NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

Division of Computer Science
Microprocessor and Digital systems Lab (MicroLab)

Functional Ultrasound Imaging: Study of the brain using the ICA modality

Diploma Thesis

by

Vaia I. Kontopoulou

Academic Supervisor: Dimitrios Soudris,
Professor, Electrical and Computer Engineering, NTUA

Athens, July 2022

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF ELECTRICAL
AND COMPUTER ENGINEERING
DIVISION OF COMPUTER SCIENCE
MICROPROCESSOR AND DIGITAL SYSTEMS LAB

Functional Ultrasound Imaging: Study of the brain using the ICA modality

Diploma Thesis

of

Vaia I. Kontopoulou

Academic Supervisor: Dimitrios Soudris,
Professor, Electrical and Computer Engineering, NTUA

Approved by the following examination committee, the 18th of July 2022.

.....	.....	.....
Dimitrios Soudris	Panagiotis Tsanakas	Georgios Matsopoulos
Professor, NTUA	Professor, NTUA	Assosiate Professor, NTUA

Athens, July 2022.

.....

Vaia I. Kontopoulou
Electrical and Computer Engineer, National Technical University of Athens

Abstract

The study of functional configuration of the brain and the nervous system of living organisms, has been the principal target as well as the biggest challenge of Neuroscience and its corresponding scientific fields. The structural complexity of the nervous system and also the scale of its fundamental phenomena and interactions comprise some of the biggest challenges related to this research area. In recent years there has been a surge in hardware set-up and computational methods development in the fields of Medical Imaging and Computational Science, paving the way for rapid advancement in approaching these challenges. The increasing need for greater spatio-temporal resolution and portability potential is the main reason that the functional ultrasound modality was introduced as a novel approach in mapping the functional responce of the brain through the dynamic quantification of the cerebral blood volume. In the present study we explore the potential of the ICA method with regard to the analysis and information extraction from functional ultrasound data, regarding the cerebral functionality of rats, in a visual stimulation experimental set-up. The exploration of the ICA modality in combination with the analysis of the resulting fUS datasets, is quantified using the Icasso clustering software, and it revolves around three basic axis: Researching the effect of different preprocessing parameters of the functional ultrasound signal, in the ICA results, research of the effect of added noise in the ICA output and exploration of the ICA modality from a dynamic analysis perspective of the fUS data, with respect to time. In addition, we performed an extensive review of the existing research in the field of mapping the brain functionality using fUS, as well as in the area of brain functional data analysis. The five datasets which we use in the present study have resulted from 2D visual-stimulation conducted on mice at the Neuroscience department of the Erasmus Medical Center in Rotterdam.

List of Figures

List of Tables

List of Acronyms

BOLD Blood-Oxygen Level Dependent.

BSS Blind Source Separation.

CT Computed Tomography.

DMN Default Mode Network.

fMRI functional magnetic resonance imaging.

fUS functional ultrasound.

Hip Hippocampus.

ICA Independent Component Analysis.

MBS Main Blood Supply.

MRI Magnetic Resonance Imaging.

PET Positron Emission Tomography.

RSP Retrosplenial cortex.

SNR Signal-to-Noise Ratio.

SSp primary Somatosensory area.

Contents

Chapter 1

Introduction

The biological brain consists of a complex system of nerve cells and synapses, organized in topologies with basic common anatomical and physiological characteristics per species, but also with a high degree of randomness per individual. The study of the brain is one of the most fascinating fields of modern research and has been largely developed due to the technological development of methods for imaging and mapping the interconnection of the various centers of the brain and their response to external stimuli.

In the present study we are dealing with a relatively new method of brain functional imaging, functional ultrasound, which presents significant advantages over conventional functional imaging methods and provides opportunities to cover and enhance key points of the human nervous system research.

1.1 Statement of the problem

The imaging of the functional interconnections of the brain is usually approached in two ways, one direct and one indirect. In the first case we are concerned with the imaging of electrical neural signals, which are also the main focus of brain response studies. The second case concerns methods of indirect imaging of nerve signals, through the mapping of the hemodynamic response of the brain. The correlation between the hemodynamic and neural signals is based on the coupling of the respective systems which is expressed by the perfusion of the capillaries that surround the nerve cells, every time they become activated. Direct imaging of the brain function can be performed with the use of techniques such as the use of special dyes that are sensitive to the voltage level, calcium imaging, electroencephalography etc. On the other hand, regarding the indirect imaging of the brain, techniques such as functional magnetic resonance imaging (fMRI), Positron Emission Tomography (PET), internal optical imaging and photoacoustic imaging are widely used. The present study focuses on the analysis of the brain function using indirect imaging and as a result, we will refer to the corresponding techniques in greater detail and in particular, to the functional ultrasound

(fUS) method.

The best possible spatio-temporal sensitivity is achieved so far using optical imaging methods ($\sim 10\mu m$, $\sim 10ms$), which require opening or thinning of the skull in order to acquire measurements. Furthermore, the range of light frequencies used does not allow the imaging of the brain in great depth. In this field, photoacoustic imaging techniques display greater penetration but with still low absolute values ($\sim 1mm$).

In-depth imaging of the brain with very good spatio-temporal resolution is achieved (until recently) using the fMRI and PET techniques. The later achieves three-dimensional imaging of the brain and its various processes, using radioactive elements which are detected by special sensors. Due to its low spatial resolution, this method is usually combined with some additional anatomical imaging modality such as Computed Tomography (CT) scan or Magnetic Resonance Imaging (MRI). The fMRI method is the most common functional imaging technique and is based on the Blood-Oxygen Level Dependent (BOLD) signal measurement, the value of which at any given time depends on the level of oxygenation of the blood. The detection of changes in this signal is based on the spatial and temporal inhomogeneity of the detected magnetic field in the brain, due to the different magnetization values of deoxyhemoglobin (oxygen-free hemoglobin) and oxyhemoglobin (oxygenated hemoglobin). Fmri is a non-invasive imaging method, widely used in both research and clinical applications. However, achieving high spatial imaging accuracy requires the application of a strong magnetic field and has the downside of reducing the time resolution and the Signal-to-Noise Ratio (SNR) of the resulting images [1]. As a result, it is difficult to functionally visualize short-term transient phenomena. The size and cost of MRI machines also makes them unprofitable and does not allow fMRI to be performed during surgeries.

The modality of functional ultrasound constitutes a novel proposal in the field of functional imaging, with features that solve many of the problems of fMRI. The high spatio-temporal resolution and portability of this method, make its application ideal for imaging the brain, even without the use of contrast factors. In the introductory work of Macé et al. (2011) [2], fUS was used to visualize the rat brain and map its activation patterns with high accuracy and a high signal-to-noise ratio. The fUS method uses the power Doppler technique in order to capture highly sensitive images, with the ability to display very small vessels. The acquisition of μDoppler images at a high frame-rate (ultrafast imaging), enables the visualization of the blood flow in the vascular system, which is shaped by the neuronal response.

The analysis of functional ultrasound data is mainly performed using the tecniques and algorithms used in the fMRI data analysis. The fUS method being a relatively new application of ultrasound, research into the analysis of the data produced is at an early stage and borrows algorithms from fMRI, which is the golden rule of the imaging techniques concerning the functionality of the brain. In the present work we will approach the analysis of fUS data using the Independent Component Analysis (ICA) algorithm, a model-free, data-driven approach to functional brain imaging, widely used in the case of fMRI.

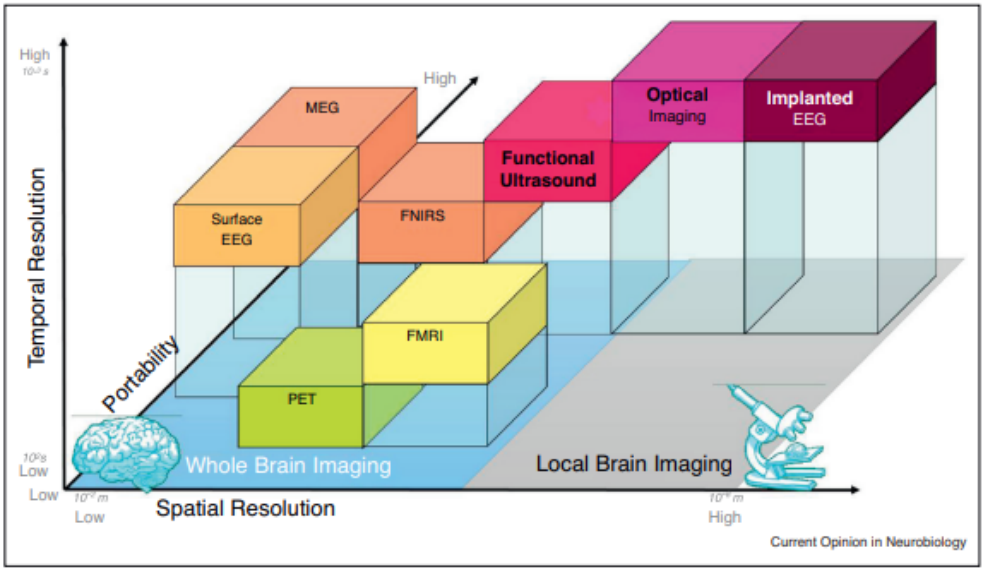


Figure 1.1: Comparative representation of brain imaging methods, in terms of their portability and spatio-temporal resolution [3].

1.2 Purpose statement of the Thesis

The research purpose of my thesis consists of the exploration of the ICA modality with regard to its potetial in the analysis of functional ultrasound data of the brain. The research is based on the explorative study of the five functional ultrasound datasets that were developed during experiments using mice as subjects, and were performed in the Erasmus Medical Center in Rotterdam. Due to the fact that the imaging modality of the functional ultrasound provides us with a limited scientific bibliography regarding its application in brain imaging, the present study uses processing techniques that have been applied succesfully with the much more popular fMRI imaging technique. Our analysis is focused on the use of the ICA method in the analysis of fUS data, with the use of the Icasso clustering software.

1.2.1 A few things about the ICA

The ICA technique is a statistical method for the analysis and decomposition of a composite dataset in a number of independent subsets of data (decomposition-based approach). The ICA does not make any a priori assumptions with regard to the dataset it aims to analyze and does not consider a particular model for the signal sources it is called upon to highlight (model-free). The set of problems which consider the decomposition of compound information signals into source-signals, even in the case where very limited information is available regarding these sources, is called Blind Source Separation (BSS). The ICA method is one of the most popular algorithms

which tackles problems of this particular set, eg. the classic coctail party problem. The aim of this particular application is the detection and extraction of the voice signal of a specific sound source inside an environment where different voice sounds and noise overlap.

The statistical independence and non gaussian distribution of the signal sources it aims to disentangle, constitute the basic assumptions of the ICA method. There are also some ambiguities and limitations which arise during the application of the ICA method, regarding the interpretation of its results. We will refer to these points analytically in Chapter 2.

1.2.2 Main research directions

In the present thesis we will research the ICA technique in three basic axis:

1. Preprocessing of the input dataset of the method
2. Noise tolerance
3. Dynamic analysis ot the results of the ICA method

In particular, regarding the first part of the analysis, we will explore the effect of the input data dimension on the stability and repeatability of the independent signal sources, defined as a result by the ICA. In the second part of the analysis we add gaussian noise in the input of the algorithm, observing once again the robustness of the results. In the third part of the analysis we research the ICA method with regard to its results which correspond to consecutive subsets of the total input dataset, with respect to time. We compare these results with the results yielding from the analysis of the total input dataset in each case.

