F} = \mathbf{HY} + \mathbf{N}. \tag{4.10}$$

The goal of the deconvolution procedure is to reconstruct matrix $\mathbf{Y}$. For this purpose, the hemodynamic response function h_n, embedded in the convolution matrix $\mathbf{H}$, is assumed. The HRF is generated through the following function, a modified version of the HRF used in [31]:

$$h_n = p_3 \left(\frac{(t_n)^{p_1 - 1} p_2^{p_1} e^{-p_2 t_n}}{\Gamma(p_1)} \right), \tag{4.11}$$

where $t_n = nT_s$ with T_s being the sampling period. Furthermore, parameters p_1 and p_2 determine the shape of the HRF, and p_3 scales the HRF. The HRFs is scaled to unit amplitude without loss of generality. The parameters used for creating the HRF presented in Fig. 4.7 are $p_1 = 4.00$, $p_2 = 1.50$, and $p_3 = 2.98$, chosen such that the HRF shape is similar to the one presented in [29]. The resulting HRF is shown in Fig. 4.7.

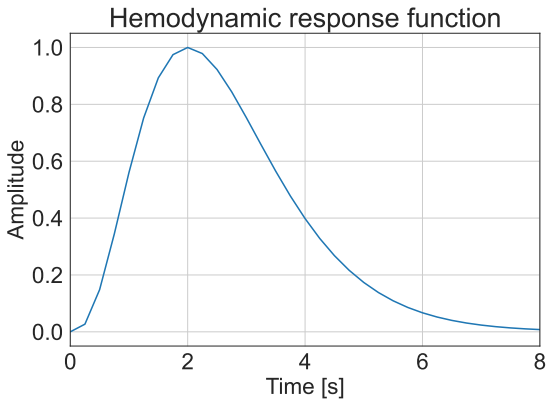


Figure 4.7: Hemodynamic response function used in the convolution matrix $\mathbf{H}$.

With the assumed HRF, the problem of reconstructing matrix $\mathbf{Y}$, containing the underlying activity $\mathbf{y}_m$ of each region m, can be solved by minimizing the noise term for each region m:

$$\hat{\mathbf{y}}_m = \arg\min_{\mathbf{y}_m} \quad ||\mathbf{f}_m - \mathbf{H}\mathbf{y}_m||_2^2. \tag{4.12}$$

However, minimizing the least squares problem is ill-posed due to $\mathbf{H}$ being wide. A popular choice in fMRI analyses is to add constraints or so-called regularizers that either enforce sparsity (PFM) or piece-wise constant (TA) solutions [22]. However, in TA, rapid changes on a short time scale in the activity-inducing signal are not emphasized. In PFM, these rapid changes are promoted by sparse regularization, assuming only a few entries of the activity-inducing signal are non-zero. However, PFM allows for both positive and negative activation, not in accordance with the biological functioning of neurons. Therefore, in this work, a constraint on the non-negativity of the solution is imposed. Now, including non-negativity without sparse regularization, the problem can be solved using non-negative least squares (NNLS):

$$\begin{aligned} \hat{\mathbf{y}}_m = \arg\min_{\mathbf{y}_m} \quad & ||\mathbf{f}_m - \mathbf{H}\mathbf{y}_m||_2^2 \\ \text{s.t.} \quad & \mathbf{y}_m \geq 0. \end{aligned} \tag{4.13}$$

where the constraint implies non-negativity for each element in vector $\mathbf{y}_m$.

Furthermore, by penalizing the number of non-zero entries by enforcing a sparsity regularizer in the cost function using the non-negative least absolute shrinkage and selection operator (NNLASSO), the notion of PFM is expanded:

$$\begin{aligned} \hat{\mathbf{y}}_m = \arg\min_{\mathbf{y}_m} \quad & \frac{1}{2N}||\mathbf{f}_m - \mathbf{H}\mathbf{y}_m||_2^2 + \lambda||\mathbf{y}_m||_1 \\ \text{s.t.} \quad & \mathbf{y}_m \geq 0. \end{aligned} \tag{4.14}$$

The addition of the regularizer supporting sparsity can be regarded as an adjustable addition to the cost function to force the solution in the desired direction, in this case, a more sparse solution. For each mouse and time course $\mathbf{f}_m$ of an ROI, the sparsity

can be different, as it is dependent on the amount of activity over time, as well as the true level of sparsity during activity. Therefore, the regularization parameter λ used for deconvolution with NNLASSO is selected per mean time course $\mathbf{f}_m$. The regularization parameter is selected so the reconstruction error RE is not significantly enlarged. This reconstruction error is measured by the MSE between the observation $\mathbf{f}_m$ and the reconstructed observation $\hat{\mathbf{f}}_m = \mathbf{H}\hat{\mathbf{y}}_m$:

$$RE(\lambda) = \frac{1}{N}||\mathbf{f}_m - \hat{\mathbf{f}}_m||_2^2. \tag{4.15}$$

To be specific, λ is set such that the relative RE between a non-sparse, essentially an NNLS solution, and a very sparse, essentially an all-zero solution, does not exceed a certain percentage. This way, the imposed sparsity levels are similar across the different optimization problems. In this case, the regularization parameter is selected such that the relative error introduced by the sparsity constraint does not exceed 1%, visualized in Fig. 4.8.

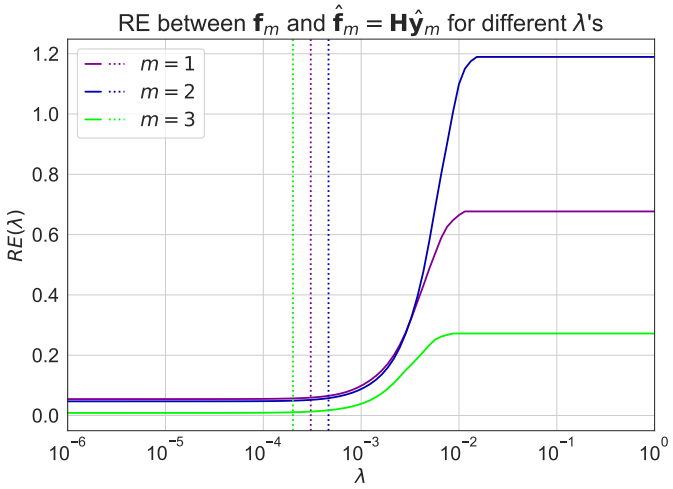


Figure 4.8: Determination of regularization parameters. Dotted lines represent the chosen value for λ for each ROI m, for which the relative reconstruction error does not exceed 1%.

This results in the following deconvolved time courses, being a reconstruction of the underlying activity of neural populations, depicted in Fig. 4.9. As visible, the reconstructed underlying activities follow the mouse movement closely but are not solely dependent on the movement. Since the hemodynamic response is a slow response to the underlying activity, it is also expected that the underlying activity follows the behavior of the mouse more closely. In the deconvolution procedure, the first and last $L-1$ indices of the underlying activity $\mathbf{y}_m$ cannot be estimated accurately. Therefore, these indices are discarded afterward, losing $L-1$ samples from the original data. Effectively, the last $L-1$ reconstructed underlying activity samples and corresponding video samples are lost. The first $L-1$ samples, corresponding to a short time before the beginning of the fUS acquisition, are already not of use, as the fUS data and mouse video are recorded simultaneously.

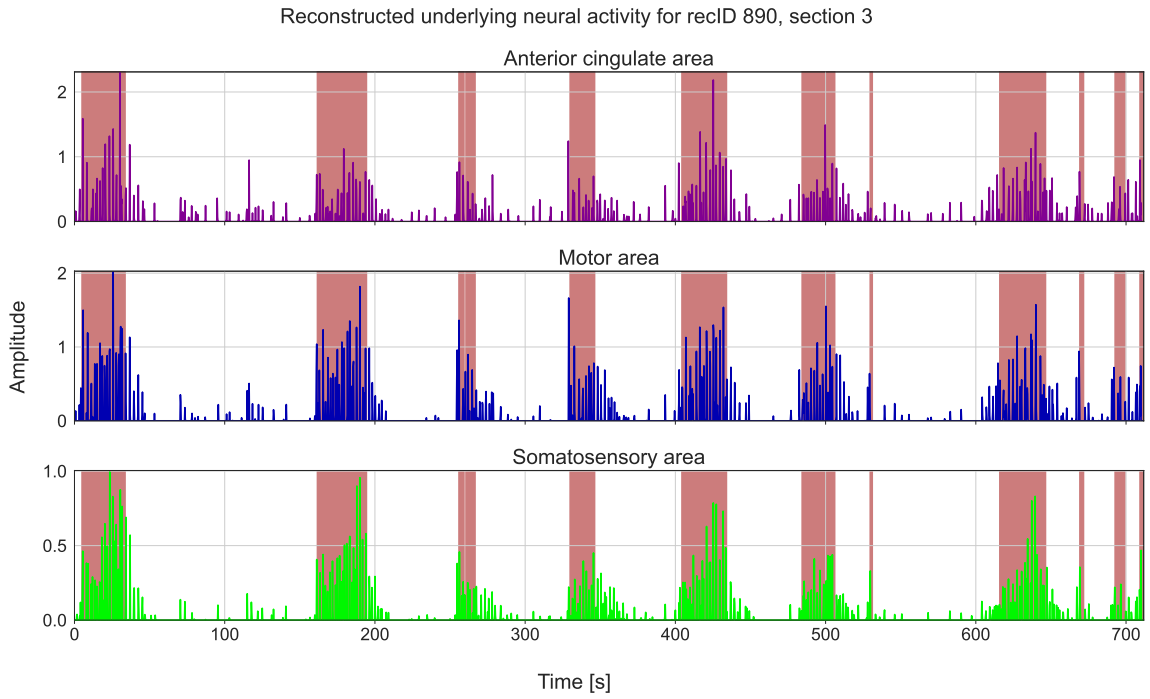


Figure 4.9: Deconvolved time courses. The movement of the mouse, extracted from the video recording, is depicted in salmon color.

4.3 HMM inference

Once the deconvolution procedure has been performed, functional networks and their dynamical properties must be learned from the deconvolved data. In this section, the inference methods accompanying the HMM are explained. The software for HMM inference of [32] is used.

The methods for the HMM are evaluation, decoding, and learning. The evaluation and decoding methods, described in Sec. 4.3.1 and Sec. 4.3.2, are used for solving the learning problem of Sec. 4.3.3.

Furthermore, it is convenient to introduce the following notation at this point. Having an fUS data recording, let the 'observation' $\mathbf{Y}$ denote the realization of the multivariate stochastic process Y as described in Sec. 3.2, where thus each column $\mathbf{y}_n \in \mathcal{R}^M$ of $\mathbf{Y}^T$ is a realization of Y at discrete time index $n = \{1, ..., N\}$. Also, let vector $\mathbf{z}$, containing the set of elements $\{z_1, ..., z_N\}$, denote the realization of the stochastic process Z.

4.3.1 Evaluation

In this section, the evaluation problem is discussed, following the structure of [33, 34]. The evaluation problem concerns the computation of the probability of an observed

sequence $\mathbf{Y}$, given the model parameters $\boldsymbol{\theta} = \{\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}\}$,

$$P(\mathbf{Y}|\boldsymbol{\theta}) = \sum_{\mathbf{z}} P(\mathbf{Y}, \mathbf{z}|\boldsymbol{\theta}) = \sum_{\mathbf{z}} P(z_1|\boldsymbol{\theta}) \prod_{n=1}^{N-1} P(z_{n+1}|z_n, \boldsymbol{\theta}) \prod_{n=1}^{N} P(\mathbf{y}_n|z_n, \boldsymbol{\theta}), \tag{4.16}$$

which requires the summation over all possible hidden state sequences $\mathbf{z}$, leading to K^N summations. Instead, the probability of an observed sequence can be computed recursively, via

$$P(\mathbf{Y}|\boldsymbol{\theta}) = \sum_{j} P(\mathbf{Y}, z_N = j|\boldsymbol{\theta}) = \sum_{j} \alpha_N^j, \tag{4.17}$$

where $j \in \mathcal{S}$, and the forward variable α_n^j describes the joint probability of partial observations up until discrete time n and state $z_n = j$, given the model $\boldsymbol{\theta}$. This leads to the discussion of the forward algorithm.

Forward algorithm

The forward variable is updated in a recursive fashion according to:

$$\alpha_{n+1}^j = \left[\sum_{i=1}^{K} \alpha_n^i a_{ij} \right] b_j(\mathbf{y}_{n+1}), \tag{4.18}$$

where $b_j(\mathbf{y}_{n+1})$ describes the emission probability of $\mathbf{y}_{n+1}$ in state j. The algorithm is initiated by

$$\alpha_1^i = \pi_i b_i(\mathbf{y}_1), \tag{4.19}$$

and finally terminated by

$$P(\mathbf{Y}|\boldsymbol{\theta}) = \sum_{i=1}^{K} \alpha_N^i. \tag{4.20}$$

For the derivation of Eq. 4.18, Eq. 4.19 and Eq. 4.20, the reader is referred to Sec. C.1.

4.3.2 Decoding

The second problem concerns the decoding problem. The goal is to compute the state that maximizes the posterior probability:

$$z_n = \arg\max_i P(z_n = i|\mathbf{Y}, \boldsymbol{\theta}). \tag{4.21}$$

By defining γ_n^i, the forward algorithm can be used to compute a part of the denominator of Eq. 4.22:

$$\gamma_n^i = P(z_n = i|\mathbf{Y}, \boldsymbol{\theta}) = \frac{P(\mathbf{Y}, z_n = i|\boldsymbol{\theta})}{P(\mathbf{Y}|\boldsymbol{\theta})}. \tag{4.22}$$

The computation of $P(\mathbf{Y}, z_n = i|\boldsymbol{\theta})$ can efficiently be performed by the forward-backward algorithm:

$$P(\mathbf{Y}, z_n = i|\boldsymbol{\theta}) = P(\mathbf{y}_1, ..., \mathbf{y}_n, z_n = i|\boldsymbol{\theta})P(\mathbf{y}_{n+1}, ..., \mathbf{y}_N|z_n = i, \boldsymbol{\theta}) = \alpha_n^i \beta_n^i, \tag{4.23}$$

where β_n^i is the backward variable, defined as:

$$\beta_n^i = P(\mathbf{y}_{n+1}, ..., \mathbf{y}_N|z_n = i, \boldsymbol{\theta}). \tag{4.24}$$

The backward algorithm characterizes itself by initialization of β_N^i:

$$\beta_N^i = 1, \quad \forall i \in \mathcal{S}, \tag{4.25}$$

and the updates, of which the proof can be found in Sec. C.2, are given by:

$$\beta_n^i = \sum_{j=1}^{N} \beta_{n+1}^j b_j(\mathbf{y}_{n+1}) a_{ij}. \tag{4.26}$$

Finally, writing Eq. 4.22 in full:

$$\gamma_n^i = \frac{P(\mathbf{Y}, z_n = i|\boldsymbol{\theta})}{P(\mathbf{Y}|\boldsymbol{\theta})} = \frac{\alpha_n^i \beta_n^i}{\sum_{i=1}^{N} \alpha_n^i \beta_n^i}, \tag{4.27}$$

where both the forward and backward algorithms can be used to efficiently calculate α_n^i and β_n^i. Then, solving Eq. 4.21 results in a sequence of states that are optimal on individual discrete time instances n. Thus, the expected number of correct states is maximized. However, the sequence consisting of individually most probable states may not be the most probable sequence. Therefore, the Viterbi algorithm has been developed, and this decoding algorithm is treated in the next section.