1.2.3 The fUS dataset

Our work handles functional ultrasound data, created during a series of optical stimulation experiments dealing with the hemodynamic response of the mouse brain. As a result, the ICA method in this particular case is called upon, to return a set of independent sources for the hemodynamic responce signal. This particular set does not necessarily have a clear correspondance with the known anatomical and physiological brain regions of the mouse. We will refer to the the processing steps of the fUS data and the interpretation of the ICA results, in more detail, during the rest of the thesis.

1.3 Thesis outline

In Chapter 2, the fUS and ICA techniques are presented in detail, with a special focus on their theoretical principle, their operation and their various applications. In Chapter 3 we present a literature review, of the research field that includes the functional ultrasound modality and the analysis methods used for exploring its results. In Chapter 4, the whole research course of our study is presented, starting from the acquisition of the experimental results, to their preprocessing strategy and their feeding to

the ICA implementation algorithm. The ICA algorithm and the Icasso software which quantifies the results of our research are presented in this chapter. The experimental results of the thesis and their scientific interpretation are displayed in Chapter 5. Finally, the conlusions that arise from the present study are displayed in Chapter 6, where we also propose some future research directions, regarding the use of the ICA method in the analysis of functional ultrasound data from the brain.

Chapter 5

Experimental Results

In the present study, the stability of the independent signal components derived by the ICA, is measured using the Iq stability index. The calculation of the mean of the Iq indices for all of the independent components that emerge with the application of the algorithm, is used to calculate the stability for the entire investigated system, for a specific input set of parameters. Based on the afforementioned assumptions, in the rest of our study we consider as a reference set (reference combination of number of independent components, total system stability and calculation running time) the choice of 20 independent signal components with a PCA dimension of 20, and 100 resampling cycles (Icasso parameters). The result of this analysis is used in the rest of the study as a reference frame for the comparison of the products of Icasso, but also for matching the images of the independent components derived by the ICA, on the output grid.

With respect to the image matching, it is realised using the munkres algorithm [4], having as an input the matrix of correlations between the independent input components and the reference frames. Establishing this correspondence is an essential step for spatial classification of the independent component images, in order to enable overview and comparison of the results.

The five datasets used in this work, were processed in the same way in order to establish the reliability of both the experimental measurements and the computational processing. For each of the datasets, two grids of reference components were produced, one with ICA dimension 20 and PCA dimension 20 and another with ICA dimension 100 and PCA dimension 100. Both the reference sets (for each input data set) were calculated with 100 ICA resampling cycles. In figures 5.1 to 5.5 we present the reference results of 20 components, for each of the five datasets used in the present study. The reference frames of 100 components are presented in the Appendix.

In addition to the presentation and commentary of the experimental results of the present study, in this chapter we will also refer to the possibilities of their extension.

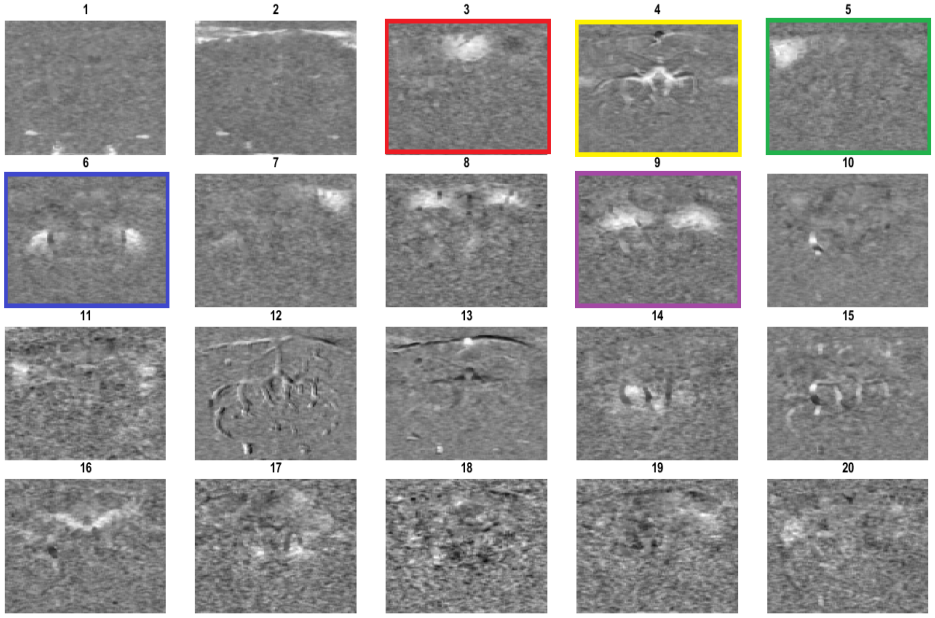


Figure 5.1: Reference components of dataset 14_14_53.

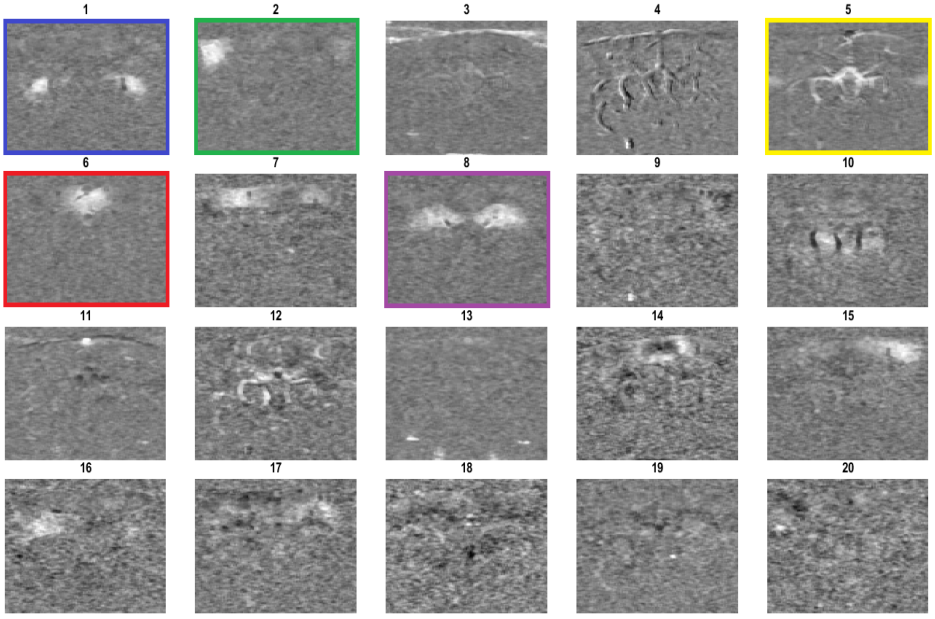


Figure 5.2: Reference components of dataset 14_30_03.

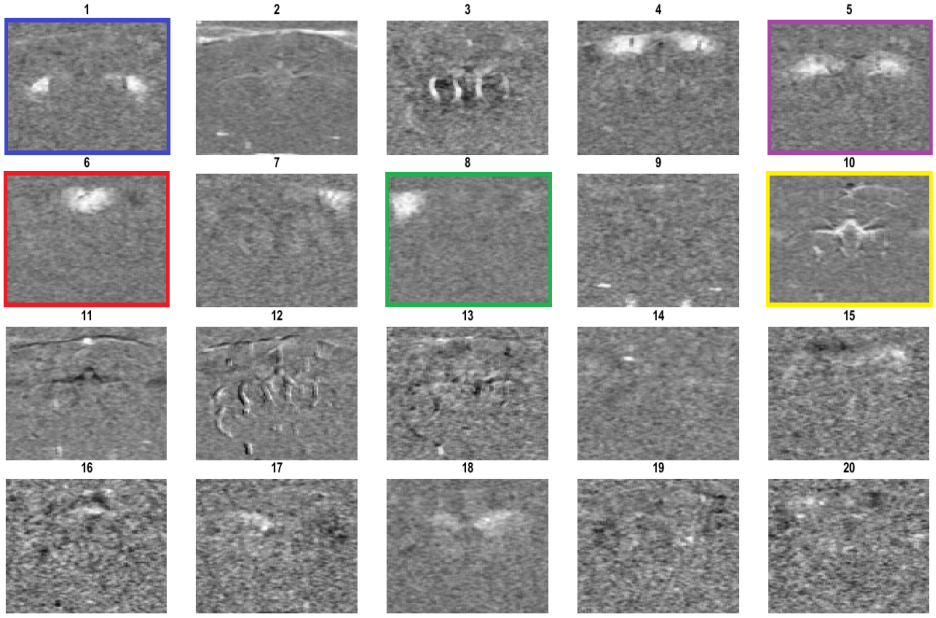


Figure 5.3: Reference components of dataset 14_39_20.

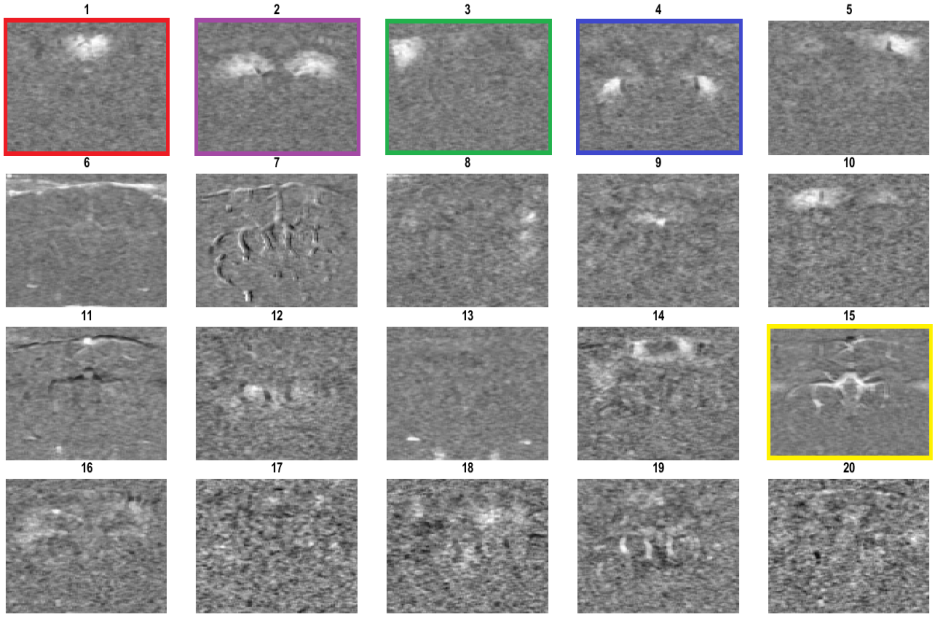


Figure 5.4: Reference components of dataset 14_43_46.