Viterbi decoding

In this section, the Viterbi algorithm is discussed, following the structure of [34]. In the Viterbi algorithm, the goal is to find the most probable state sequence, given data and the model parameters collected in $\boldsymbol{\theta}$:

$$\mathbf{z}^* = \arg\max_{\mathbf{z}} P(\mathbf{z}|\mathbf{Y}, \boldsymbol{\theta}), \tag{4.28}$$

and the highest probability along a single path, ending in state $j \in \mathcal{S}$ is computed by:

$$\delta_n^j = \max_{z_1, ..., z_{n-1}} P(z_1, ..., z_{n-1}, z_n = j, \mathbf{Y}|\boldsymbol{\theta}). \tag{4.29}$$

The procedure is as follows. First, an initialization step is applied:

$$\delta_1^j = \pi_j b_j(\mathbf{y}_1) \tag{4.30}$$

$$\psi_1^j = 0, \tag{4.31}$$

where ψ_n^j tracks the most probable path. Then, the following recursion is applied:

$$\delta_{n+1}^j = \max_{i \in \mathcal{S}} \delta_n^i a_{ij} b_j(\mathbf{y}_{n+1}) \tag{4.32}$$

$$\psi_{n+1}^j = \arg\max_{i \in \mathcal{S}} \delta_n^i a_{ij} b_j(\mathbf{y}_{n+1}). \tag{4.33}$$

Thus, ψ_{n+1}^j is the most probable next state j, and tracks the most probable path. Thereafter, the termination step is applied:

$$q^* = \max_{i \in \mathcal{S}} \delta_N^i \tag{4.34}$$

$$z_N^* = \arg\max_{i \in \mathcal{S}} \delta_N^i. \tag{4.35}$$

To obtain the Viterbi path, backtracking from the solution obtained in Eq. 4.35 has to be performed. This is illustrated in Eq. 4.36:

$$z_n^* = \psi_{n+1}(z_{n+1}^*), \quad n = N-1, N-2, ..., 1. \tag{4.36}$$

4.3.3 Learning

The final problem encountered is the learning problem. In the learning problem, the HMM parameters have to be learned from data. As seen before, an HMM is characterized by a set of parameters $\boldsymbol{\theta} = \{\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}\}$. The problem encountered is that having observations $\mathbf{Y}$, the hidden states $\mathbf{z}$ and parameters $\boldsymbol{\theta}$ are unknown. Thus maximizing $p(\mathbf{Y}, \mathbf{z}|\boldsymbol{\theta})$ is a difficult procedure, as it requires marginalization over the hidden states $\mathbf{z}$. Therefore, an iterative method is used called expectation maximization (EM). For the inference of HMM parameters, the Baum-Welch algorithm is commonly used and is an instance of the EM algorithm. It consists of the expectation step (E-step) and the maximization step (M-step). In the E-step, the following conditional expectation with respect to the hidden states $\mathbf{z}$ is computed [33], based on HMM parameters $\boldsymbol{\theta}_{old}$:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) = \mathbb{E}[\log(p(\mathbf{Y}, \mathbf{z}|\boldsymbol{\theta}))|\mathbf{Y}, \boldsymbol{\theta}_{old}] \tag{4.37}$$

Thus, in the E-step the parameters $\boldsymbol{\theta}_{old}$ are fixed and the expected state assignments are computed. Thereafter, the M-step is performed, in which the state assignments are fixed and the parameters are updated, according to:

$$\boldsymbol{\theta} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) \tag{4.38}$$

This iteration is repeatedly performed until a maximum number of iterations is achieved, or the model log-likelihood exceeds a certain threshold. It is important to note that the solution produced by the Baum-Welch algorithm is not guaranteed to converge to the global maximum; instead, it often converges to a local maximum.

E-step

In the E-step, the conditional expectation $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old})$ is computed such that subsequently in the M-step, the parameters $\boldsymbol{\theta}$ can be estimated relatively easy. For the computation of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old})$, the computation of two probabilities is required, as derived in Sec. C.3 [33]. First, γ_n^i, that is the probability of having $z_n = i$, given observations $\mathbf{Y}$, has to be computed. As encountered in Sec. 4.3.2, the forward-backward algorithm can be used to compute γ_n^i in Eq. 4.27. Secondly, the joint probability of the system being in state i at discrete time n and state j at discrete time $n+1$, given the observations $\mathbf{Y}$ has to be computed [34]:

$$\begin{aligned} \xi_n^{ij} = P(z_n = i, z_{n+1} = j | \mathbf{Y}, \boldsymbol{\theta}) &= \frac{P(z_n = i, z_{n+1} = j, \mathbf{Y} | \boldsymbol{\theta})}{P(\mathbf{Y} | \boldsymbol{\theta})} \\ &= \frac{\alpha_n^i a_{ij} b_j(\mathbf{y}_{n+1}) \beta_{n+1}^j}{\sum_{i=1}^K \alpha_n^i \beta_n^i}. \end{aligned} \tag{4.39}$$

Both γ_n^i and ξ_n^{ij} will play a crucial role in the maximization of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old})$, as discussed in next section.

M-step

In the M-step, the state assignments are fixed and the parameters in $\boldsymbol{\theta}$ are updated. The derivations of the update equations presented below can be found in Sec. C.3 [33]. First of all, the initial state probabilities $\boldsymbol{\pi} = \{\pi_i, ..., \pi_K\}$, that are the probabilities of being in state i at discrete time $n = 1$, are updated by:

$$\pi_i = \frac{\gamma_1^i}{\sum_{j=1}^K \gamma_1^j} \tag{4.40}$$

Then, the transition probabilities a_{ij}, that is the transition from state i to state j, are given by:

$$a_{ij} = \frac{\sum_{n=1}^{N-1} \xi_n^{ij}}{\sum_{n=1}^{N-1} \gamma_n^i} \tag{4.41}$$

For the Gaussian HMM, i.e. an HMM with state-specific Gaussian distributed emission distributions $f_{B_i}(\mathbf{y}_n) = \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ with $\mathbf{B} = \{B_i : i \in \mathcal{S}\}$ and $B_i = \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$, the parameter updates are as follows:

$$\hat{\boldsymbol{\mu}}_i = \frac{\sum_{n=1}^N \gamma_n^i \mathbf{y}_n}{\sum_{n=1}^N \gamma_n^i} \tag{4.42}$$

$$\hat{\boldsymbol{\Sigma}}_i = \frac{\sum_{n=1}^N \gamma_n^i (\mathbf{y}_n - \hat{\boldsymbol{\mu}}_i)(\mathbf{y}_n - \hat{\boldsymbol{\mu}}_i)^T}{\sum_{n=1}^N \gamma_n^i} \tag{4.43}$$

4.3.4 Practical EM implementation

The EM algorithm is initialized as follows. First, the observations are clustered using the K-means algorithm. From this clustering result, a one-hot encoding of the state sequence $\mathbf{z}$ is made. Subsequently, the state-specific observation distributions $f_{B_i}(\mathbf{y}_n)$ with $i \in \mathcal{S}$ are initialized by performing the M-step once. Furthermore, the initial probabilities $\boldsymbol{\pi}$ and state transition probability matrix $\mathbf{A}$ are uniformly initialized.

For the EM procedure, a total of 50 iterations are used to learn the model parameters, which turns out to be sufficient in terms of convergence of the model log-likelihood in practice. Finally, Viterbi decoding is applied to infer the most probable state sequence $\mathbf{z}$.

4.4 Evaluation metrics

It is important to define metrics that measure the temporal dynamics of functional networks, such that NRQ 2 can be answered.

NRQ 2. Is it possible to differentiate normal and mutated mice based on the temporal dynamics of functional networks?

For this purpose, four different metrics have been chosen to possibly reveal group differences between the mice, and these metrics are discussed below.

Inferred state transition probability matrices

The first and perhaps most straightforward parameter for evaluating group differences is the inferred state transition probability matrix $\mathbf{A}$. To analyze group differences, two HMM inference procedures can be applied to the subsets of the data, created based on the phenotype of the mice. The resulting difference between the inferred state transition probability matrices $\mathbf{A}_{WT}$ and $\mathbf{A}_{HOM}$ can be measured by computing the MSE:

$$MSE(\mathbf{A}_{WT}, \mathbf{A}_{HOM}) = \frac{1}{K^2}||\mathbf{A}_{WT} - \mathbf{A}_{HOM}||_2^2. \tag{4.44}$$

However, solely investigating this metric is not informative. To test whether an effect is measured, the mice phenotype labels are shuffled randomly, and the MSE is repeatedly computed. Then, by visualizing the distribution of the computed MSEs, it can be determined whether an actual effect is measured.

Fractional occupancy

The second metric is the fractional occupancy of states, measuring the fraction of time a state $k \in \mathcal{S} = \{1, ..., K\}$ is active:

$$\textit{fractional occupancy}_l(k) = \frac{1}{N}\sum_{n=1}^{N}(z_n = k), \tag{4.45}$$

where l specifies the specific mouse and N is the total discrete-time samples per mouse.

Mean state life time

The third metric is the mean state life time, that is the average duration of a state $z_n = k$.

Mean inter state time

The last metric used is the mean inter state time, measuring the elapsed time between two consecutive occurrences of the same state $z_n = k$.

5 Simulation

In this chapter, it is explained how synthetic data is generated and analyzed to support answering ERQ 3:

ERQ 3. How do different noise intensities influence the inference of temporal dynamics?

To be specific, the influence of noise on the state transition probability matrix is analyzed in this chapter. The advantage of having synthetic data is that certain parameters, such as the HRF parameters and the number of states K of the HMM, can be held equal to the true parameters to isolate the errors purely introduced by noise. For the purpose of analyzing the influence of noise, four different simulation cases with different state transition probability matrices have been created to measure the performance of the applied methods across different scenarios. For every simulation case, different noise intensities are tested. Furthermore, the performance of both NNLS and NNLASSO deconvolution methods are tested.

The layout of this chapter is as follows. First, the generation of synthetic data is discussed in Sec. 5.1. Subsequently, the deconvolution and HMM inference procedures, as described in Sec. 4.2 and Sec. 4.3, are applied to the synthetic data. The results of the analyses are found in Sec. 5.2. After that, a discussion on the results is held in Sec. 5.3

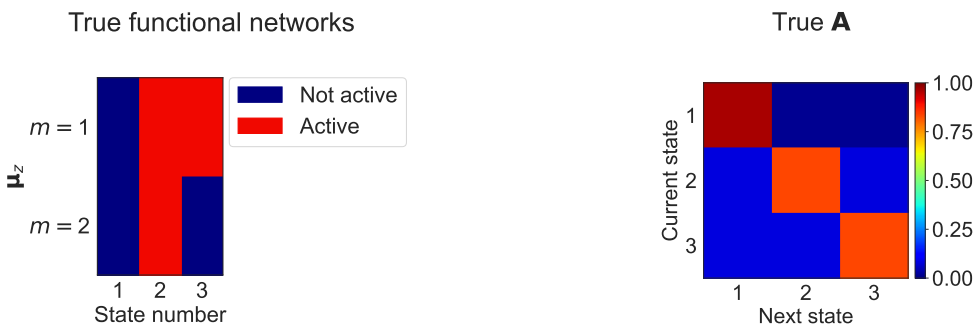
5.1 Synthetic data generation

Synthetic data is generated following the signal model of Ch. 3. Thus, synthetic data is generated through the convolution of binary impulse sequences, representing the underlying neural population activity, and the HRF. The impulse sequences are a multivariate binary sequence following a Markov process and thus a probability transition matrix. Also, each state has a specific binary activation pattern in which each of the ROIs is active or not. Furthermore, a state included in which no ROIs are active. Although it is unlikely that there is no activity in the entire brain, this situation is considered since only a 2D slice of the brain is imaged. For visual explanation purposes, the generation of synthetic data generation for the first simulation case, having sticky transitions and no off-diagonal transitional preferences, is shown in this section.

Binary impulse sequences

The binary impulse sequences are generated via three activity patterns, or functional networks, including a no-activity state, depicted in Fig. 5.1a. Thus, the number of

states is equal to $K = 3$. Those functional networks follow a state transition probability matrix $\mathbf{A}$, shown in Fig. 5.1b.



(a) (b)

Figure 5.1: **(a)** True functional networks. **(b)** True state transition probability matrix.

For the first simulation case, the binary impulse sequences, generated following the functional networks and their transition probabilities, are shown in Fig. 5.2.

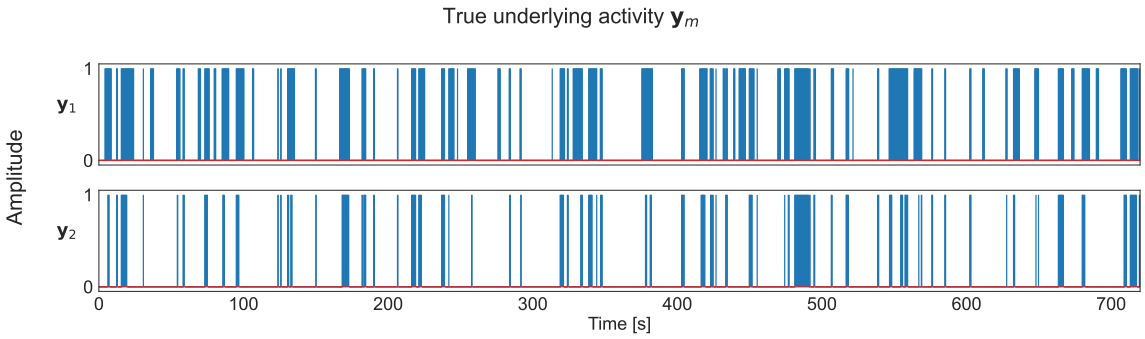


Figure 5.2: Binary impulse sequence representing underlying neural activity.

Synthetic fUS data

To create synthetic data, the binary impulse sequence and the HRF of Fig. 4.7 are convolved and Gaussian white noise is added to the clean fUS signals, according to the signal model described in Sec. 3.1. This results in the following synthetic fUS time courses for a noise intensity $\sigma = 0.1$, depicted in Fig. 5.3.

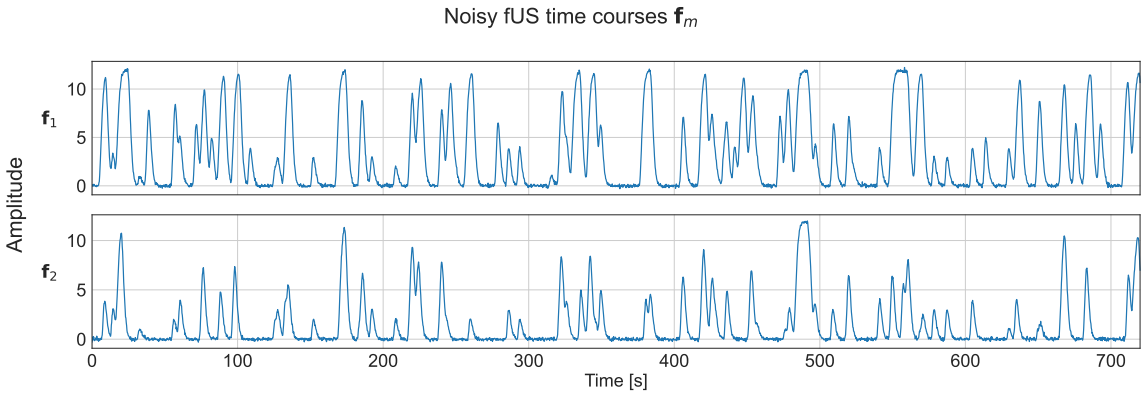


Figure 5.3: Noisy synthetic fUS data with noise intensity $\sigma = 0.1$

5.2 Results

In this section, the results of the synthetic data simulation are presented. After the methods described in Ch. 4 are applied to the synthetic data, a comparison between the NNLS and NNLASSO deconvolution methods is made. Subsequently, the influence of noise on the inference of the temporal dynamics is evaluated.

5.2.1 Comparison between NNLS and NNLASSO deconvolution methods

After generating synthetic data, a subsequent deconvolution procedure is applied on the time courses of each region m separately, using NNLS and the NNLASSO. For the first simulation case, this results in the following deconvolved time courses for NNLS, depicted in Fig. 5.4. For the NNLASSO deconvolution procedure, the regularization parameter is chosen such that the reconstruction error $RE(\lambda)$ per time course is not larger than 1%. The result of deconvolution using the NNLASSO can be found in Sec. A.2.

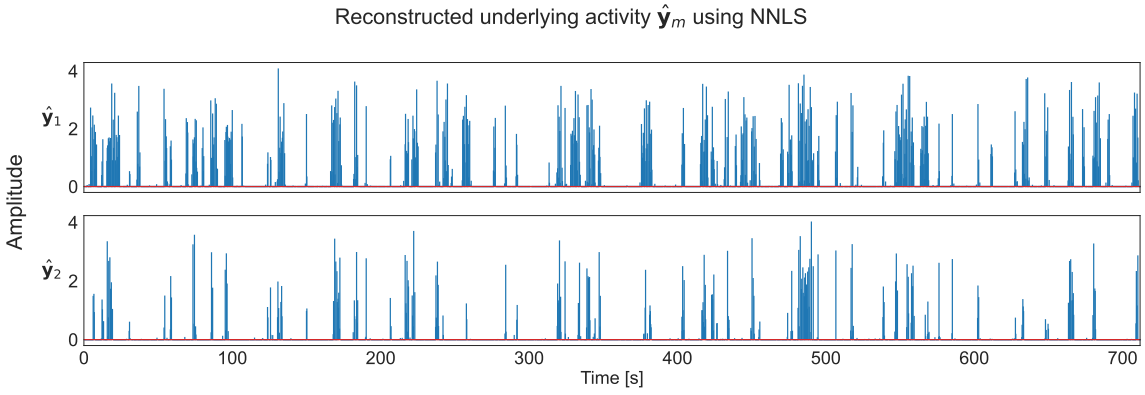


Figure 5.4: Deconvolved synthetic fUS data using NNLS for noise intensity $\sigma = 0.1$.

Subsequently, an HMM inference procedure is applied to the deconvolved data. For the first simulation case, using NNLS as deconvolution method, the true state sequence and the most probable state sequence obtained using Viterbi decoding are shown in Fig. 5.5:

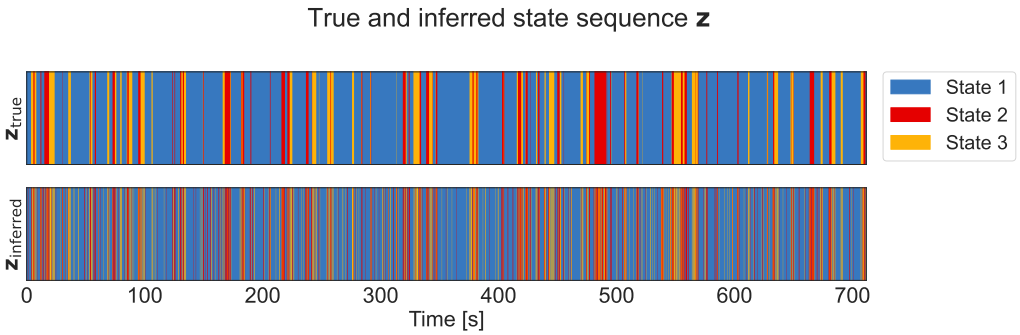


Figure 5.5: True and inferred state sequence after Viterbi decoding, using NNLS deconvolution for time courses with noise intensity $\sigma = 0.1$.

It can be noticed that the states with one or more active regions are inferred relatively often compared to the true state sequence. It is observed that using NNLS as a deconvolution method in combination with the HMM leads to a significant amount of spurious activations. Now, using the fact that these 'observations' $\mathbf{y}_n$ are two-dimensional, the correct and inferred state classifications are shown in Fig. 5.6.

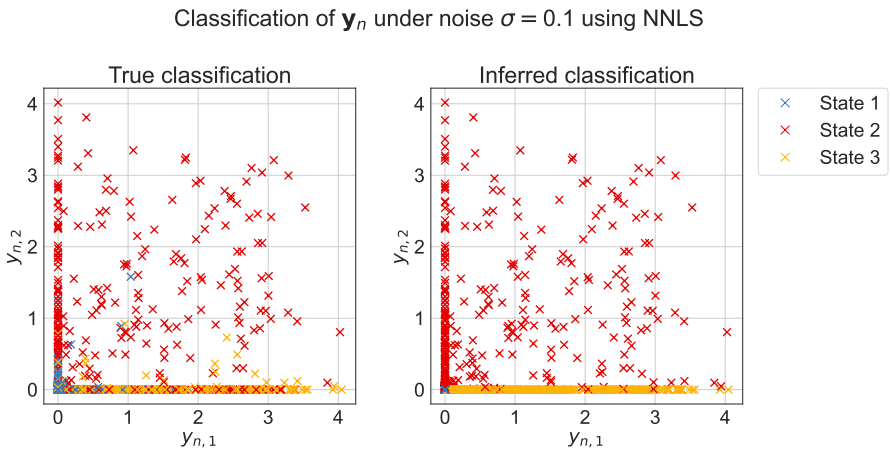


Figure 5.6: True and inferred state classification using NNLS for noise intensity $\sigma = 0.1$.

Here, a significant result can be found regarding the HMM model assumptions and inference procedure. It can be observed that spurious activations are not classified in the correct cluster, mainly due to the Gaussian HMM being not able to deal with truncated distributions. The likelihood of a 'spurious' sample under the actual functional network distribution is small, as this 'spurious' sample has a less sparse activity pattern than the functional network that generated the sample, leading to the wrong classification

of the sample. However, the opposite case is observed, where samples with a more sparse activity pattern than the functional network generating the sample can be classified correctly if such a wrong transition is unlikely under the inferred state transition probability matrix. Thus, a classification bias is present toward more densely activated functional networks. On the positive side, although the Gaussian HMM cannot deal with truncated distributions, it might still be useful, as the clusters already overlap significantly, and classifying such samples seems a rather challenging task. Also, it can be noted that the functional networks are inferred correctly.

Now, performing an NNLASSO deconvolution procedure on the same time courses and subsequently applying Viterbi decoding, the most probable state sequence is obtained and visualized in Fig. 5.7:

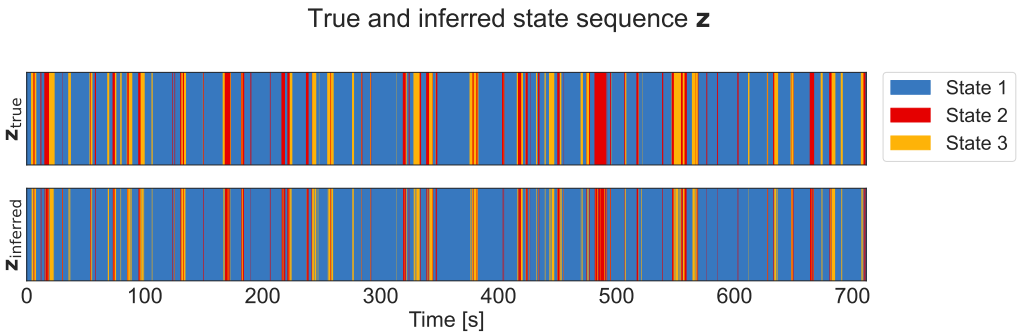


Figure 5.7: True and inferred state sequence using the NNLASSO for noise intensity $\sigma = 0.1$.

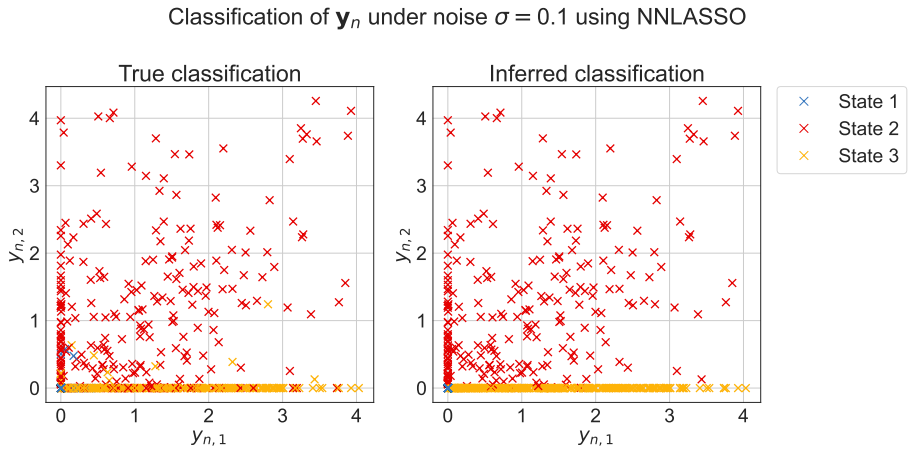


Figure 5.8: True and inferred state classification using the NNLASSO for noise intensity $\sigma = 0.1$.

where it can be noted that the inferred state sequence $\mathbf{z}_{\text{inferred}}$ matches the true state sequence $\mathbf{z}_{\text{true}}$ better compared to using NNLS. This similarity is mainly due to the first state (blue) being selected more often due to the sparse regularization, which suppresses spurious activations and better matches the ground truth. This can also be observed

on the left of Fig. 5.8, where more data points $\mathbf{y}_n$ are now approximately on location (0,0).

Finally, the resulting reconstruction of the temporal dynamics captured by $\mathbf{A}$ using NNLS or the NNLASSO is considered. The results of learning matrix $\mathbf{A}$ from deconvolved time courses with noise intensity $\sigma = 0.1$ using NNLS and the NNLASSO are shown in Fig. 5.9 and Fig. 5.10, respectively.

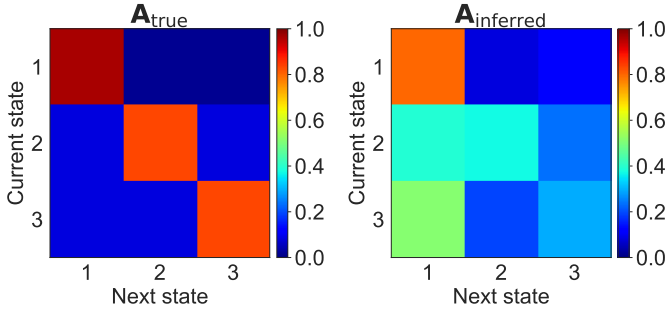


Figure 5.9: True and inferred state transition probability matrix using NNLS for noise intensity $\sigma = 0.1$.

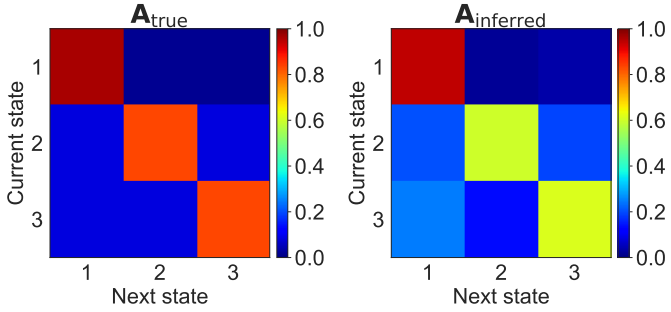


Figure 5.10: True and inferred state transition probability matrix using the NNLASSO for noise intensity $\sigma = 0.1$.

From visual observation, it is clear that NNLS is unable to recover the structure of the temporal dynamics. On the other hand, using the NNLASSO, the dynamics are recovered relatively accurately, maintaining a large part of the structure in matrix $\mathbf{A}$. However, the first significant difference can be observed in the state self-transition probability, also called the state stickiness, which is significantly reduced. In the deconvolved time courses, it is observed that reconstruction errors are made during long periods of sustained activity, resulting in less state stickiness. Also, a slight bias can be noticed from state 2 to state 3. This bias can be explained by considering the functional network pattern of state 2, where a single error in one of the two deconvolved time courses leads to a self-transition or a transition to state 3.

5.2.2 Noise influence

As a next step, the influence of noise is further tested by measuring the ability to recover the underlying dynamics of the functional networks, expressed in the state transition probability matrix $\mathbf{A}$. A total of four simulation cases have been constructed, covering a wide range of possible brain dynamics and noise intensities $\sigma \in \{0.1, 0.3, 0.5\}$, where the noise intensity of experimental fUS data is considered to vary between $\sigma = 0.3$ and $\sigma = 0.5$. The transition probability matrices of the four simulation cases are constructed such that the resulting time courses resemble a realistic fUS time course, listed in Sec. A.1 for $\sigma = 0.5$, where the functional networks remain the same as before. For all results shown in the sections below, a total of 8 chunks consisting of 720 seconds of deconvolved data are concatenated for robust inference of matrix $\mathbf{A}$ under noise intensity $\sigma = 0.5$. The convergence of the inferred matrix $\mathbf{A}$ to the true matrix $\mathbf{A}$ is evaluated in Sec. A.3 for varying synthetic data lengths. In the remainder of this chapter, only results will be presented that use the NNLASSO deconvolution method, as this technique outperforms NNLS in terms of reconstruction of $\mathbf{A}$.

Sticky states, no transitional preferences

The first simulation case, already shown before, has a state transition probability matrix $\mathbf{A}$ that has no clear structure, i.e., a state transition probability matrix without off-diagonal transition probabilities. This simulation case is constructed to show whether no structure is learned from data if there truly is no transitional structure in the underlying activity. Furthermore, it should be noted that the state transition probabilities are rather sticky, i.e., a high preference for self-transitions. The result of inferring matrix $\mathbf{A}$ is shown in Fig. 5.11.

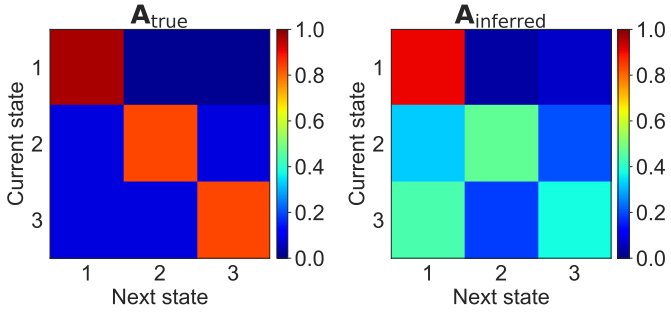


Figure 5.11: Inference of $\mathbf{A}$ with sticky states under noise intensity $\sigma = 0.5$.

It can be noted that a significant transition bias is present toward state 1, due to reconstruction errors introduced during the deconvolution procedure.

Sticky states, transitional preference towards sparse activation patterns

The second simulation case concerns a state transition probability matrix $\mathbf{A}$ having a transitional preference from state 2 to state 3. It can be observed that the inferred $\mathbf{A}$

does maintain the structure by a small portion for noise intensity $\sigma = 0.5$.