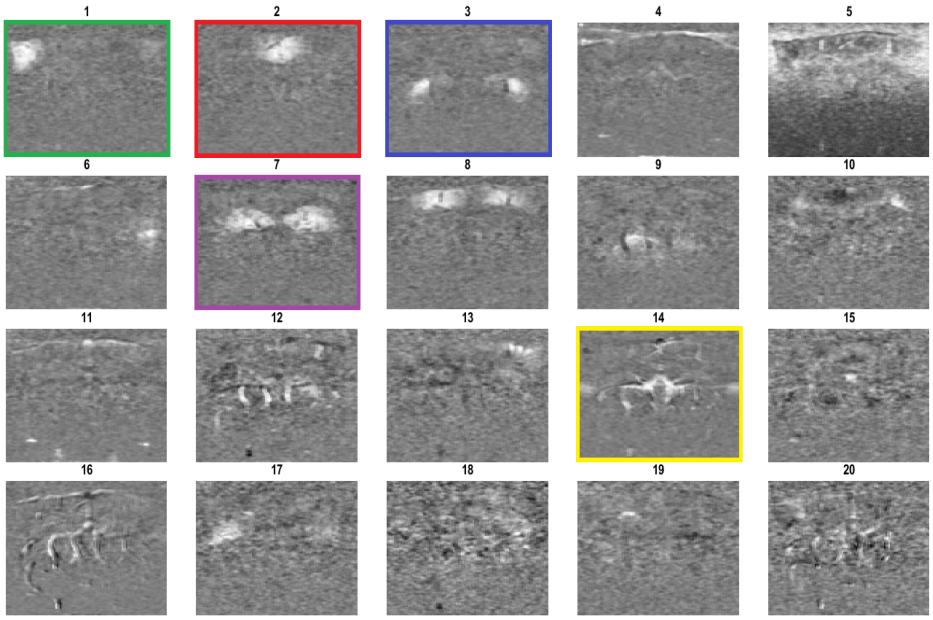


Figure 5.5: Reference components of dataset 14_48_15.

5.1 Varying PCA dimension, with a fixed number of independent signal sources

5.1.1 Experimental Results

The first analytic experiment we performed, concerns the effect of the PCA dimension applied to the input data, as one of the preprocessing steps before feeding it to the ICA algorithm. Keeping constant the number of independent signal sources we ask to be determined by ICA, we let the dimension of the PCA analysis to gradually vary from 20, up to 200 with a step equal to 20. We use the component mapping function so that any changes in the components are visually detected at each step of the analysis.

Because we do not want to use a large number of component vectors in the ICA analysis (due to the computational complexity and large processing time requirements of the Fastica algorithm), PCA produces the dominant components of the data set to feed into the ICA. In this way, however, the reliability of the signal is reduced while some phenomena of short duration - but possibly of great importance - are lost from the analysis.

Figure 5.6 shows the result of using Icasso to apply the ICA in order to select the 20 most robust independent signal sources with varying PCA dimension from 20 to 200. This chart shows the average values of the "cluster" quality index Iq, for the fUS experiments 1 to 5, as well as the percentages of Icasso results per Iq range, for each of the data sets.

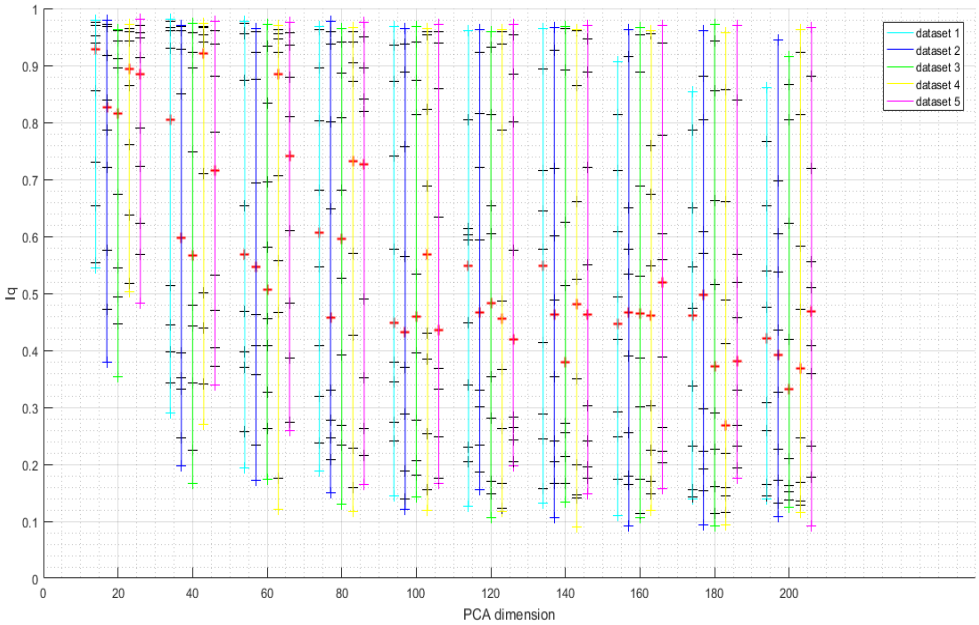


Figure 5.6: Stability index Iq for the clustering produced by Icasso with varying PCA, for each of the five datasets.

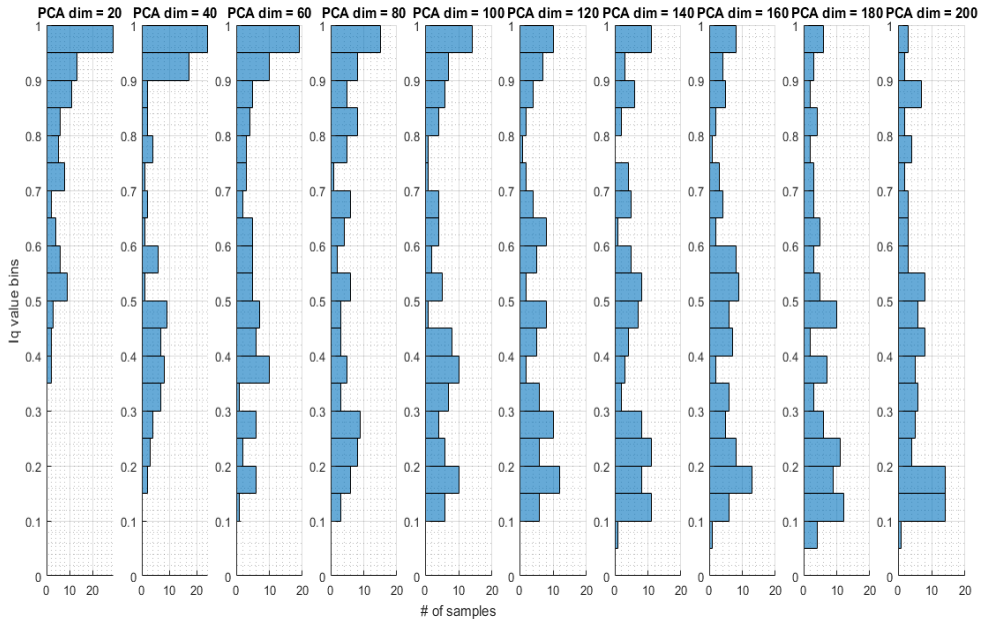


Figure 5.7: Comparative visualization of the Iq-value histograms of the total results, with 100 points per PCA dimension.

We observe that the maximum Iq value for each data set and for each PCA dimension remains close to 1, while the minimum and maximum curves for all the experiments are clearly depicted in the chart. These curves become more apparent in Figure 5.7 where they depict for each PCA dimension, the percentage concentration of the 100 points of the 5 datasets in total.

On the other hand, the median of Iq values decreases as the PCA dimension increases with the relationship between the two quantities not necessarily being linear. The median value also appears to show a fairly defined range of values for the five datasets, with a maximum difference of 0.37 for 60 eigenvectors and a minimum of 0.07 for 160 eigenvectors.

Figure 5.7 is derived from combining the results of the 5 datasets, as mentioned above. The histograms show a clear shift of the number of experimental points towards lower Iq value,s as the PCA dimension of the analysis increases. In other words, the distribution of independent components given by the ICA changes gradually as the values move from around 1, to 0.

In the second part of the analysis, from the frames of the reference set (Figure 5.1), frames #3, #4, #5, #6 and #9 were selected for an indicative presentation of the results. These specific components were chosen because they appear clearly in all datasets (Figures 5.1 - 5.5 with red, yellow, green, blue and purple colors respectively) and it is feasible to visually match them with defined anatomical regions of the mouse brain. At this point, it is worth noting that of the remaining components, there were some that appeared clearly in certain datasets' reference frames (such as component 14 of dataset 14_30_03 appearing in position 14 of the dataset 14_43_46 reference results) and others that corresponded to defined anatomical regions (such as components 7 and 8 of dataset 14_43_46), which are not clearly visible in all reference datasets, and their behavior seems to be a consequence of the way the ICA selects the signal components. It has already been mentioned that the ICA results in a series of independent components that make up the input signal. However it is possible for some of the actual signal sources to be separated during analysis, which appears (visually) to be the case with components 7 and 8 of dataet 14_43_46.

For the five spatial signal components chosen to be presented, we attempted a mapping of their images to anatomical regions of the mouse brain. Their numbering hereafter follows the numbering of the reference set of 20 independent components and PCA dimension 20, of the dataset 14_43_46 (Figure 5.1).

Component 3 (red) anatomically corresponds to two regions of the retrosplenial cortex (Retrosplenial cortex (RSP) v and d), which are a key part of the brain's spatial information processing network. These areas are also part of the brain's Default Mode Network, which according to the literature shows low levels of activation when performing a specific mental task and high levels of activation when the mental task is not focused on a certain goal. From now on we will refer to this specific area with the name RSP (Retrosplenial area).

Component 4 (yellow) corresponds to the main perfusion of the brain, as shown in Figure 4.2. The specific component is not an anatomical region but a key part of

the system of blood vessels that supply the brain and for this reason it was considered noteworthy in the present analysis. In the rest of the work we will refer to the specific component with the name Main Blood Supply (MBS).

Component 5 (green) corresponds to a part of the somatosensory area of the brain and concerns the processing of the stimulus in the mouse's whiskers, which is one of the main sources of spatial information and allows the animal to perceive and react to the characteristics of its environment. Hereafter we will refer to this region as primary Somatosensory area (SSp).

Component 6 (blue) corresponds to a part of the external (lateral) angular nucleus, the primary processing center for visual information received from the retina of the eye. It is part of the thalamus and we will refer to it as the LGN (Lateral Geniculate Nucleus).

Finally, component 9 (purple) corresponds to the main two parts of the hippocampus, which according to the anatomy of the mouse brain plays a key role in learning and memory processes. We will refer to this component as Hippocampus (Hip).

Figures 5.8 - 5.12 show the values of the stability index Iq given by Icasso for the five selected components, in each of the five data sets and for each value of the PCA dimension, for a selected number of independent components equal to 20.

Here we should briefly refer to the way in which Figures 5.8 to 5.12 were produced. More specifically, for each of the five datasets we map the 20 components of the reference set (Figures 5.1 - 5.5) to the 20-component reference set of the dataset 14_14_53 (Figure 5.1), ignoring matches with a convolution value below 100. The threshold of 100 is empirical and was set with the aim of eliminating false matches as much as possible, as we have assumed from the outset that not all data acquisitions give the same independent spatial signal components. The convolution value given by the matching function has no specific or reliable value limits. However, the threshold we define cuts off the association of the most stochastic spatial components and for this reason it is used at this stage of the analysis. Next, for each value of the PCA dimension, we map the generated components to the reference set and from there, to the reference set of dataset 14_14_53 (Figure 5.1), with the constraint we mentioned earlier. Finally, we print the stability index values for all of the experiments and for each of the five selected components.

Appendix A details the results of the varied PCA dimension and 20 independent components experiments, for the 14_14_53 data set.