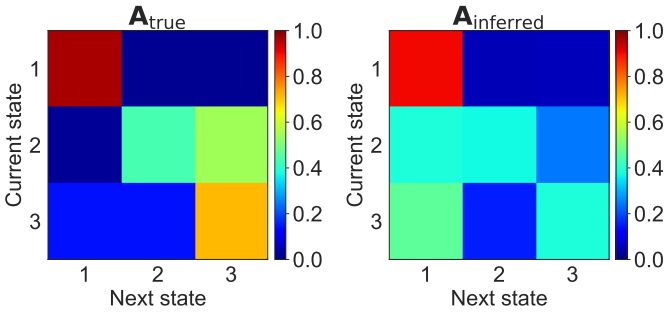


Figure 5.12: Inference of $\mathbf{A}$ with transitional preference under noise intensity $\sigma = 0.5$.

Sticky states, transitional preference towards dense activation patterns

As a third simulation, the reverse situation of the previous simulation has been created. Now, a preference to transition from state 3 to state 2 is present. Also, in this case, the inferred $\mathbf{A}$ only partly maintains the structure between the states having at least one region active. However, it should be noted that the state transition from state 3 to state 2 is more likely than the reverse, matching ground truth.

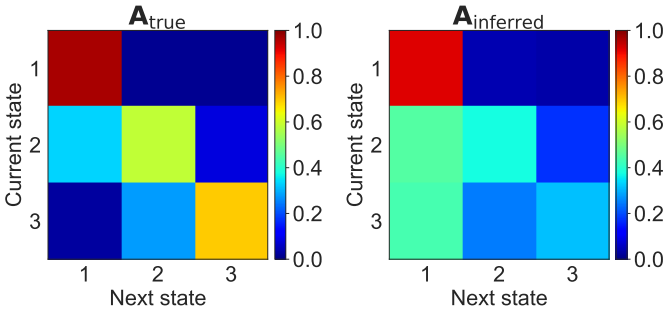


Figure 5.13: Inference of $\mathbf{A}$ with transitional preferences under noise intensity $\sigma = 0.5$.

No sticky states, transitional preference towards sparse activation patterns

Finally, a situation is tested with less state stickiness, of which the result of inferring $\mathbf{A}$ for $\sigma = 0.5$ can be found in Fig. 5.14. It becomes apparent that the developed techniques cannot restore any structure of $\mathbf{A}_{\text{true}}$. First of all, having non-sticky state self-transitions, and thus a more sparse activation pattern, leads to an increase in errors introduced by inferring the exact activation timing. Especially for larger noise intensities, estimating the exact timing appears to be a difficult task.

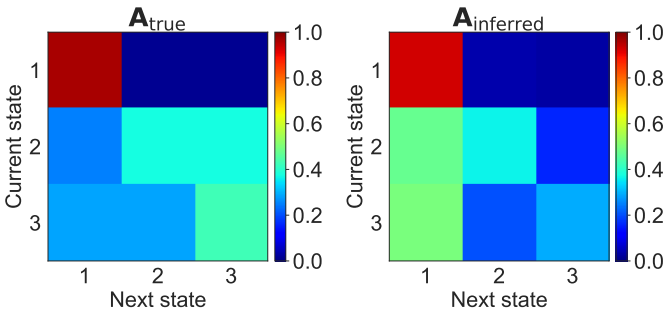


Figure 5.14: Inference of $\mathbf{A}$ without sticky states under noise intensity $\sigma = 0.5$.

5.3 Discussion

First of all, from the first simulation case, it becomes clear that performing a deconvolution using the NNLASSO outperforms NNLS. However, since sparsity is enforced in the reconstruction of $\mathbf{y}_m$, very sticky self-transitions are not favored. This effect is also clearly visible in the results of all simulation cases, where the diagonal elements of the inferred matrix $\mathbf{A}$ have a lower probability value. The introduced errors by deconvolution mainly consist of finding a solution that is too sparse and timing the exact moment of activation. Setting the regularization parameter λ at a reconstruction error of 1% is a trade-off between detecting genuine and spurious activations. Therefore, a less sparse solution effectively results in deconvolution using NNLS, which performs significantly worse. Furthermore, estimating the exact timing of underlying activity is an issue and manifests itself in not inferring the correct temporal dynamics of functional networks.

On the positive side, it is found empirically that the method can infer the actual functional networks relatively often for all simulation cases, knowing the true HRF and number of states K. Also, the overall activation pattern of the inferred state sequence $\mathbf{z}$ seems to match the ground truth closely from visual observations. Thus, the method for setting the regularization parameter, as described in Sec. 4.2, seems to be a good trade-off between detecting spurious and true activations. However a solid analysis on choosing the regularization parameter is lacking. Furthermore, for more sticky state transition probability matrices, the introduced state transition preferences are still visible, despite the errors introduced in the state transition probability matrix. For lower noise intensities, improved results are found for reconstructing transition preferences. However, it appears that no clear structure can be learned for less sticky state self-transitions due to the difficulty of inferring the exact timing of underlying activity in noisy conditions.

Although significant errors are introduced, a relative comparison of the inferred state transition probability matrices of each group of mice could still disclose group differences for the experimental data analysis. However, care must be taken when interpreting the learned state transition probabilities, as the state transition probability matrix is affected by reconstruction errors.

6 Results

In this chapter, the results of the experimental data analysis are presented. After applying a deconvolution procedure using the NNLASSO on each time course, a few challenges must be tackled. First, the number of states K is determined for applying the HMM inference methods. Secondly, the consistency of functional network inference is evaluated. Finally, differences between groups of mice are analyzed. These challenges are solved by starting at a high-level analysis, where all deconvolved time courses of the two groups of mice are concatenated. Subsequently, an HMM inference procedure is applied. Here, it is thus assumed that the functional networks are similar in the two groups of mice. The difference between the two groups of mice, measured by metrics such as the fractional occupancy, mean state life time, and mean inter-state time, is evaluated in the high-level analysis to avoid the possibility of differences introduced by local minima of the EM algorithm.

After the high-level analysis, a group-level analysis is performed by training two HMMs on the concatenated data of the group of wild-type (WT) and homozygous (HOM) mice, respectively. It is verified whether functional networks are altered between the two groups of mice. Also, the temporal dynamics of functional networks are evaluated by investigating the inferred state transition probability matrices.

As a final step, applying a separate HMM inference procedure on each mouse would be desirable. In the case of consistent functional networks, the dynamics of functional networks can be evaluated at a subject level, and the consistency of these dynamics can be investigated. However, it is found that the amount of data points per mouse is too few for the HMM inference procedure to learn consistent functional networks reliably.

6.1 High-level analysis

Before diving into the results of learning functional networks and their temporal dynamics, it is crucial to analyze global differences between the two groups of mice. For this purpose, it is analyzed whether the hyperactivity is translated into more movement by computing the fraction of time the mice are moving. Also, it is analyzed what fraction of time any brain region is active and whether there are group differences in the amount of global brain activity. As visible in Fig. 6.1, the two groups of mice cannot be separated based on their movement or global brain activity. As no apparent global differences are found, it is necessary to decompose the brain activity into networks of functional connectivity by applying an HMM with $K > 2$.

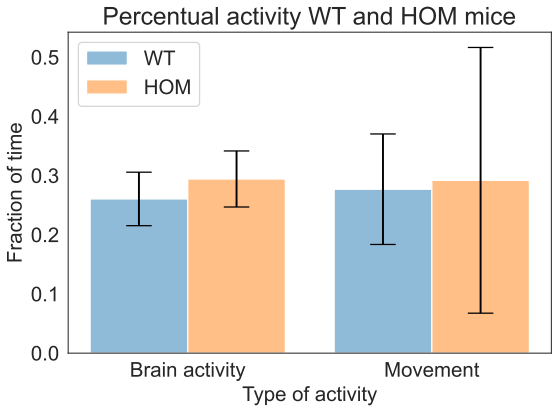


Figure 6.1: Final time courses after the pre-processing stage.

6.1.1 Determination of the number of states

In order to determine the number of states K, the uniqueness of functional networks, differences in the mean fractional occupancy of states per group of mice, and the biological plausibility of functional networks are investigated. One of the outcomes of the HMM inference procedure is a set of state vectors, where each state vector $\boldsymbol{\mu}_z$ indicates whether a certain ROI is active or not. The uniqueness of these state vectors, or functional networks, is evaluated against the number of inferred states. In Fig. 6.2, the number of unique functional networks is thus plotted against different values for K, and it is visible that only a unique decomposition of activity can be made up to $K = 5$.

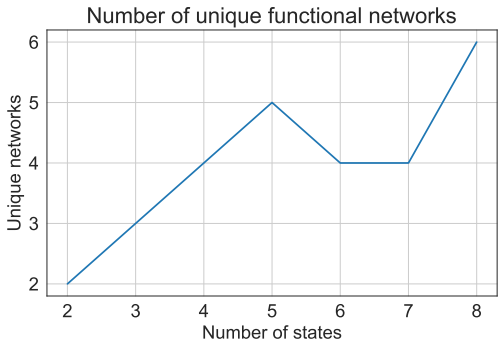


Figure 6.2: Number of unique functional networks versus the number of states.

Introducing a larger number of states, $K > 2$, leads to a subdivision or decomposition of the activity patterns in the brain, with thus a maximum of five unique functional networks. For these varying numbers of states, the fractional occupancy of the two groups is evaluated. It is found that no clear distinction can be made using this specific metric.

Moving on to the biological plausibility of the functional network decomposition, for $K = 4$, an interesting set of functional networks is found. By plotting the inferred mean $\boldsymbol{\mu}_z$ of the multivariate Gaussian distributions, the functional networks are visualized in Fig. 6.3. It is shown that brain activity consists of brain states having activity in all three regions, both the motor and somatosensory area, or only the anterior cingulate area. This finding is in correspondence with literature [24], where it is also demonstrated that neural activity in the anterior cingulate area does not always follow neural activation of the motor or somatosensory area. In contrast, activity in either the motor or somatosensory area is not independently observed. This finding motivates the choice of $K = 4$ for further analyses.

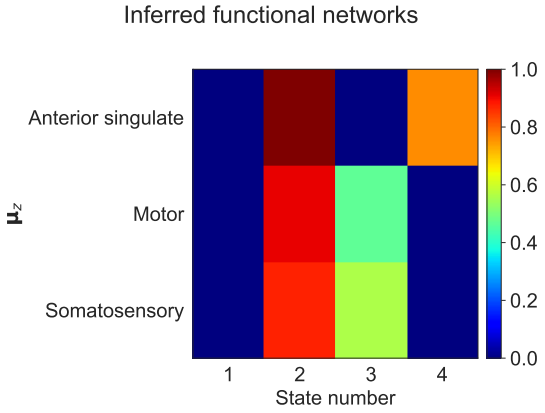


Figure 6.3: Visualization of inferred functional networks for $K = 4$.

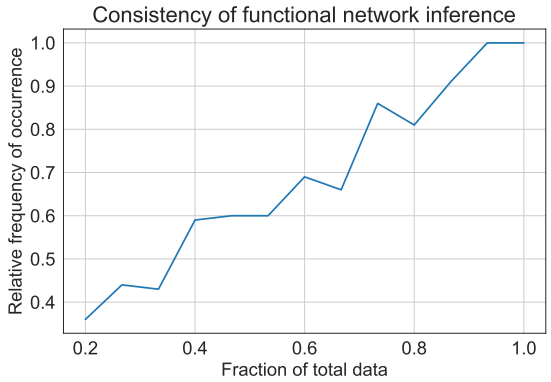


Figure 6.4: Consistency of inferred functional networks for $K = 4$.

6.1.2 Consistency of functional network inference

In order to evaluate the consistency of the found functional network decomposition for $K = 4$, the data is subsampled. Subsampling of the data is performed by leaving out some recordings. Subsequently, the functional networks for $K = 4$ are inferred from the reduced data set. By reducing data size and performing 100 Monte Carlo runs per number of recordings that are left out, the consistency of functional network inference is analyzed by computing the relative frequency of occurrence of the functional network decomposition of Fig. 6.3. As visible in Fig. 6.4, even for half the amount of data, the functional network decomposition of Fig. 6.3 is inferred the most times. Also, it can be noted that for the full data length, different K-means initializations of the EM algorithm do not affect functional network inference.

6.1.3 Comparison of WT and HOM mice

To compare the temporal dynamics of WT and HOM mice, the fractional occupancy, mean state life time, and mean inter-state time, as described in Sec. 4.4, are computed. The mean fractional occupancy of the groups of mice is shown in Fig. 6.5.

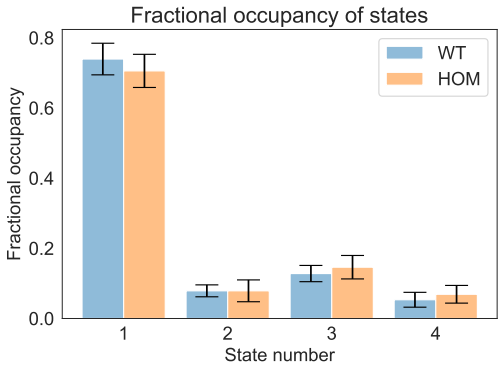
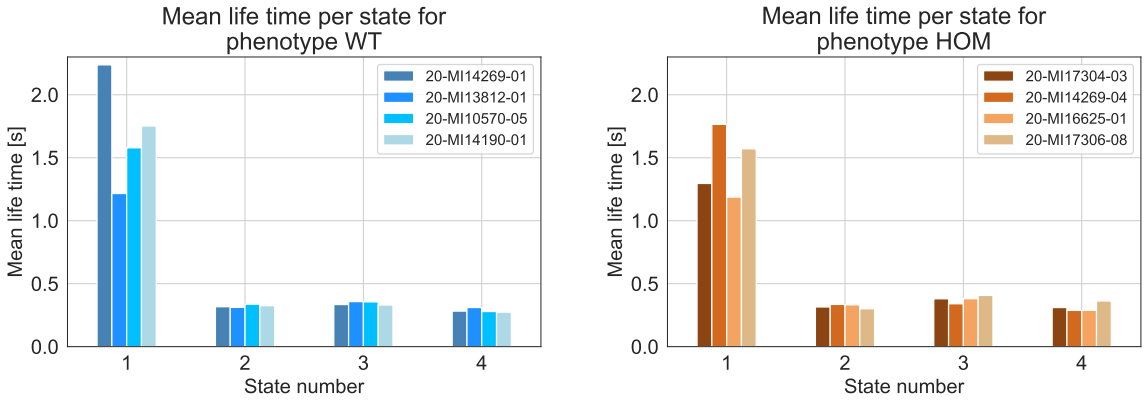


Figure 6.5: Fractional occupancy of states.

Although small group differences in mean fractional occupancy can be noticed, the WT and HOM mice cannot be separated due to the individual differences in fractional occupancy. The fractional occupancy per mouse can be found in Fig. B.4.

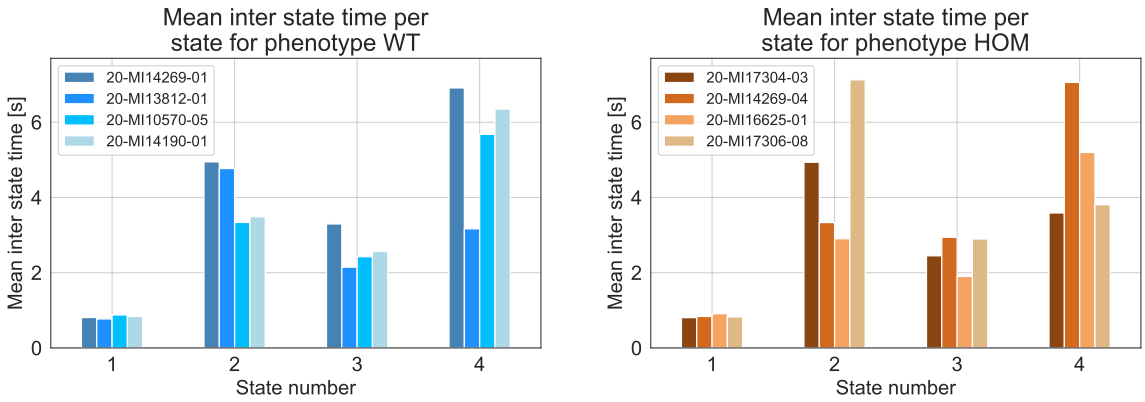
Now, the mean state life time is computed for each mouse, visualized in Fig. 6.6. Again, no clear differences are found, except for a small difference in the state life time of the first state, corresponding to no brain activity. Furthermore, it can be seen that the mean inter-state times of states containing brain activity are rather low, and intra-group differences are also relatively small. The sparse regularization potentially introduces this effect, not favoring state stickiness, also observed in the simulations in Ch. 5.



(a) (b)

Figure 6.6: Mean state life time for mice of group **(a)** WT **(b)** HOM. Legends contain labels of animals in correspondence with Tab. B.1.

As the last metric, the mean inter-state time is computed per mouse, as illustrated in Fig. 6.7. Also, in this case, the individual differences are significant, and overlap between the groups of mice is present.

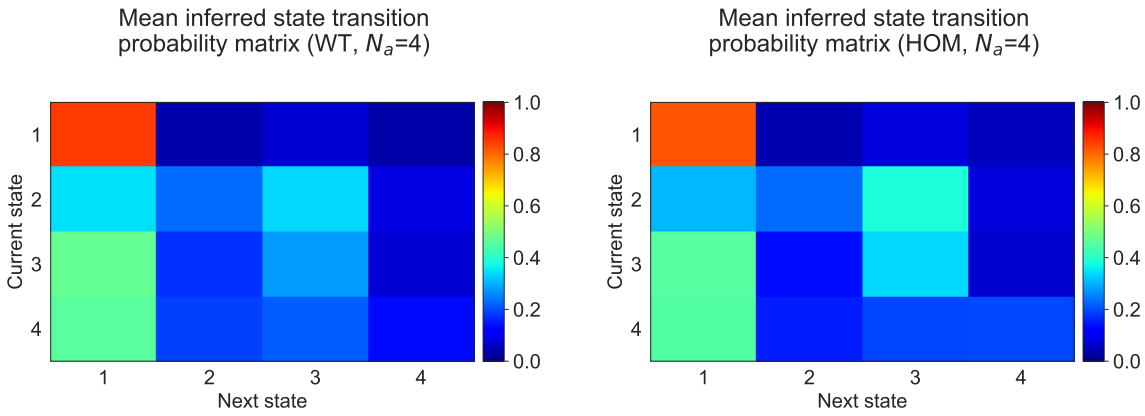


(a) (b)

Figure 6.7: Mean inter state time for mice of group **(a)** WT **(b)** HOM. Legends contain labels of animals in correspondence with Tab. B.1.

6.2 Group-level analysis

Now, possible differences between the two groups of mice are analyzed on a group level by training two HMMs using $K = 4$ on the concatenated data of the group of normal and mutated mice, respectively. It is found that the functional networks are not altered and consistently learned, also for multiple initializations of the EM algorithm, as visible in Sec. B.5. While having these consistent functional networks, the state transition probability matrices of the WT and HOM mice are displayed in Fig. 6.8, where these matrices are computed on $N_a = 4$ animals per group. It can be observed that the dynamics of the groups of mice are different regarding the transition probabilities to state 3, i.e., the state in which both the motor and somatosensory are active.

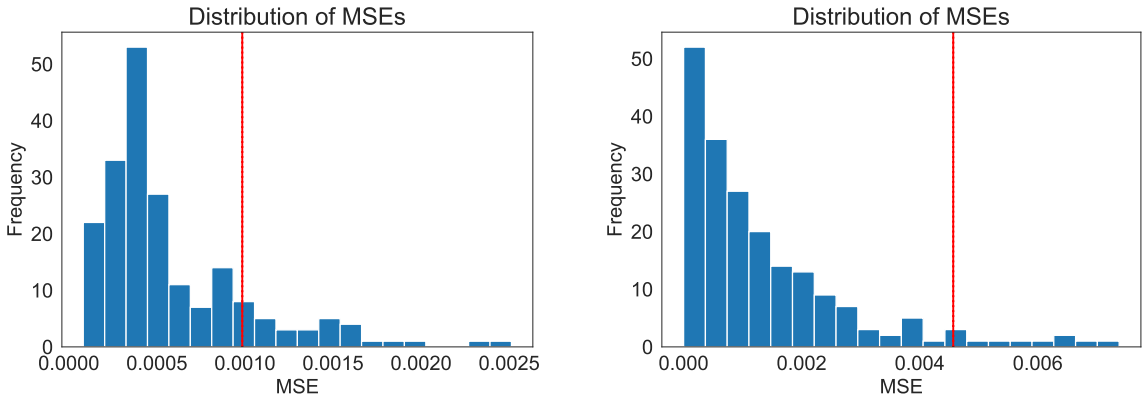


(a) (b)

Figure 6.8: State transition probability matrices per group of mice: **(a)** WT **(b)** HOM.

To analyze whether a genuine effect is measured, the labels of recordings are interchanged, as described in Sec. 4.4. Subsequently, the $MSE(\mathbf{A}_{WT}, \mathbf{A}_{HOM})$ is computed between the inferred state transition probability matrices $\mathbf{A}_{WT}$ and $\mathbf{A}_{HOM}$ of the WT and HOM mice, respectively. Performing 200 Monte Carlo runs, the result of the measured effect versus the distribution of random effects is illustrated in Fig. 6.9a.

Now, it should be noted that the transition probability matrix measures two quantities simultaneously. First, the tendency to and the sustainment of any brain activation are measured by the transition probabilities in the first row and column, respectively. Second, the transitional structure between states that contain any activity is measured. As previously illustrated in Fig. 6.1, the WT and HOM cannot be separated based on the amount of brain activity. Therefore, the first row and column are excluded from the state transition probability matrix. As there are also no clear effects on the dynamics of the fourth state, the fourth state is also excluded. For this analysis of the partial state transition probability matrices, the transition probabilities are recalculated, and a similar procedure as before is followed. The results are shown in Fig. 6.9b.



(a) (b)

Figure 6.9: Distribution of MSEs on the **(a)** full and **(b)** partial transition probability matrices computed from random phenotype assignments (blue) versus the true partial MSE (solid red). Stability of the true MSE is evaluated by showing 95% confidence intervals (stipled red) established through performing 200 Monte Carlo runs on the initialization of the EM algorithm.

As can be seen, it cannot be fully established whether the observed difference in the state transition probability matrix is a coincidence or not. However, with $\sim$ 96% of the random groups of recordings resulting in lower differences, it can be claimed that a genuine effect is measured with the risk of 4% of being wrong. Performing analyses with a restriction on the minimum number of interchanged phenotypes leads to similar results for both approaches. Also, performing similar analyses interchanging the labels of mice instead of recordings lead to similar results, although it should be noted that fewer random combinations are possible.

7 Discussion

In this chapter, a discussion on the results is held. It is evaluated how the results answer the engineering and neuroscientific research questions in Sec. 7.1 and Sec. 7.2, respectively.

7.1 Engineering reseach questions

Inference of functional networks and temporal dynamics

Concerning the inference of functional networks and temporal dynamics of these functional networks, it is found that a signal model is required to model the data from a generative perspective. The signal model is composed of two components, namely a convolutive fUS data model and a state-space model, being the HMM. After constructing this signal model, it is found that methods such as deconvolution using the NNLASSO and EM are necessary to learn networks of functional connectivity from fUS data. For the deconvolution procedure, also NNLS is considered in combination with the HMM inference procedure, which leads to a significant increase in inferred spurious activations.

Pre-processing

Before directly applying the methods to experimental data, a pre-processing stage is designed. The pre-processing stage reduces problem complexity and performs data preparation for the developed methods.

As there is significant variability in the blood vessel structure of mice, a dynamic functional connectivity analysis on very small neural populations is not considered possible in a reliable manner. Thus, the choice is made to consider the activity of neural populations at a larger scale, namely on the level of anatomical regions. Since fUS data only contains blood vessels, and thus no anatomical regions are visible, these regions have to be detected. Using a combination of sICA and warping the Paxinos brain atlas on the blood vessels, anatomical regions in fUS data can be detected successfully. However, this method requires expert knowledge for placing landmarks used in the brain atlas warping stage, which also appeared to be difficult for anatomical regions, such as the thalamus, located deeper in the brain. Next, a data dimensionality reduction step is performed. Here, by computing the mean fUS time course of each anatomical region, it is found that enough detail is maintained, and problem complexity is successfully reduced.

Concerning data preparation, it is found that fUS data requires a global noise reduction step. The developed method can significantly reduce the observed scattering artifacts. However, the assumption is made that the global noise is homogeneous across

the 2D fUS images, which did not hold for all mice. After the earlier described data dimensionality reduction step, drift correction is applied by performing linear regression on the baseline activity of a mean fUS time course. Finally, the possible offset in the mean fUS time course can quite easily be removed. It should be noticed that the performance and selected thresholds of all involved pre-processing methods are subject to data interpretation. However, using video recordings showing the behavior of mice as prior knowledge, it is believed that reliable thresholds are established.

Noise influence

The influence of noise is analyzed by performing several simulation cases with increasing levels of advanced state transition probability matrices. Concerning the inference of the state transition probability matrix under noisy conditions, it is found that the general structure of the matrix is affected by a combination of timing and sparsity errors, which are introduced by the deconvolution procedure. The true structure appeared to be less recoverable for increasing noise intensities, although, for relatively large state-stickiness, the signature of the true structure is still visible. Furthermore, it must be noted that it is hard to quantify the noise intensity present in experimental data due to varying signal power.

Consistency

The functional networks for the three identified ROIs and $K = 4$ states are inferred in a relatively consistent manner within the two groups of mice. Across several subsampled data sets, the demonstrated decomposition of brain activity is inferred the most often. However, the functional networks for four regions, including the thalamus, could not be consistently inferred. This inconsistency is mainly due to a significant data loss, as the thalamus is not identified in half of the total number of recordings. Furthermore, due to a lack of data per mouse, the reproducibility of functional networks and consistency of inference in each mouse could not be evaluated. Also, it is found that during the simulations the method can relatively consistently infer the actual functional networks that generated the synthetic fUS data.

7.2 Neuroscientific reseach questions

Biologic plausibility

The consistent decomposition of brain activity into a functional network comprising the motor and somatosensory area, and a functional network in which only the anterior cingulate area is active, is consistent with literature [24]. This finding shows that the developed methods can reveal functional networks in a relatively consistent manner and that fUS does indirectly measure the activity of underlying neural populations. Also, in the work of [24, 25], it is found that neural activity related to behavior is spread through the entire brain. This result is consistent with the identification of a brain state in which all regions are active. However, it is impossible to verify if this state is identified purely based on the actual underlying activity of neural populations

or whether the global blood supply in the brain also induces an all-active brain state. Also, it must be noted that it is unknown and hard to verify whether any artefacts contribute to the reconstructed neural activity in the mice brain.

Differentiation of mice

A possible difference in brain dynamics between WT and HOM mice is demonstrated using the developed method. The choice is made to use metrics that measure the temporal dynamics of functional networks independently from the amount of movement, as several types of behaviors could invoke activity in neural populations. First of all, no differentiation can be made between the groups of mice using the fractional occupancy and mean inter-state time, as there are significant intra-group differences. These differences are likely induced by the variability of spontaneous behaviors a mouse can perform in combination with a relatively short recording time. The variability in brain activation due to spontaneous behavior might thus cause the inability to separate classes using these metrics. Also, the mean state life time seems to be affected by the combination of noise and sparse regularization, as expected from simulation results. However, it is unknown if the brain states truly last as short as reconstructed, even during long periods of sustained movement, as the ground truth is unknown.

Concerning the state transition probability matrices, a transitional bias towards the functional network containing the motor and somatosensory area is observed in HOM mice. By creating groups with randomly assigned recordings, the significance of the observed transitional effect is measured. Although a smaller effect is measured for $\sim$ 96% of the random combinations, it cannot be fully established whether the observed effect is a coincidence or not. However, the effect is statistically significant enough to claim a difference in brain dynamics, with the risk of 4% being wrong. As a few random combinations result in a larger difference, an overlap between the groups of mice in brain dynamics is expected. Performing the Monte Carlo analysis on the measured effect using phenotype groups of randomly assigned animals is, in principle, preferred over recording-wise randomization. However, significantly fewer random combinations can be made with a few animals, and a robust statistical analysis cannot be performed. Also, the group overlap effects cannot be mitigated by inserting a limit on the minimum number of interchanged phenotypes at this low number of animals and recordings.

8 Conclusion

First of all, it is concluded that, by the use of a signal model, a method consisting of a deconvolution procedure in combination with a state-space inference approach has successfully been constructed for dynamic functional connectivity analysis on fUS data. For the deconvolution procedure, the NNLASSO is deployed to suppress spurious activations. As an unsupervised machine learning method to learn functional networks and their temporal dynamics, EM for HMMs is leveraged.

In order to reduce problem complexity, the fUS data is interpreted on the level of anatomical regions. The anatomical regions are successfully identified using a combination of sICA and warping a brain atlas on top of the fUS blood vessels. It is established that this reduction of complexity still preserves enough detail to discover biologically plausible functional networks and their dynamics. For data preparation, first, a global noise reduction step involving recording-dependent thresholds. Also, region-wise thresholds are involved in the drift correction step, dependent on the variability of brain activity. It is found that, although these thresholds are necessary for the developed methods to function, generalization across several recordings and larger data sets is expected to be difficult and must be executed carefully.

From synthetic data simulations, it can be concluded that the developed methods struggle under noisy conditions, as reconstructing the exact timing of the activity of neural populations is more complicated. Also, due to sparse regularization, state stickiness is less preserved. However, transitional dynamics captured by the state transition probability matrices are still inferred, although the impact of noise is significantly present.

After all, an interesting decomposition of brain activity into biologically plausible functional networks is found in correspondence with literature. These functional networks are inferred relatively consistent across different subsampled fUS data sets. Also, a transitional bias towards the functional network comprising the motor and somatosensory area is observed in HOM mice. This effect can be related to a difference in mice phenotype with 96% certainty.

8.1 Future work

Pre-processing pipeline

Several improvements can be made to establish a more robust pre-processing pipeline. These improvements are firstly concerned with improving data quality. Tracing the origin of the observed scattering effects and extremely slow drifts and subsequently mitigating their presence would eventually result in a robust pipeline with fewer thresholds.

However, besides improving data quality, the currently used methods can also be improved. Several thresholds are present in the pre-processing pipeline, leading to much analytic flexibility in the pipeline, which is not desired. Where most of the thresholds could be generalized across different recordings, the threshold for global noise reduction, although dynamic already, requires a manual determination per recording. A different approach using wavelets could improve the global noise reduction step by a more accurate determination of the global noise present in the fUS data.