5.1.2 Observations and Conclusions

In this first experiment of the present study, we investigated the influence of the PCA preprocessing dimension on the quality of ICA results. Using the stability index Iq, we observe that the results of the Icasso iterations with an increasing value of the PCA dimension, show a decrease in the overall stability for each of the five datasets and a shift of the ICA estimates to lower stability values. This is something already expected, based on the theory of data preprocessing with PCA. As we increase the

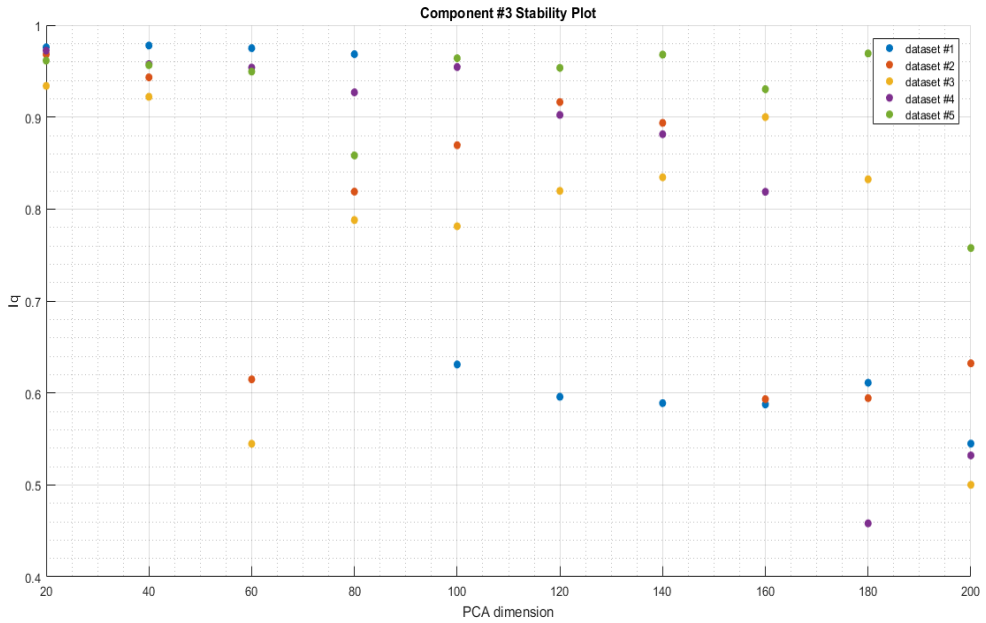


Figure 5.8: Iq stability values for the RSP signal component, for the five input datasets of the ICA with varying PCA dimension.

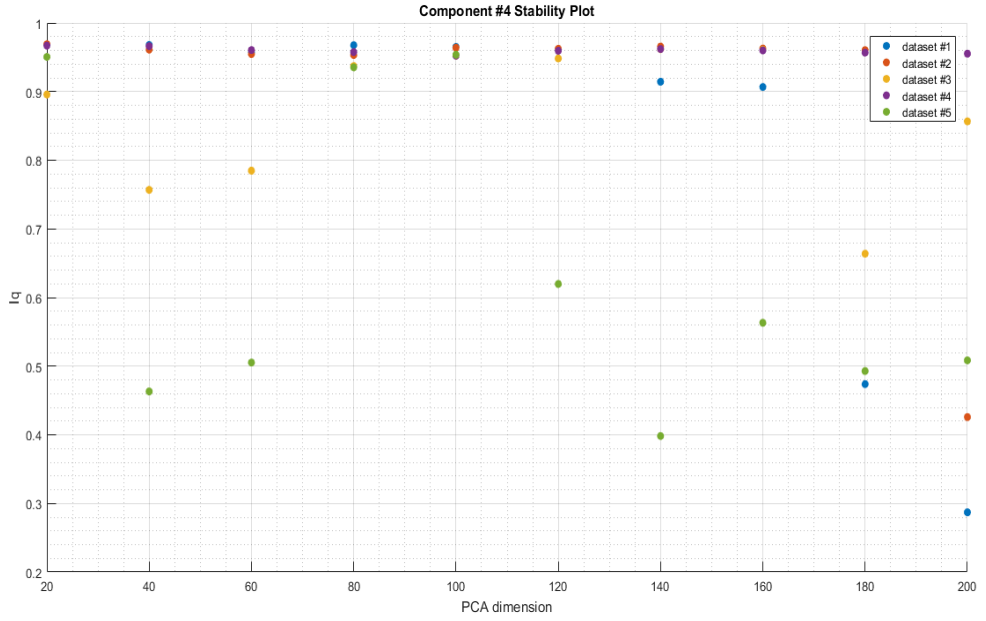


Figure 5.9: Iq stability values for the MBS signal component, for the five input datasets of the ICA with varying PCA dimension.

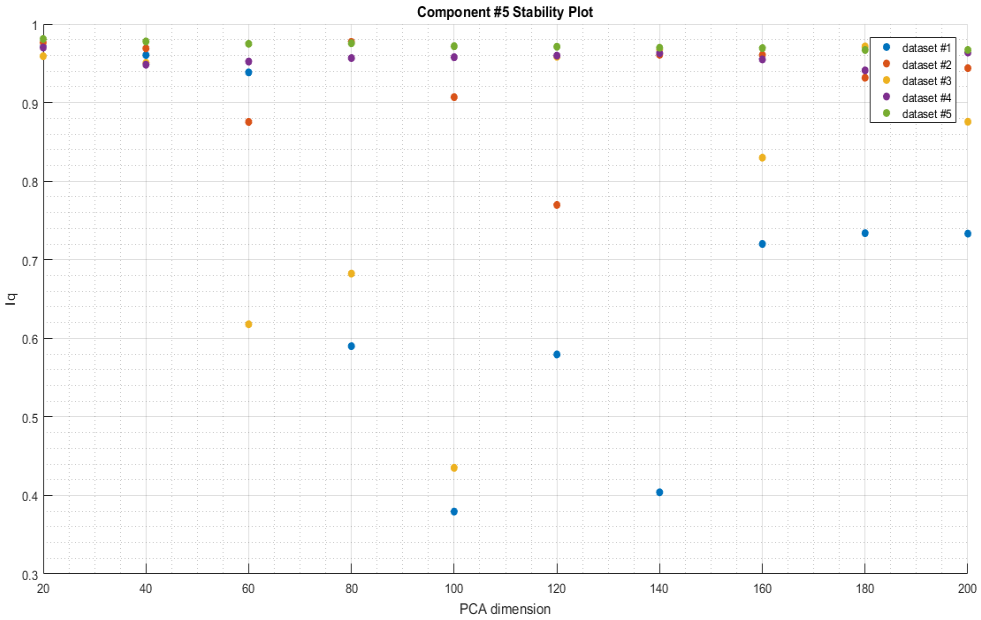


Figure 5.10: Iq stability values for the SSp signal component, for the five input datasets of the ICA with varying PCA dimension.

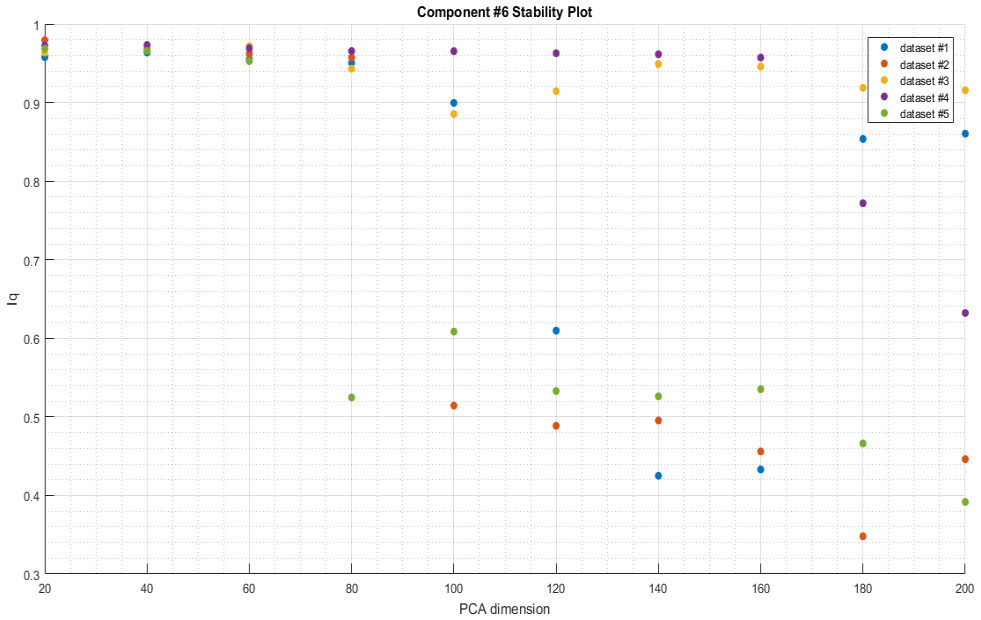


Figure 5.11: Iq stability values for the LGN signal component, for the five input datasets of the ICA with varying PCA dimension.

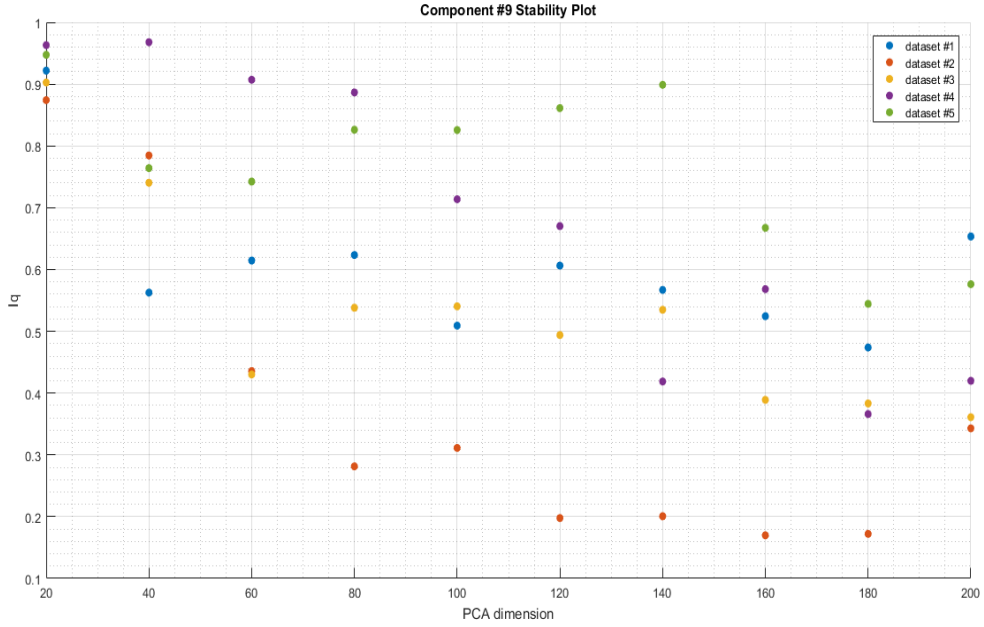


Figure 5.12: Iq stability values for the Hip signal component, for the five input datasets of the ICA with varying PCA dimension.

dimension of the input data set fed to the ICA algorithm, the dimension of the input eigenvector matrix increases and the algorithm "struggles" to identify the same independent signal components. Also of interest is the fact that, as the PCA dimension increases, the maximum Iq value remains almost unchanged, while about one-tenth of the twenty independent components maintain an Iq value above 0.7-0.75 for all values of the input data dimension. In the plots depicting the stability of the individual signal components for each data set (Figures 5.8 - 5.12, Appendix Figures 15 - 16) we notice that there is a stability profile for each of the independent components. This means that not all components follow the same profile in terms of the Iq index, but the specific results depend both on the choice of the component and on the given experimental detection of the input signal. Between different PDI sequences, for most of the independent components given by the ICA, large variations in the value of the stability index appear, with some of the curves following a downward trend with increasing PCA dimension (e.g., Figures 15b', 15g', 16a') and others displaying a more complex behavior (eg Figures 15d', 15e', 15f', 15g'). The conclusion that emerges from the above observations is that for the final selection of the meaningful fUS signal components, and in general for the evaluation of the ICA results, a further delving into the functional profile of the components resulting from the algorithm might be needed, in combination with the parameters of the initial functional ultrasound experiment so that the stability profile per component can be evaluated.

Regarding the interpretation and inference of conclusions for the brain areas #3,

#4, #5, #6 and #9 (RSP, MBS, SSp, LGN and Hip respectively), we can say the following:

From the stability profile of the RSP region (Figure 5.8) we see that it appears in a consistent manner in every analytic experiment, with a drop in its stability (on average) as the PCA dimension increases. This particular area plays an active role in brain function, when it does not require concentration on a specific task. While for PCA dimension equal to 20 it shows high stability, the difficulty in detecting it in some of the datasets for larger values of that particular pre-processing dimension may indicate that it does not play an essential role in brain function during this particular visual activation experiment.

Looking now at the activation profile of the MBS region (Figure 5.9) we see that it shows very high stability in almost all datasets and for all PCA dimensions (with the exception of dataset 14_48_15). Because this particular area is a focal point of the tree of vessels that supply the mouse brain with blood, it is safe to infer the reason for its constant presence in the results. Regarding the exception of the data set 14_48_15, one guess we could make is to suggest that in this particular experiment there was a lower brain perfusion on average, related to certain parameters of the experimental setup that were not available in the present work.

A similar stable activation profile is displayed by the SSp area (Figure 5.10) whose function is linked to the spatial perception of the animal through its whisker. Because the function of this region is (at least theoretically) indirectly linked to the visual activation underlying the experiment we are studying, additional experimental activation data are needed to draw a firm conclusion about why it exhibits such high stability. It is possible, however, that its stability is due to the de facto significance of the mouse's whisker in each case, or even to the fact that, due to the projection of arbitrary and spatially meaningless images into the animal's field of vision - a projection which hinders its visual ability to navigate space-, the mouse relies heavily on its whisker sense in order to perceive its environment, hence the intense activity of this particular brain region.

Observing Figure 5.11 in which the stability of the LGN region during the experiments is depicted, we observe high stability for low values of the PCA dimension, and dual behavior of the estimates for PCA values of 100 and above. In half of the datasets, this region shows consistently high stability, while in the rest we have a gradual while the PCA dimension increases. Since, as mentioned above, we do not have a lot of data on the parameters of the experimental setup for obtaining the functional ultrasound measurements, it would be fallible to conclude on the differences regarding this parameter between the different data sets, which may be due to the acquisition order of the measurements, the animal's fatigue or even to its habituation to the visual stimulus. As the LGN region is functionally involved in the processing of visual stimuli, its stability during the experiments and following analysis is to be expected. The difference in the behavior of the data sets during the analysis is therefore likely to lie in fundamental differences regarding the way the measurements were taken.

Regarding the Hip region (Figure 5.12) which is associated with the learning and

memory process, it shows a very wide range of values across the five data sets for each PCA dimension. We also notice that three out of the five data sets maintain a constant value of region-specific stability (sets 14_14_53, 14_39_20, 14_48_15) and the other two (14_30_03, 14_43_46) show a constant decrease in its stability. This uniformity of results - even in groups - is of great interest to the analysis of the functioning of the specific area, and may in the future be able to provide safe conclusions regarding the interpretation of its activation.

As a general conclusion in this particular analysis experiment, it can be taken that the preprocessing of the input data of the ICA with PCA, reduces the complexity of the problem of determining the main signal components while at the same time it carries the risk of losing some sources with a short duration of activation, but possibly of importance for the brain function. Each of the spatial components identified by the ICA, exhibits a distinct activation behavior, largely intertwined with both its theoretical functionality and its functional relationship to the given experimental setup and its technical parameters. A more accurate interpretation of the results requires knowledge of the functional characteristics of the spatial components in depth, as well as the characteristics of the experimental setup for obtaining the functional ultrasound measurements.

5.2 Data Analysis with Added Noise and a Fixed Number of Signal Sources

5.2.1 Experimental Results

Here we applied ICA analysis with the number of independent signal components being equal to 20, and a PCA dimension equal to 20. This particular combination was chosen (as mentioned in section 4.4.3) as a reference system and is used in the rest of the thesis in the total input data set of the algorithm. We add to each of the analysis pipelines, Gaussian noise with snr values from 5 to 30 (with a step of 5) and we use Icasso to calculate the stability indices of the results, the time courses of the activation of the components and compare these results with the results that the ICA gives without the extra noise.

The choice of the lower limit for the snr (snr=5) is justified by the fact that, for lower snr values the algorithm fails to detect a sufficient number of independent components. In the case of snr=5, for the five data sets, four to five distinct components are detected. On the other hand, the upper limit (snr=30) was chosen since the algorithm succesfully detects almost all of the twenty main components.

It should also be mentioned that the upper limit of the iterations of the algorithm by Icasso was increased from 300 to 1000, since the algorithm - especially for snr values from 5 to 15 - does not sufficiently converge in the majority of the added noise cases.

Figure 5.13 shows the overall stability index for each of these analysis cases, for an increasing snr value. As expected, we observe an overall increase in stability in parallel

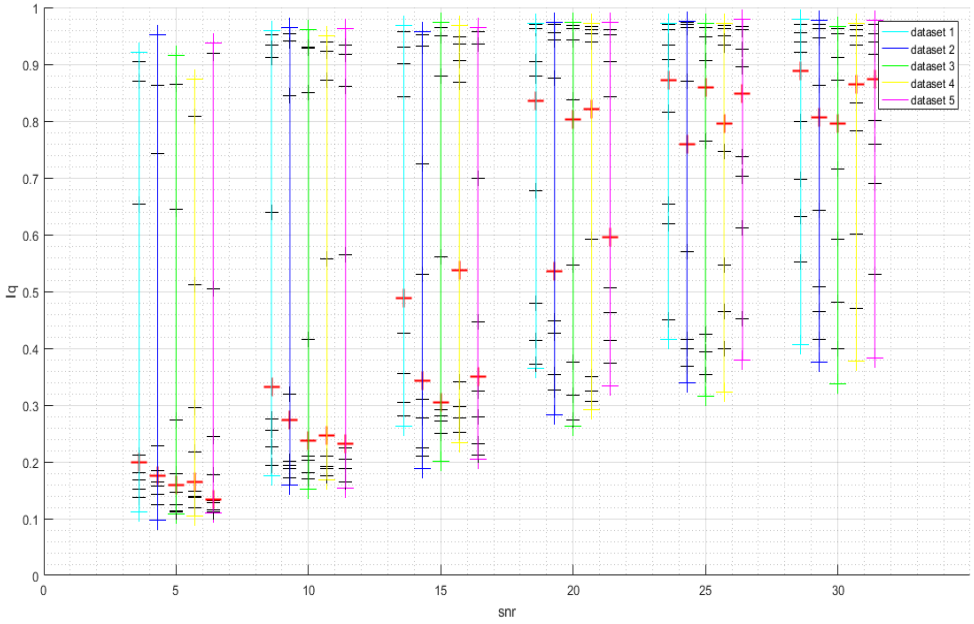


Figure 5.13: Iq stability index for the clustering produced by Icasso with varying snr value, for each of the five datasets.

to noise reduction. We observe that for all the cases, the maximum Iq value remains close to 1, with the minimum moving from 0.1 to higher values as the value of the snr increases. On the other hand, the median increases steadily with the increasing snr (with the exception of the third data set at snr 25 and the fifth data set at snr 30) and shows a small range of values at the extreme snr values (5-10 and 25-30) and a greater range in the middle values (20-25). Of interest is the fact that the signal components derived from the experiments mostly display extreme Iq values for low snr, with the centimeter distances approaching - and the median moving towards 1 - as the snr increases. This is most evident in figure 5.14 where the histograms of the total results of the five data sets are presented. The distribution has two centers: one near Iq=1 and the second near the minimum Iq. This minimum increases and at the same time it contains a lower percentage of the total energy of the distribution as the value of the snr increases. This histogram form simulates the form of the corresponding images 5.6-5.7 for a variable number of eigenvectors, with the low peak of the distribution increasing and containing a lower percentage of the total energy of the distribution as the number of eigenvectors of the signal fed to the ICA decreases.

As in the previous part of the fUS data analysis (section 5.1), we present in Figures 5.15 - 5.19 the values of the stability index Iq for each of the spatial components RSP, MBS, SSp, LGN and Hip of the dataset 14_14_53, as described in paragraph 5.1. These images were produced in the manner described above for Figures 5.8 - 5.12.

Finally, in Figures 5.20 and 5.21, the spatial signal components for each of the

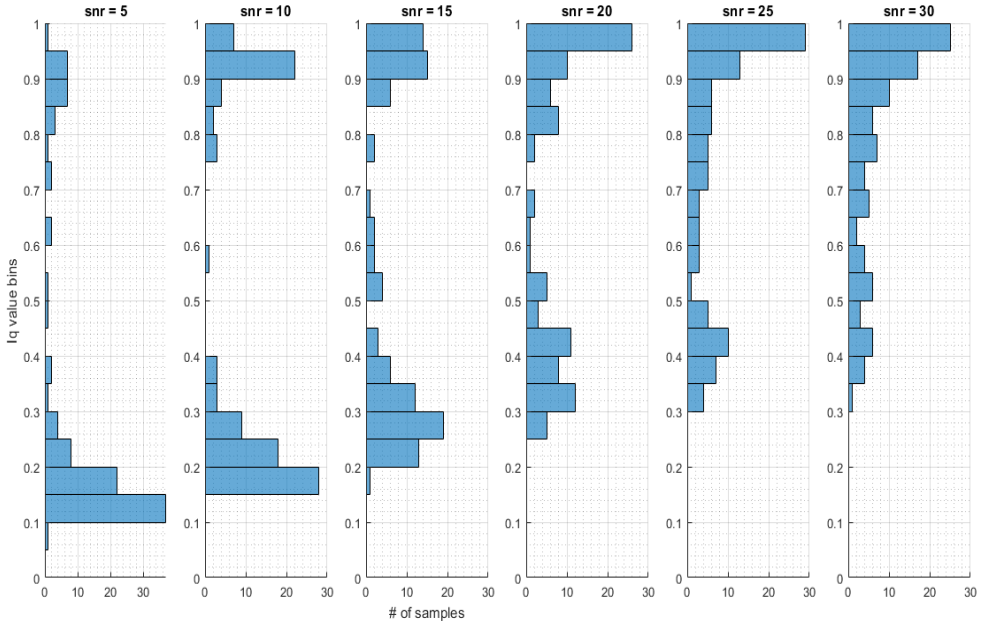


Figure 5.14: Comparison of the Iq-value histograms of the total results, with 100 points per snr value.