Furthermore, as previously discussed, the atlas warping procedure is a labor-intensive and challenging procedure, as the warping is performed using manually identified landmarks. However, these landmarks are not clear for deeper located regions such as the thalamus, resulting in less accurate brain atlas warping and anatomical region detection. In order to improve, a joint MRI and fUS setup could be considered to more clearly map the non-linear deformation of the mice brain and subsequently identify blood vessels surrounding anatomical regions of interest.

Experimental data analysis

First, more measurements are desired for a more robust statistical experimental data analysis. In order to reduce intra-group differences, more measurements per mouse should be performed. Also, more mice should be added to the analysis to reveal more robust inter-group differences. Furthermore, gathering more data would enable the ability to include the thalamus in the analysis.

Second, the detection of mouse behavior could be expanded with automated identification of, e.g., whisker movement, grooming, and other behaviors. Improved behavior tracing would subsequently enable the improvement of metrics evaluating the temporal dynamics of functional networks, leading to more specific analyses of the temporal dynamics with respect to behavioral treats. For now, the metrics could already be expanded by accounting for the movement of the mice.

Furthermore, the influence of selecting the regularization parameter could be analyzed in further detail. The chosen method to select the regularization parameter, such that the relative reconstruction error does not increase by more than 1%, is a trade-off between recovering genuine and spurious activations. Although the ground truth is unknown, it should be analyzed how varying this threshold affects the final results. Apart from sparse regularization on the underlying activity directly, sparse regularization on the first-order derivative of the underlying activity could be considered as well, resulting in more piece-wise constant solutions.

Finally, it should be analyzed how variability in the HRF influences the finally inferred results. To start with, this could be analyzed by means of a simulation.

Bibliography

[1] A. E. Urai, B. Doiron, A. M. Leifer, and A. K. Churchland, "Large-scale neural recordings call for new insights to link brain and behavior," pp. 11–19, 1 2022.

[2] C. J. Stoodley, A. M. D'Mello, J. Ellegood, V. Jakkamsetti, P. Liu, M. B. Nebel, J. M. Gibson, E. Kelly, F. Meng, C. A. Cano, J. M. Pascual, S. H. Mostofsky, J. P. Lerch, and P. T. Tsai, "Altered cerebellar connectivity in autism and cerebellar-mediated rescue of autism-related behaviors in mice," *Nature Neuroscience*, vol. 20, no. 12, pp. 1744–1751, 12 2017.

[3] V. Zerbi, M. Pagani, M. Markicevic, M. Matteoli, D. Pozzi, M. Fagiolini, Y. Bozzi, A. Galbusera, M. L. Scattoni, G. Provenzano, A. Banerjee, F. Helmchen, M. A. Basson, J. Ellegood, J. P. Lerch, M. Rudin, A. Gozzi, and N. Wenderoth, "Brain mapping across 16 autism mouse models reveals a spectrum of functional connectivity subtypes," *Molecular Psychiatry*, vol. 26, no. 12, pp. 7610–7620, 12 2021.

[4] R. Botvinik-Nezer, F. Holzmeister, C. F. Camerer, and et al., "Variability in the analysis of a single neuroimaging dataset by many teams," *Nature*, 2020. [Online]. Available: https://doi.org/10.18112/openneuro.

[5] C. M. Bennett, A. A. Baird, M. B. Miller, and G. L. Wolford, "Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction," in *Human Brain Mapping*, 2009.

[6] E. Mace, G. Montaldo, B. F. Osmanski, I. Cohen, M. Fink, and M. Tanter, "Functional ultrasound imaging of the brain: Theory and basic principles," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 60, no. 3, pp. 492–506, 2013.

[7] H. Berger, "Über das Elektrenkephalogramm des Menschen," *Archiv für Psychiatrie*, p. 25, 4 1929.

[8] S. M. Osovets, D. A. Ginzburg, V. S. Gurfinkel', L. P. Zenkov, L. P. Latash, V. B. Malkin, P. V. Mel'nichuk, and E. B. Pasternak, "Electrical activity of the brain: Mechanisms and interpretation," *American Institute of Physics*, pp. 103–150, 1983.

[9] C. Iadecola, "The Neurovascular Unit Coming of Age: A Journey through Neurovascular Coupling in Health and Disease," pp. 17–42, 9 2017.

[10] C. Huneau, H. Benali, and H. Chabriat, "Investigating human neurovascular coupling using functional neuroimaging: A critical review of dynamic models," *Frontiers in Neuroscience*, vol. 9, no. DEC, 2015.

[11] H. S. Wei, H. Kang, I. Y. D. Rasheed, S. Zhou, N. Lou, A. Gershteyn, E. D. McConnell, Y. Wang, K. E. Richardson, A. F. Palmer, C. Xu, J. Wan, and M. Nedergaard, "Erythrocytes Are Oxygen-Sensing Regulators of the Cerebral Microcirculation," *Neuron*, vol. 91, no. 4, pp. 851–862, 8 2016.

[12] T. Deffieux, C. Demené, and M. Tanter, "Functional Ultrasound Imaging: A New Imaging Modality for Neuroscience," pp. 110–121, 10 2021.

[13] E. Mace, G. Montaldo, I. Cohen, M. Baulac, M. Fink, and M. Tanter, "Functional ultrasound imaging of the brain," *Nature Methods*, vol. 8, no. 8, pp. 662–664, 8 2011.

[14] C. Rabut, M. Correia, V. Finel, S. Pezet, M. Pernot, T. Deffieux, and M. Tanter, "4D functional ultrasound imaging of whole-brain activity in rodents," *Nature Methods*, vol. 16, no. 10, pp. 994–997, 10 2019.

[15] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, "Dynamic functional connectivity: Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360–378, 10 2013.

[16] X. Liu and J. H. Duyn, "Time-varying functional network information extracted from brief instances of spontaneous brain activity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 11, pp. 4392–4397, 3 2013.

[17] X. Di and B. B. Biswal, "Dynamic brain functional connectivity modulated by resting-state networks," *Brain Structure and Function*, vol. 220, no. 1, pp. 37–46, 1 2015.

[18] P. Tewarie, L. Liuzzi, G. C. O'Neill, A. J. Quinn, A. Griffa, M. W. Woolrich, C. J. Stam, A. Hillebrand, and M. J. Brookes, "Tracking dynamic brain networks using high temporal resolution MEG measures of functional connectivity," *NeuroImage*, vol. 200, pp. 38–50, 10 2019.

[19] S. Keilholz, C. Caballero-Gaudes, P. Bandettini, G. Deco, and V. Calhoun, "Time-Resolved Resting-State Functional Magnetic Resonance Imaging Analysis: Current Status, Challenges, and New Directions," pp. 465–481, 10 2017.

[20] F. I. Karahanoglu and D. Van De Ville, "Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks," *Nature Communications*, vol. 6, 7 2015.

[21] V. D. Calhoun, J. Liu, and T. Adali, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," *NeuroImage*, vol. 45, no. 1 Suppl, 2009.

[22] E. Uruñuela, T. A. W. Bolton, D. Van De Ville, and C. Caballero-Gaudes, "Hemodynamic Deconvolution Demystified: Sparsity-Driven Regularization at Work," 7 2021. [Online]. Available: http://arxiv.org/abs/2107.12026

[23] X. Liu, C. Chang, and J. H. Duyn, "Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns," *Frontiers in Systems Neuroscience*, vol. 7, no. DEC, 12 2013.

[24] N. A. Steinmetz, P. Zatka-Haas, M. Carandini, and K. D. Harris, "Distributed coding of choice, action and engagement across the mouse brain," *Nature*, vol. 576, no. 7786, pp. 266–273, 12 2019.

[25] C. Stringer, M. Pachitariu, N. Steinmetz, C. B. Reddy, M. Carandini, and K. D. Harris, "Spontaneous behaviors drive multidimensional, brainwide activity," *Science*, vol. 364, no. 6437, 2019.

[26] D. Vidaurre, S. M. Smith, and M. W. Woolrich, "Brain network dynamics are hierarchically organized in time," *Proceedings of the National Academy of Sciences*, vol. 114, no. 48, pp. 12 827–12 832, 11 2017. [Online]. Available: https://www.pnas.org/content/114/48/12827

[27] H. Eavani, T. D. Satterthwaite, R. E. Gur, R. C. Gur, and C. Davatzikos, "Unsupervised Learning of Functional Network Dynamics in Resting State fMRI," *Inf Process Med Imaging*, pp. 426–437, 2013.

[28] A. Mittal, S. Linderman, J. Paisley, and P. Sajda, "Bayesian recurrent state space model for rs-fMRI," *Machine Learning for Health*, 11 2020. [Online]. Available: http://arxiv.org/abs/2011.07365

[29] A. O. Nunez-Elizalde, M. Krumin, C. B. Reddy, G. Montaldo, A. Urban, K. D. Harris, and M. Carandini, "Neural correlates of blood flow measured by ultrasound," *Neuron*, 3 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0896627322001775

[30] A. Hyvärinen, J. Karhunen, and E. Oja, "Independent Component Analysis," Tech. Rep., 2001.

[31] A. K. Aydin, W. D. Haselden, Y. Goulam Houssen, C. Pouzat, R. L. Rungta, C. Demené, M. Tanter, P. J. Drew, S. Charpak, and D. Boido, "Transfer functions linking neural calcium to single voxel functional ultrasound signal," *Nature Communications*, vol. 11, no. 1, 12 2020.

[32] S. Linderman, B. Antin, D. Zoltowski, and J. Glaser, "SSM: Bayesian Learning and Inference for State Space Models," 2020. [Online]. Available: https://github.com/lindermanlab/ssm

[33] J. Miller, "Hidden Markov Models: Lecture Notes on Advanced Stochastic Modeling," Duke University, Durham, Tech. Rep., 2016.

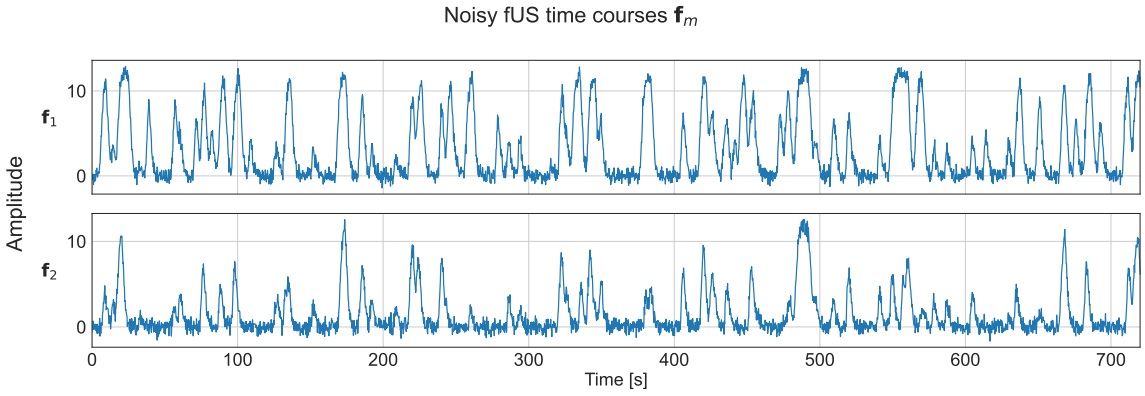
[34] J. Daniel and J. H. Martin, "Speech and Language Processing A.1 Markov Chains," Tech. Rep., 2021.

Synthetic data analysis

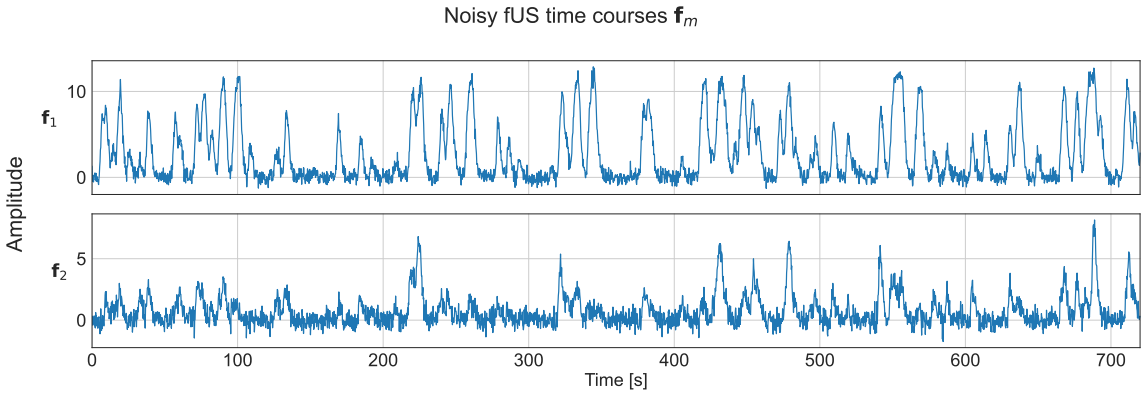
A

In this chapter, supplemental material on the synthetic data analysis is presented. For each simulation case, the noisy fUS time courses are shown for $\sigma = 0.5$. Furthermore, deconvolution using the NNLASSO for synthetic data with noise intensity $\sigma = 0.1$ is shown. Also, the convergence of inferring $\mathbf{A}$ is illustrated by means of a simulation with varying length of the concatenated data.

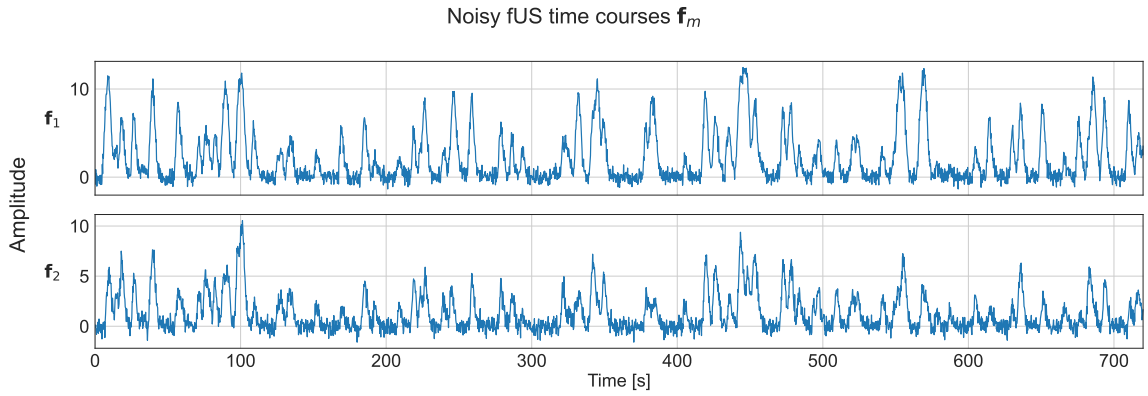
A.1 Synthetic noisy fUS time courses



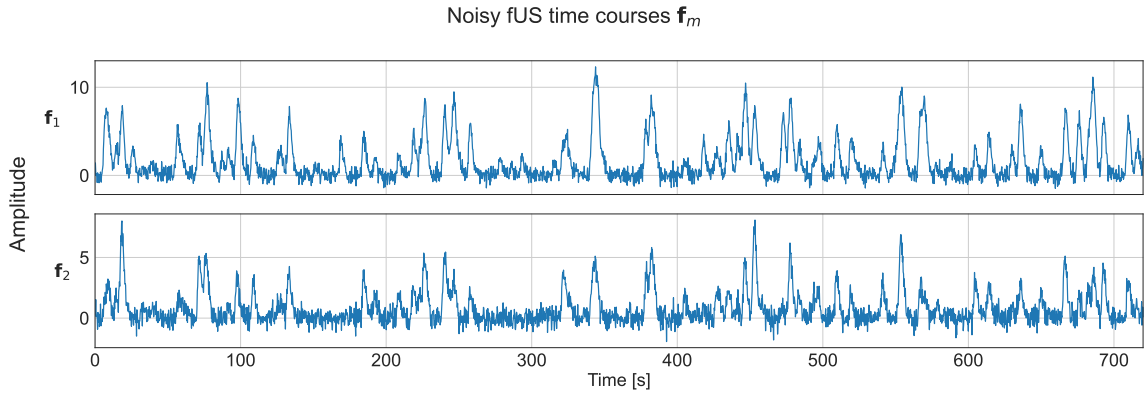
(a)



(b)



(c)



(d)

Figure A.1: Noisy synthetic fUS data with noise intensity $\sigma = 0.5$. **(a)** Sticky states, no transitional preferences. **(b)** Sticky states, transitional preference towards sparse activation patterns. **(c)** No sticky states, transitional preference towards sparse activation patterns. **(d)** No sticky states, transitional preference towards sparse activation patterns.