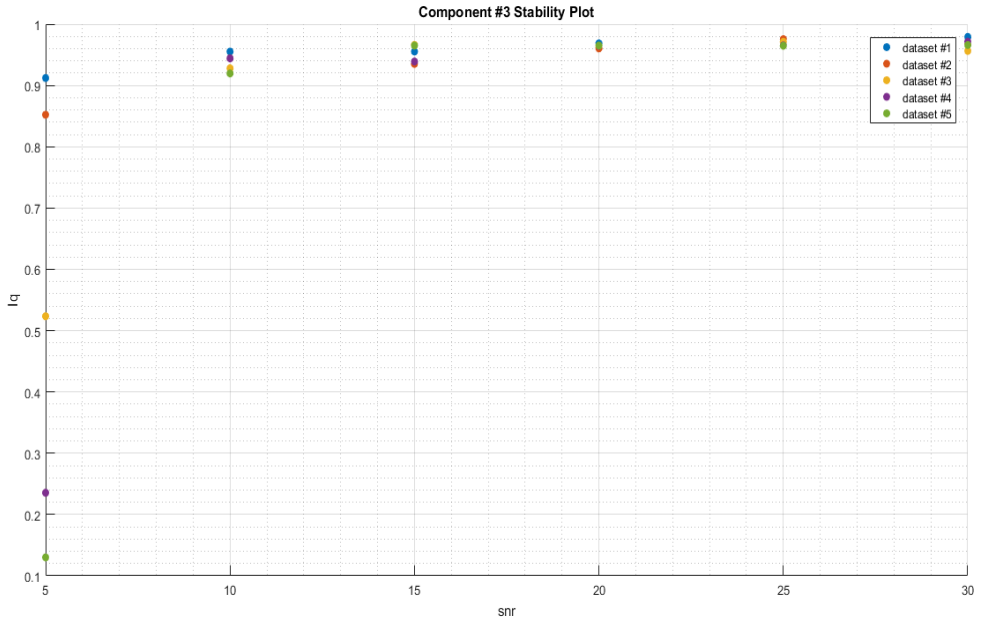


Figure 5.15: Iq stability values for the RSP signal component, for the five input datasets of the ICA with varying snr.

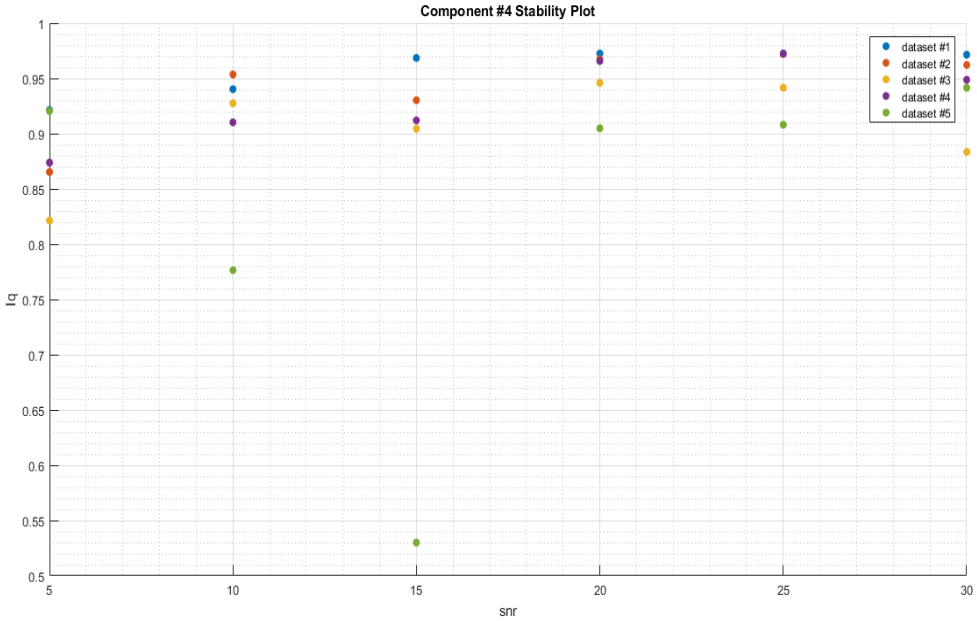


Figure 5.16: Iq stability values for the MBS signal component, for the five input datasets of the ICA with varying snr.

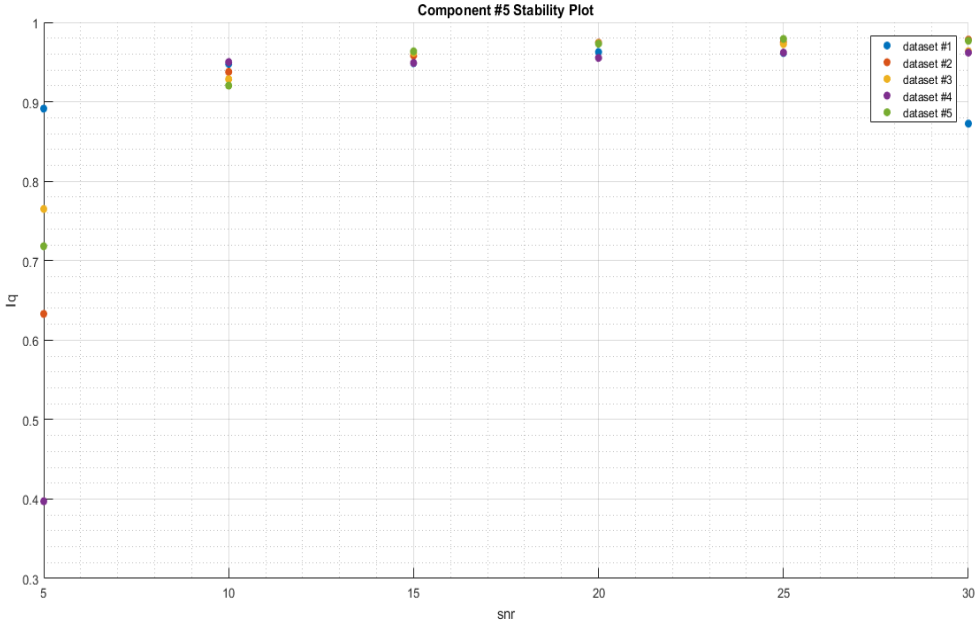


Figure 5.17: Iq stability values for the SSp signal component, for the five input datasets of the ICA with varying snr.

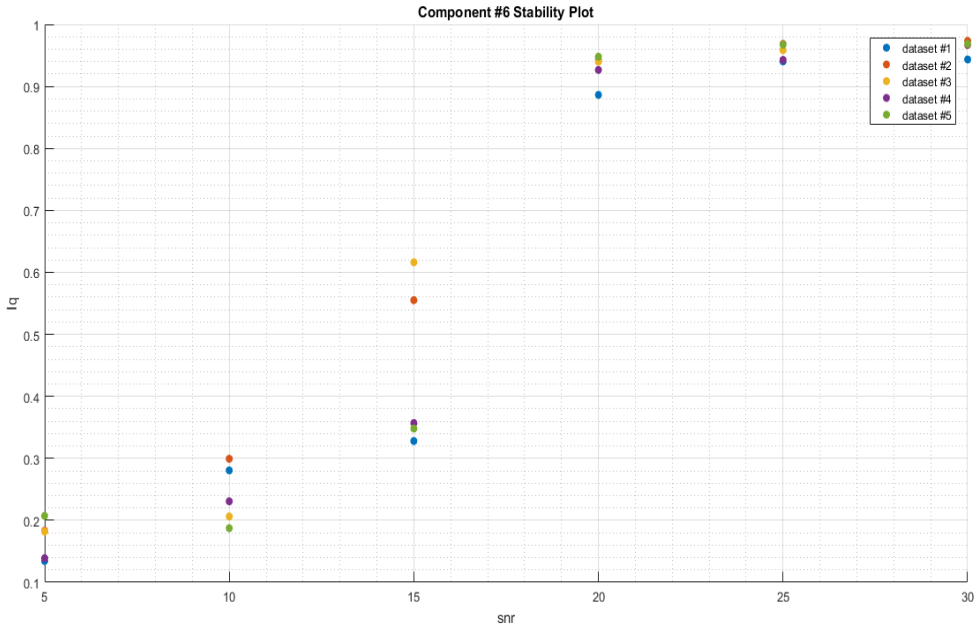


Figure 5.18: Iq stability values for the LGN signal component, for the five input datasets of the ICA with varying snr.

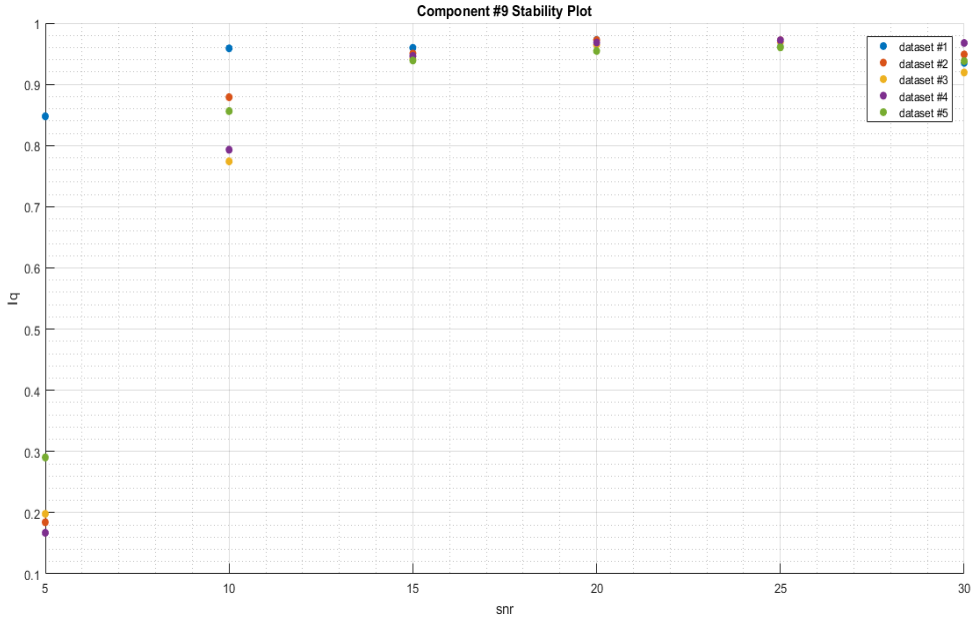


Figure 5.19: Iq stability values for the Hip signal component, for the five input datasets of the ICA with varying snr.

variable snr value analysis cases of the 14_14_53 data set are presented. In the first column of Figures 5.20 and 5.21, the reference components of the specific data set are presented (as shown in Figure 5.1) and then, in each of the next six columns, the evolution of these components as produced by the ICA and Icasso is presented, with a variable snr value of the input signal. The corresponding results for the remaining four datasets are listed in Appendix A.