A.2 Deconvolution of noisy synthetic fUS data

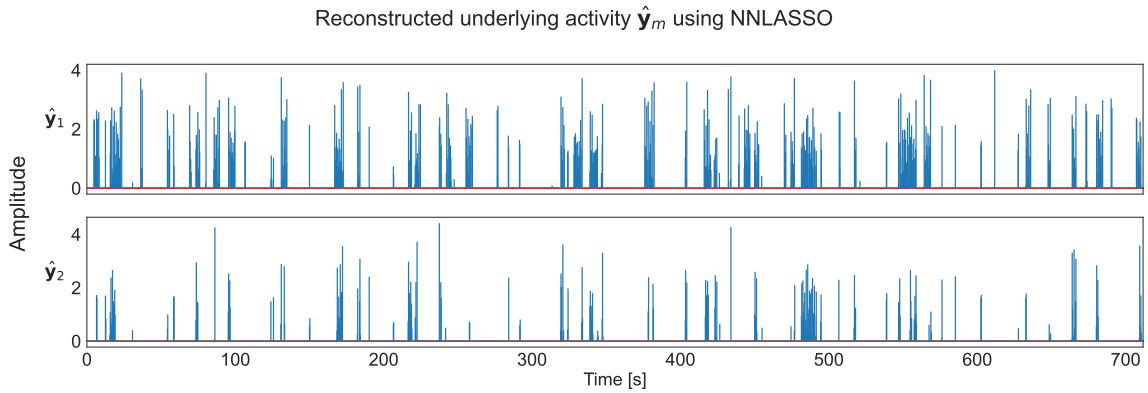


Figure A.2: Deconvolved synthetic fUS data using the NNLASSO for noise intensity $\sigma = 0.1$.

A.3 Convergence of inferring A

To determine the consistency of inferring $\mathbf{A}$, the convergence of the MSE is evaluated in Fig. A.3a for varying lengths of concatenated deconvolved data, using a recording duration, equal to the experimental data, of 720 seconds. The reconstruction error between the true and inferred state transition probability matrix $\mathbf{A}$ is measured by the MSE:

$$MSE = \frac{1}{K^2}||\mathbf{A}_{\text{true}} - \mathbf{A}_{\text{inferred}}||_F^2. \tag{A.1}$$

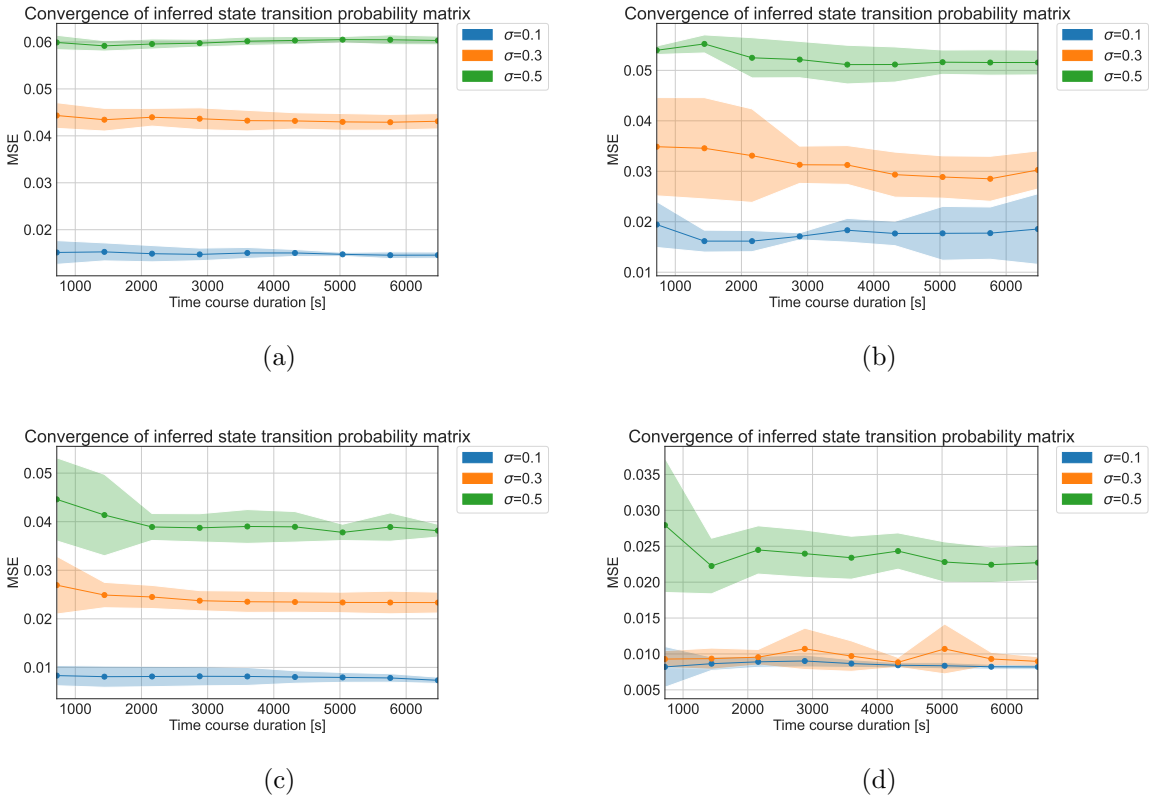


Figure A.3: Convergence of the MSE. For each noise intensity σ, 5 Monte Carlo runs are performed from which 95% confidence intervals are constructed. **(a)** Sticky states, no transitional preferences. **(b)** Sticky states, transitional preference towards sparse activation patterns. **(c)** No sticky states, transitional preference towards sparse activation patterns. **(d)** No sticky states, transitional preference towards sparse activation patterns.

Experimental data analysis

B

In this appendix, all additional information and figures of the experimental data analysis can be found.

B.1 Recording table

Table B.1: fUS recordings

recID	sectionID	animalID	phenotype
830	3	20-MI13812-01	WT
890	3	20-MI13812-01	WT
1002	3	20-MI14269-04	HOM
1181	3	20-MI10570-05	WT
1033	3	20-MI14190-01	WT
1124	3	20-MI16625-01	HOM
959	4	20-MI14269-01	WT
959	6	20-MI14269-01	WT
1102	1	20-MI17304-03	HOM
1105	1	20-MI17304-03	HOM
1181	1	20-MI10570-05	WT
1124	1	20-MI16625-01	HOM
1184	1	20-MI17306-08	HOM
1033	2	20-MI14190-01	WT
1181	23	20-MI10570-05	WT

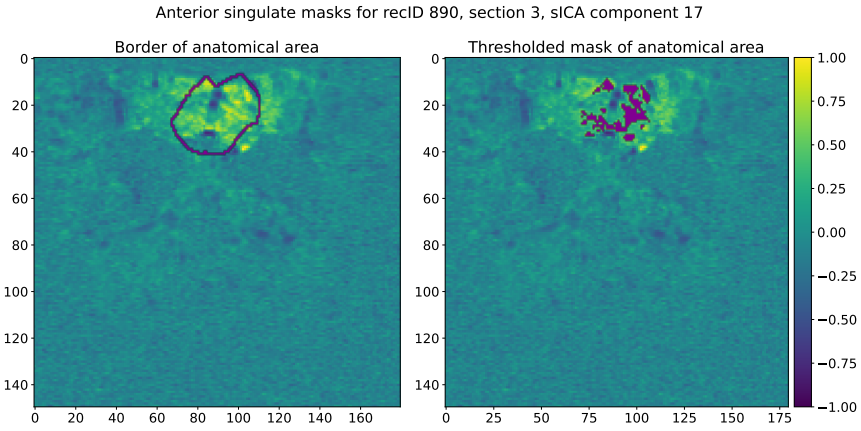
B.2 Intrinsic activity detection thresholds

In Tab. B.2 below, the relative tresholds are listed for detecting data points corresponding to intrinsic activity, that should not be included in the linear regression. For the anterior cingulate area, a slightly higher threshold is chosen, since the signal-to-noise ratio is lower for anatomical regions that are located deeper in the brain.

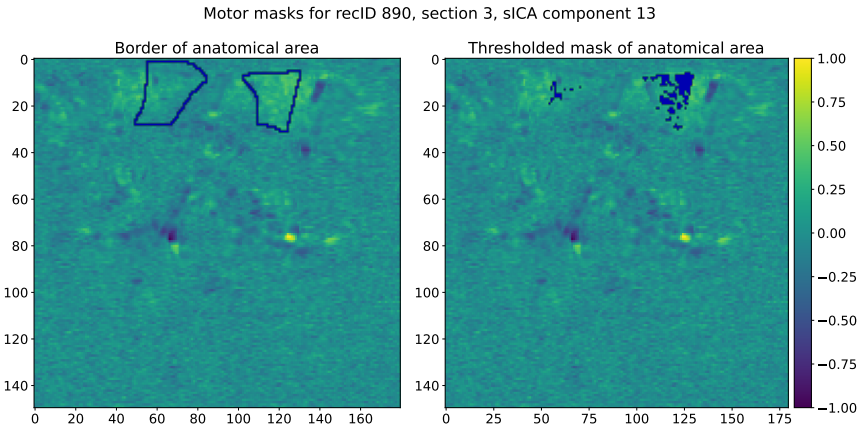
Table B.2: Relative thresholds per region

Region	Anterior cingulate	Motor	Somatosensory
Threshold [%]	5	3	3

B.3 Detection of other anatomical ROIs



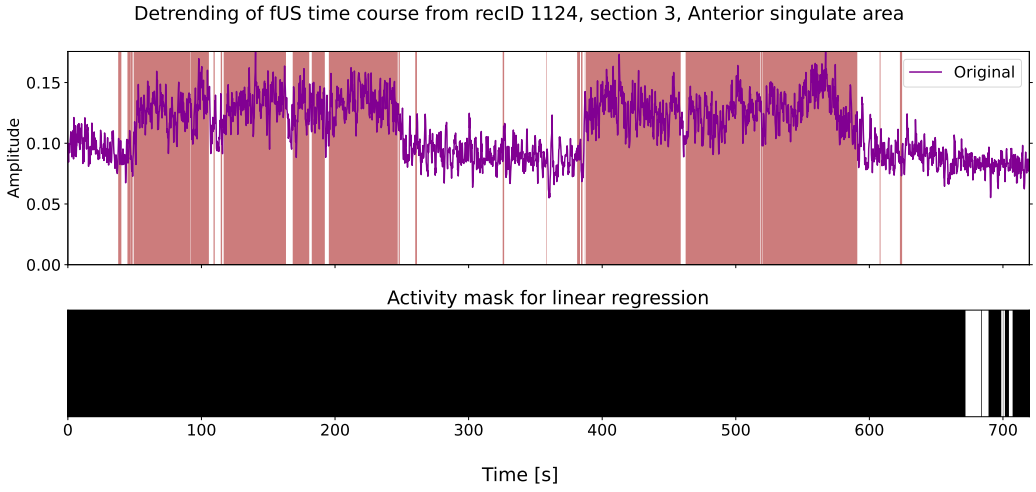
(a)



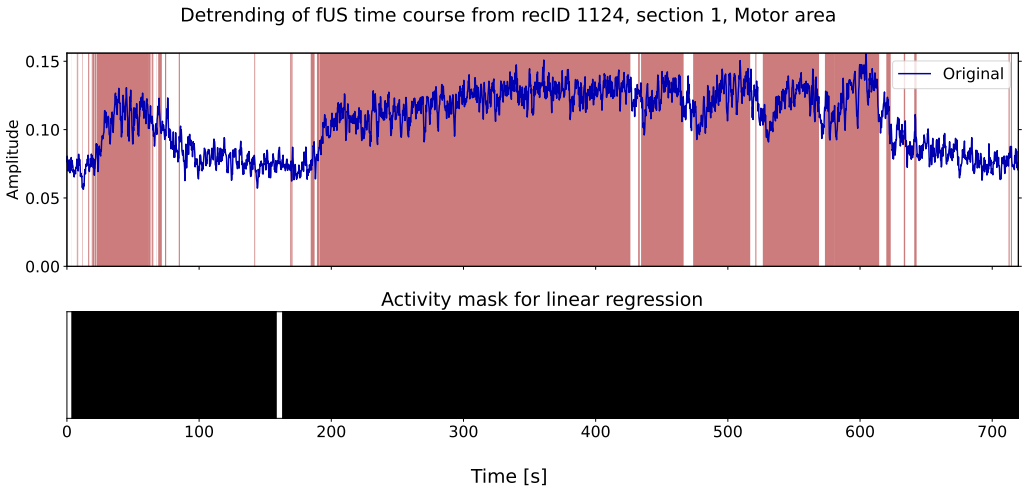
(b)

Figure B.1: ROI detection of **(a)** anterior cingulate area and **(b)** motor area. **Left:** Border of regional mask. **Right:** Thresholded pixels within regional mask.

B.4 Unsuccessful detrending of time courses



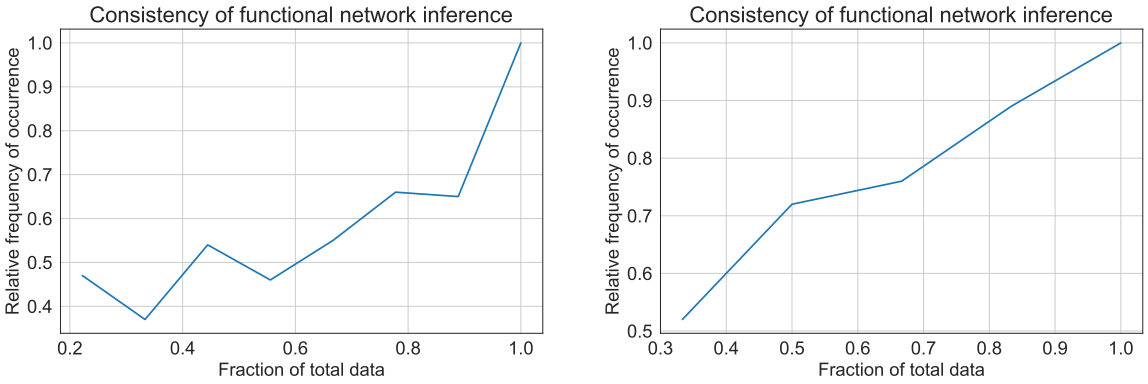
(a)



(b)

Figure B.2: Unsuccessful detrending of time course belonging to **(a)** anterior cingulate area **(b)** motor area.