5.2.2 Observations and Conclusions

From the plot of the percentage distributions of the stability values for the results of each data set (Figure 5.13) and also from the histograms of the stability values for the entire data set (Figure 5.14), the negative effect of increasing the input noise on the ICA results is evident , as the algorithm fails to detect all the stable independent signal components. It is interesting to note that the minimum component stability value detected by the ICA increases with an increasing snr, the maximum remains constant at values between 0.9 and 1, while one tenth of the samples show stability values above 0.9 for all datasets and for all snr values with the exception of snr=5. This fact indicates the existence of a small number of independent signal components that appear reliably, with high stability, in each data set. From Figures 5.15 to 5.19 and from the Appendix Figures 17 and 18, we note indicatively that the RSP component and the MBS, SSp and Hip signal components, have stability values above 0.9 for all (RSP component) or most datasets, for all snr values except 5.

In general for all spatial components of the 14_14_53 dataset we observe an average increase in stability as the value of snr increases, which is obviously expected theoretically. On the other hand, observing the evolution of the stability index of each component separately for a given fUS experiment, an upward trend is not always followed. This is likely due to a mismatch of the component of the given dataset with some of the components of the reference set of 14_14_53.

From Figures 5.20 and 5.21, we notice that each spatial signal component has a different tolerance to noise. We observe that the RSP, MBS, SSp and Hip components are distinct for all the different snr values we considered, while the LGN component is distinct for an snr value of 20 and over. We also observe that the ICA fails to detect the majority of independent signal components for an snr value equal to 5, which justifies the choice of the lower bound for this part of the analysis. It is also worth commenting that, while some signal components (e.g. component 10) are detected by the ICA for the same snr values as some of the more clear anatomical brain regions (LGN component), they do not show a solid anatomical form, which supports the theory that the signal sources detected by the ICA do not necessarily correspond to distinct anatomical regions of the brain but are part of a functional ensemble, without a clear anatomical mapping. Another explanation is of course the fact that the ICA often separates the main signal components into an unknown linear combination. However, functional ultrasound mapping the perfusion of the brain, whose functional diagram is related but not identical to the anatomical-physiological diagram of the brain, both

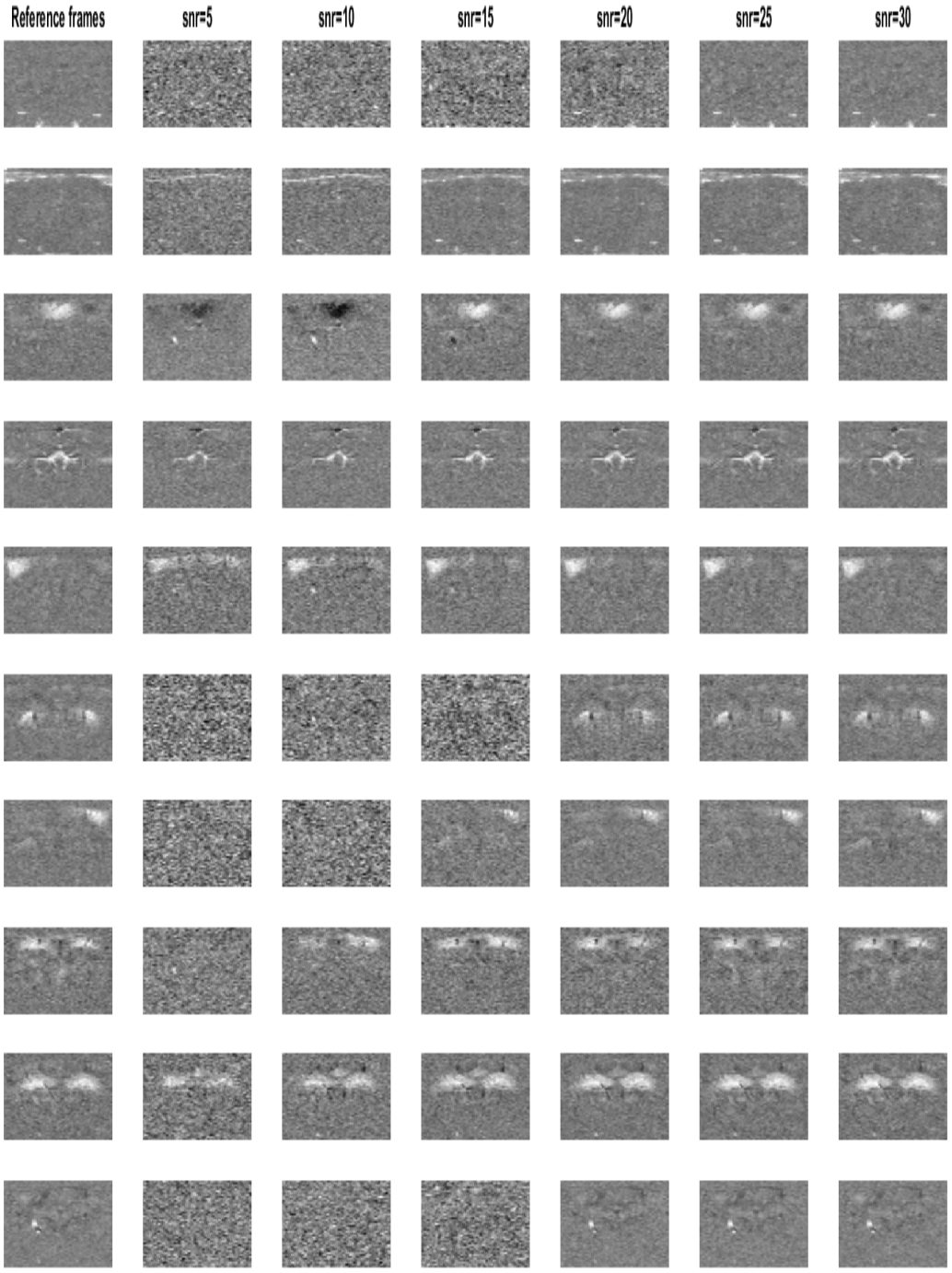


Figure 5.20: Presentation of the evolution of the first 10 out of the 20 principal spatial components of the dataset 14_14_53, produced by Icasso for a varying snr value of the input signal. In the first column we present the reference components, as obtained by Icasso with an input PCA dimension of 20, and in the following columns we display the results of the ICA algorithm for the different snr values of the input signal.

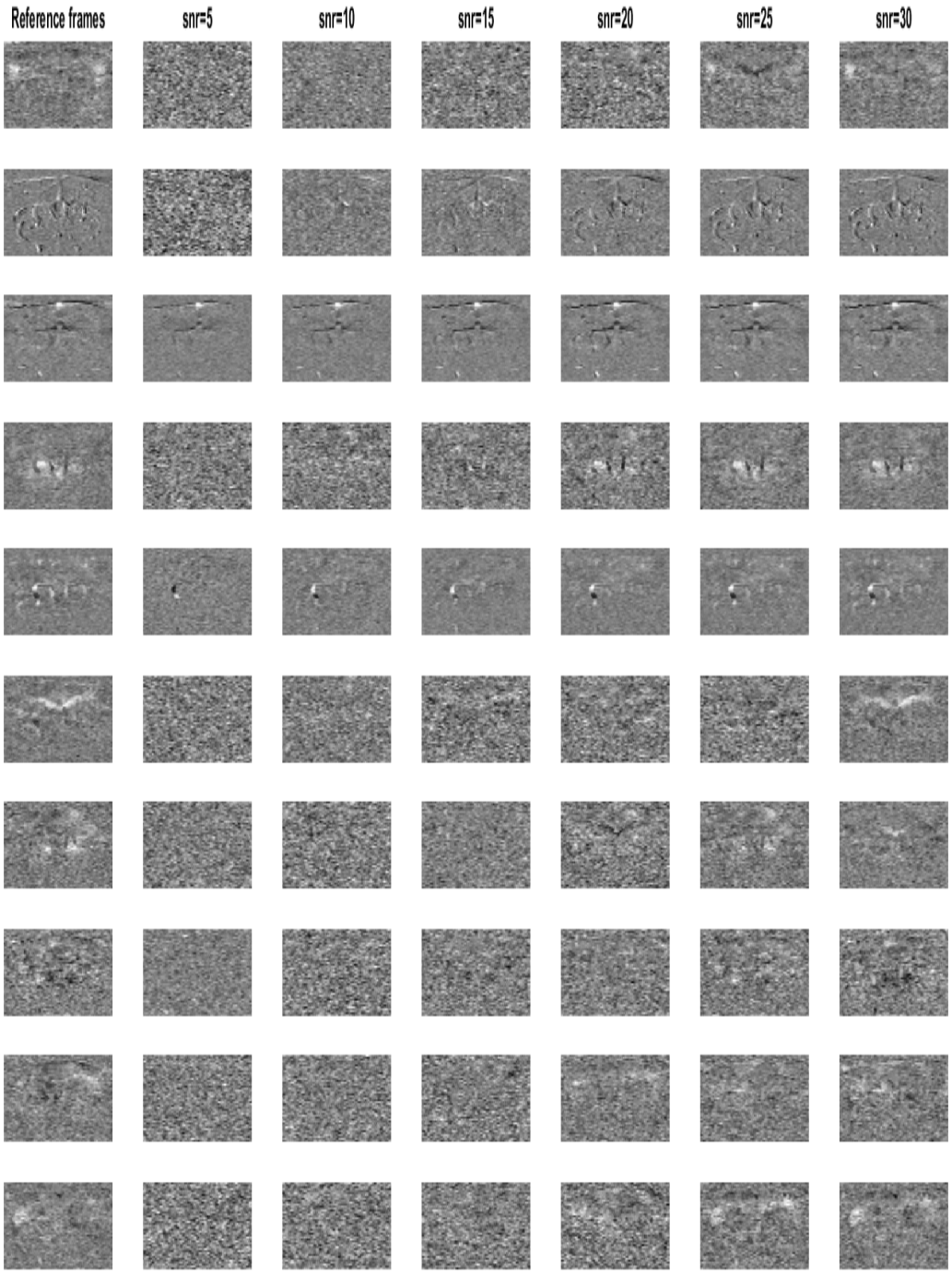


Figure 5.21: Presentation of the evolution of the next 10 out of the 20 principal spatial components of the dataset 14_14_53, produced by Icasso for a varying snr value of the input signal.

hypotheses are likely to be valid to some extent.

5.3 Dynamic Data Analysis

5.3.1 Experimental Results

In the third part of the analysis of the functional ultrasound data, we used a dynamic analysis approach. Taking the input signal after the preprocessing steps described above, we repeat the analysis using Icasso, in time frames of 200 samples ($\sim 42sec$). The parameters used for the ICA are: 20 signal components and a PCA dimension equal to 20. The number of resampling cycles used is 30, due to the long duration of the analytic experiments.

To match the components of the input signal, the reference frames of 100 components that have been calculated separately for each dataset were used as main reference components. In this way, the activation sequences were calculated for the five datasets, namely in each of the 19 time frames of the dataset, it was determined which of the 100 reference components were activated. To enable a better comparison of the results between the five datasets (derived from five independent repetitions of the same experiment, in the same experimental animal) we mapped the reference frames of sets 1,3, 4 and 5 to the reference frames of set 2. In this way, the following five activation maps were obtained where the "on" period is marked in yellow and the "off" period in blue. The numbering of the components on the y-axis corresponds to the numbering of the 100 reference components of the second data set (set 14_30_03), which are listed in Appendix A.