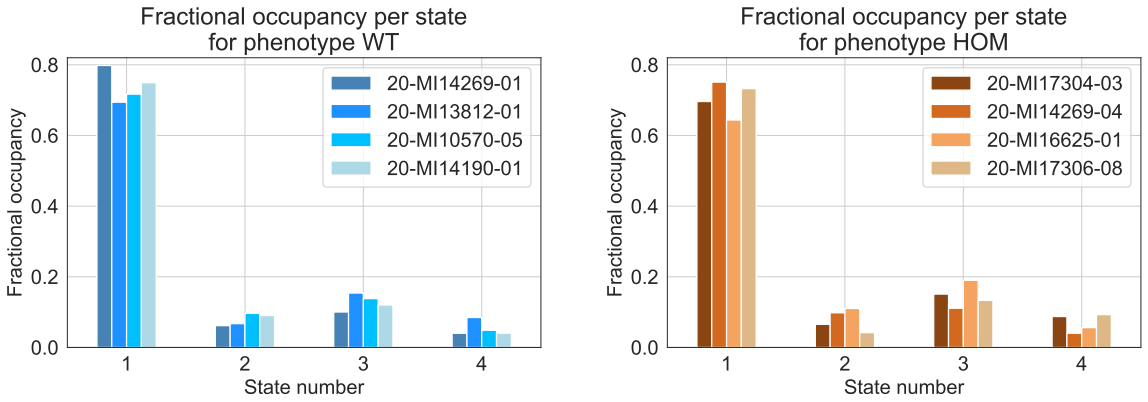
B.5 Consistency of functional network inference



(a) (b)

Figure B.3: Consistency of functional network inference per group of **(a)** WT **(b)** HOM mice.

B.6 Fractional occupancy of states per mouse



(a) (b)

Figure B.4: Individual fractional occupancies of **(a)** WT **(b)** HOM mice. Legends contain labels of animals in correspondence with Tab. B.1.

C Mathematical derivations

C.1 Forward algorithm

In this section, the initialization step, recursive step and termination step of the forward algorithm are derived.

C.1.1 Initialization

Initialization step for forward variable α_1^i.

$$\begin{aligned}\alpha_1^i &= P(\mathbf{y}_1, z_1 = i|\boldsymbol{\theta}) \\ &= P(\mathbf{y}_1|z_1 = i, \boldsymbol{\theta})P(z_1 = i|\boldsymbol{\theta}) \\ &= b_i(\mathbf{y}_1)\pi_i\end{aligned}$$

C.1.2 Recursion

Recursive computation of α_n^j. Let $j \in \mathcal{S} = \{1, ..., K\}$, where K is the number of states.

$$\begin{aligned}\alpha_{n+1}^j &= P(\mathbf{y}_1, ..., \mathbf{y}_{n+1}, z_{n+1} = j|\boldsymbol{\theta}) \\ &= \sum_{z_1,...,z_n} P(\mathbf{y}_1, ..., \mathbf{y}_n, \mathbf{y}_{n+1}, z_1, ..., z_n, z_{n+1} = j|\boldsymbol{\theta}) \\ &= \left[\sum_{z_1,...,z_n} P(\mathbf{y}_1, ..., \mathbf{y}_n, z_1, ..., z_n, z_{n+1} = j|\boldsymbol{\theta})\right] P(\mathbf{y}_{n+1}|z_{n+1} = j, \boldsymbol{\theta}) \\ &= \left[\sum_{i=1}^{K} \sum_{z_1,...,z_{n-1}} P(\mathbf{y}_1, ..., \mathbf{y}_{n-1}, \mathbf{y}_n, z_1, ..., z_{n-1}, z_n = i, z_{n+1} = j|\boldsymbol{\theta})\right] b_j(\mathbf{y}_{n+1}) \\ &= \left[\sum_{i=1}^{K} \sum_{z_1,...,z_{n-1}} P(\mathbf{y}_1, ..., \mathbf{y}_n, z_1, ..., z_{n-1}, z_n = i|\boldsymbol{\theta})P(z_{n+1} = j|z_n = i, \boldsymbol{\theta})\right] b_j(\mathbf{y}_{n+1}) \\ &= \left[\sum_{i=1}^{K} P(\mathbf{y}_1, ..., \mathbf{y}_n, z_n = i|\boldsymbol{\theta})a_{ij}\right] b_j(\mathbf{y}_{n+1}) \\ &= \left[\sum_{i=1}^{K} \alpha_n^i a_{ij}\right] b_j(\mathbf{y}_{n+1})\end{aligned}$$

C.1.3 Termination

Termination step to compute $P(\mathbf{Y}|\boldsymbol{\theta})$.

$$\begin{aligned} P(\mathbf{Y}|\boldsymbol{\theta}) &= P(\mathbf{y}_1, ..., \mathbf{y}_N|\boldsymbol{\theta}) \\ &= \sum_{i=1}^{K} P(\mathbf{y}_1, ..., \mathbf{y}_N, z_N = i|\boldsymbol{\theta}) \\ &= \sum_{i=1}^{K} \alpha_N^i \end{aligned}$$

C.2 Backward algorithm

Recursive computation of β_n.

$$\begin{aligned} \beta_n^i &= P(\mathbf{y}_{n+1}, ..., \mathbf{y}_N|z_n = i, \boldsymbol{\theta}) \\ &= \sum_{j=1}^{K} P(\mathbf{y}_{n+1}, ..., \mathbf{y}_N, z_{n+1} = j|z_n = i, \boldsymbol{\theta}) \\ &= \sum_{j=1}^{K} P(\mathbf{y}_{n+2}, ..., \mathbf{y}_N|\mathbf{y}_{n+1}, z_{n+1} = j, z_n = i, \boldsymbol{\theta}) P(\mathbf{y}_{n+1}, z_{n+1} = j|z_n = i, \boldsymbol{\theta}) \\ &= \sum_{j=1}^{K} P(\mathbf{y}_{n+2}, ..., \mathbf{y}_N|\mathbf{y}_{n+1}, z_{n+1} = j, z_n = i, \boldsymbol{\theta}) P(\mathbf{y}_{n+1}|z_{n+1} = j, z_n = i, \boldsymbol{\theta}) \\ &\qquad P(z_{n+1} = j|\mathbf{y}_{n+1}, z_n = i, \boldsymbol{\theta}) \\ &= \sum_{j=1}^{K} P(\mathbf{y}_{n+2}, ..., \mathbf{y}_N|z_{n+1} = j, \boldsymbol{\theta}) P(\mathbf{y}_{n+1}|z_{n+1} = j, \boldsymbol{\theta}) P(z_{n+1} = j|z_n = i, \boldsymbol{\theta}) \\ &= \sum_{j=1}^{K} \beta_{n+1}^j b_j(\mathbf{y}_{n+1}) a_{ij} \end{aligned}$$

C.3 EM algorithm for the Gaussian HMM

In this section, the EM algorithm of the HMM is derived following [33].

C.3.1 E-step

As described before, $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old})$ is computed as follows:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) = \mathbb{E}[\log(p(\mathbf{Y}, \mathbf{z}|\boldsymbol{\theta}))|\mathbf{Y}, \boldsymbol{\theta}_{old}].$$

Using the HMM factorization:

$$\begin{aligned}\log p(\mathbf{Y}, \mathbf{z}|\boldsymbol{\theta}) &= \log p(z_1|\boldsymbol{\theta}) + \sum_{n=1}^{N-1} \log p(z_{n+1}|z_n, \boldsymbol{\theta}) + \sum_{n=1}^{N} \log p(\mathbf{y}_n|z_n, \boldsymbol{\theta}) \\ &= \sum_{i=1}^{K} \mathbb{1}(z_1 = i) \log \pi_i + \sum_{n=1}^{N-1} \sum_{i=1}^{K} \sum_{j=1}^{K} \mathbb{1}(z_n = i, z_{n+1} = j) \log a_{ij} + \\ &\quad \sum_{n=1}^{N} \sum_{i=1}^{K} \mathbb{1}(z_n = i) \log f_{B_i}(\mathbf{y}_n)\end{aligned}$$

where $f_{B_i}(\mathbf{y}_n) = \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ with $\mathbf{B} = \{B_i : i \in \mathcal{S}\}$ and $B_i = \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$. Now, taking the expectation with respect to $\mathbf{z}$ given the observations $\mathbf{Y}$, resulting in taking the expectation of the indicator functions only. This is equal to the probability of the event in the indicator function.

$$\begin{aligned}Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) &= \sum_{i=1}^{K} P(z_1 = i|\mathbf{Y}) \log \pi_i + \sum_{n=1}^{N-1} \sum_{i=1}^{K} \sum_{j=1}^{K} P(z_n = i, z_{n+1} = j|\mathbf{Y}) \log a_{ij} + \\ &\quad \sum_{n=1}^{N} \sum_{i=1}^{K} P(z_n = i|\mathbf{Y}) \log f_{B_i}(\mathbf{y}_n)\end{aligned}$$

Now, Eq. 4.27 and Eq. 4.39 can be recognized and computed to provide an analytical expression for $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old})$. Substitution leads to:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) = \sum_{i=1}^{K} \gamma_1^i \log \pi_i + \sum_{n=1}^{N-1} \sum_{i=1}^{K} \sum_{j=1}^{K} \xi_n^{ij} \log a_{ij} + \sum_{n=1}^{N} \sum_{i=1}^{K} \gamma_n^i \log f_{B_i}(\mathbf{y}_n)$$

C.3.2 M-step

First, the following maximization problem is established to maximize HMM parameter $\boldsymbol{\pi} = \{\pi_1, ..., \pi_K\}$:

$$\max_{\boldsymbol{\pi}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}), \quad \text{s.t.} \quad \sum_{i=1}^{K} \pi_i = 1$$

Using Lagrange multipliers:

$$\frac{\partial}{\partial \pi_i} \left(Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) - \lambda (\sum_{j=1}^{K} \pi_j - 1) \right) = 0$$

gives

$$\pi_i = \frac{\gamma_1^i}{\lambda} \longrightarrow \lambda = \sum_{i=1}^{K} \gamma_1^i$$

resulting in

$$\pi_i = \frac{\gamma_1^i}{\sum_{j=1}^{K} \gamma_1^j}$$

For the maximization of parameter $\mathbf{A}$ containing elements a_{ij}, that are transition probabilities, again the Lagrange multipliers are used to compute:

$$\max_{\mathbf{A}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}), \quad \text{s.t.} \quad \sum_{j=1}^{K} a_{ij} = 1 \quad \forall i \in \mathcal{S}.$$

Using Lagrange multipliers:

$$\frac{\partial}{\partial a_{ij}} \left(Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) - \lambda(\sum_{j=1}^{K} a_{ij} - 1) \right) = 0$$

gives

$$a_{ij} = \frac{\sum_{n=1}^{N-1} \xi_n^{ij}}{\lambda} \longrightarrow \lambda = \sum_{n=1}^{N-1} \sum_{j=1}^{K} \xi_n^{ij},$$

resulting in

$$a_{ij} = \frac{\sum_{n=1}^{N-1} \xi_n^{ij}}{\sum_{n=1}^{N-1} \sum_{j=1}^{K} \xi_n^{ij}} = \frac{\sum_{n=1}^{N-1} \xi_n^{ij}}{\sum_{n=1}^{N-1} \gamma_n^i}.$$

Finally, the last parameters captured by $\mathbf{B}$ have to be updated. To maximize with respect to the parameters b_i of state i, the derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old})$ with respect to B_i is computed and set equal to zero:

$$\nabla_{B_i} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) = \sum_{n=1}^{N} \gamma_n^i (\nabla_{B_i} \log f_{B_i}(\mathbf{y}_n)) = 0.$$

First, maximizing for $\boldsymbol{\mu}_i$ gives:

$$\frac{\partial}{\partial \boldsymbol{\mu}_i} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) = \sum_{n=1}^{N} \gamma_n^i (\frac{\partial}{\partial \boldsymbol{\mu}_i} \log f_{B_i}(\mathbf{y}_n)) = 0.$$

Now, neglecting the terms in $\log f_{B_i}(\mathbf{y}_n))$ independent of $\boldsymbol{\mu}_i$ and

$$\frac{\partial}{\partial \boldsymbol{\mu}_i} \log f_{B_i}(\mathbf{y}_n) = \frac{\partial}{\partial \boldsymbol{\mu}_i} (-\frac{1}{2} (\mathbf{y}_n - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_n - \boldsymbol{\mu}_i)) = \Sigma^{-1} (\mathbf{y}_n - \boldsymbol{\mu}_i).$$

Substitution gives

$$\sum_{n=1}^{N} \gamma_n^i (\frac{\partial}{\partial \boldsymbol{\mu}_i} \log f_{B_i}(\mathbf{y}_n)) = \sum_{n=1}^{N} \gamma_n^i \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_n - \boldsymbol{\mu}_i) = 0,$$

from which the new estimate of $\boldsymbol{\mu}_i$ can be obtained:

$$\hat{\boldsymbol{\mu}}_i = \frac{\sum_{n=1}^{N} \gamma_n^i \mathbf{y}_n}{\sum_{n=1}^{N} \gamma_n^i}$$

Now, maximizing for Σ_i gives:

$$\frac{\partial}{\partial \boldsymbol{\Sigma}_i} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) = \sum_{n=1}^{N} \gamma_n^i (\frac{\partial}{\partial \boldsymbol{\Sigma}_i} \log f_{B_i}(\mathbf{y}_n)) = 0.$$

Now, neglecting the terms in $\log f_{B_i}(\mathbf{y}_n)$ independent of $\boldsymbol{\Sigma}_i$, using $\boldsymbol{\Sigma}^{-1} = \mathbf{K}$ and the fact that $\log(|\mathbf{K}^{-1}|) = \log(|\mathbf{K}|^{-1}) = -\log(|\mathbf{K}|)$:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{K}_i} \log f_{B_i}(\mathbf{y}_n) &= \frac{\partial}{\partial \mathbf{K}_i} (-\frac{1}{2} (\mathbf{y}_n - \boldsymbol{\mu}_i)^T \mathbf{K}_i (\mathbf{y}_n - \boldsymbol{\mu}_i) - \frac{1}{2} \log |\mathbf{K}^{-1}|) \\ &= -\frac{1}{2} (\mathbf{y}_n - \boldsymbol{\mu}_i)(\mathbf{y}_n - \boldsymbol{\mu}_i)^T + \frac{1}{2} \boldsymbol{\Sigma}_i \end{aligned}$$

Substitution gives

$$\sum_{n=1}^{N} \gamma_n^i (\frac{\partial}{\partial \boldsymbol{\mu}_i} \log f_{B_i}(\mathbf{y}_n)) = \sum_{n=1}^{N} \gamma_n^i (-\frac{1}{2} (\mathbf{y}_n - \boldsymbol{\mu}_i)(\mathbf{y}_n - \boldsymbol{\mu}_i)^T + \frac{1}{2} \boldsymbol{\Sigma}_i) = 0,$$

from which the new estimate of $\boldsymbol{\Sigma}_i$ can be obtained, using the new estimate $\hat{\boldsymbol{\mu}}_i$:

$$\hat{\boldsymbol{\Sigma}}_i = \frac{\sum_{n=1}^{N} \gamma_n^i (\mathbf{y}_n - \hat{\boldsymbol{\mu}}_i)(\mathbf{y}_n - \hat{\boldsymbol{\mu}}_i)^T}{\sum_{n=1}^{N} \gamma_n^i}$$