Figure 5.27 shows the histograms of the five experimental PDI acquisitions, with the number of independent components (out of the 100 reference components of the 14_-30_03 dataset) activated per total number of activations, with the maximum being all of the 19 frames of the dynamic analysis of the signal . From the shape of the histograms, it is clear that the majority of components identified through the ICA appear less than 5 times in the dynamic analysis frames, while there is also a constant number of components that appear almost throughout the duration of each experiment. Table 5.2 identifies the components in order of an increasing number of activation frames.

Analyzing the results of the activation patterns, we determined the components with the largest number of activation periods, as well as the number of components with zero activation during the experiment. It is noteworthy that out of the 100 reference components used to match the 20 components of each dynamic frame, the number of inactive components for the four datasets are 41, 31, 36 and 34 respectively.

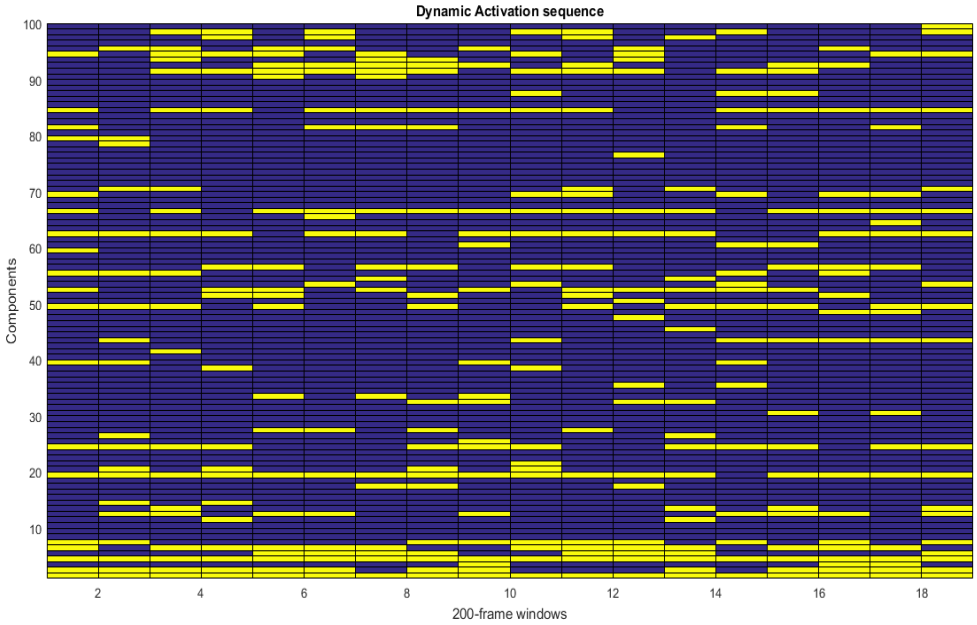


Figure 5.22: Illustration of the "on" and "off" periods of activation for the 100 spatial components of the fUS signal, using a moving time frame, for the 14_14_53 dataset.

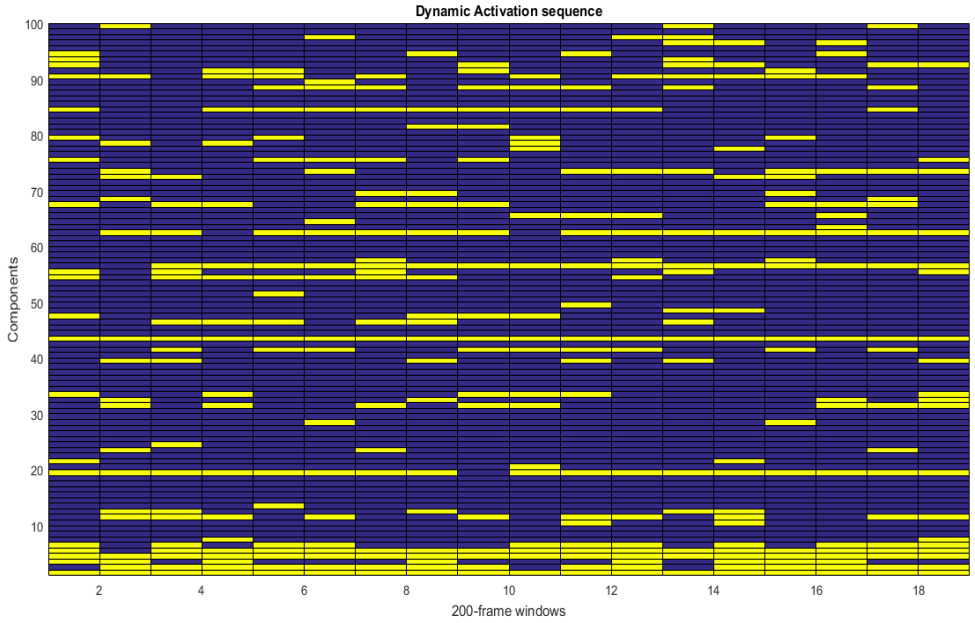


Figure 5.23: Illustration of the "on" and "off" periods of activation for the 100 spatial components of the fUS signal, using a moving time frame, for the 14_30_03 dataset.

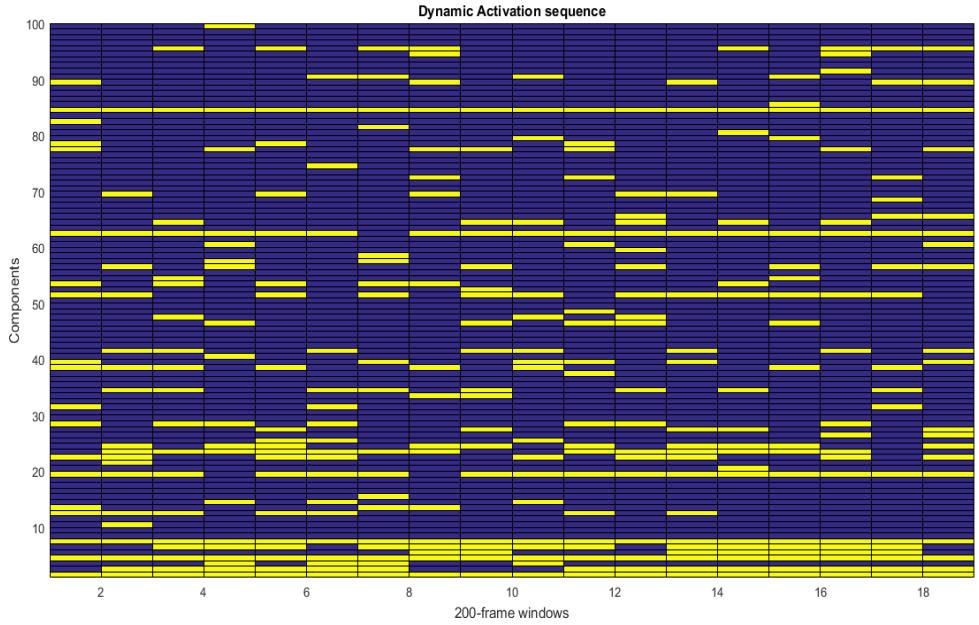


Figure 5.24: Illustration of the "on" and "off" periods of activation for the 100 spatial components of the fUS signal, using a moving time frame, for the 14_39_20 dataset.

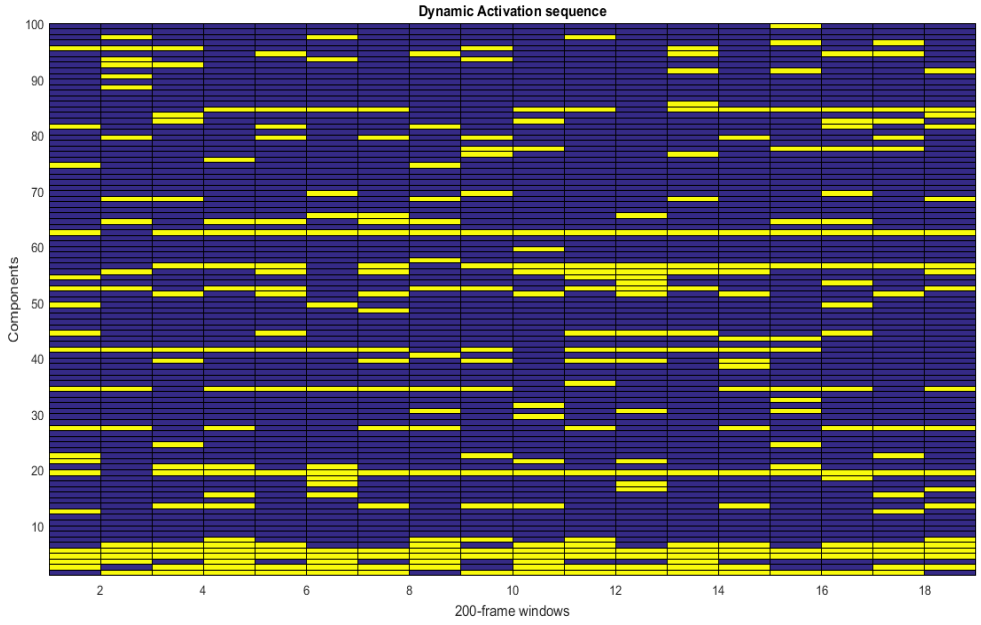


Figure 5.25: Illustration of the "on" and "off" periods of activation for the 100 spatial components of the fUS signal, using a moving time frame, for the 14_43_46 dataset.

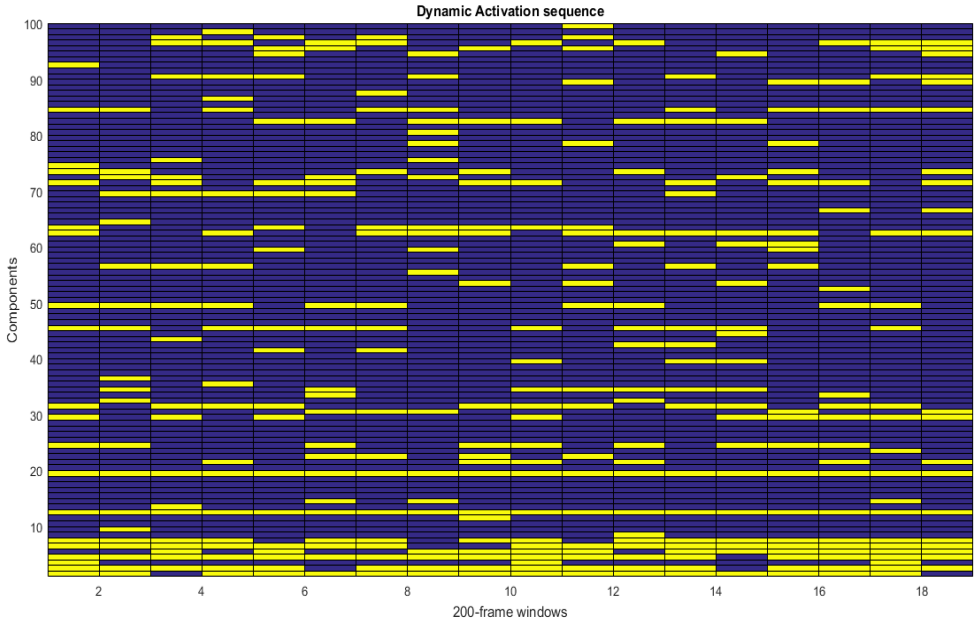


Figure 5.26: Illustration of the "on" and "off" periods of activation for the 100 spatial components of the fUS signal, using a moving time frame, for the 14_48_15 dataset.

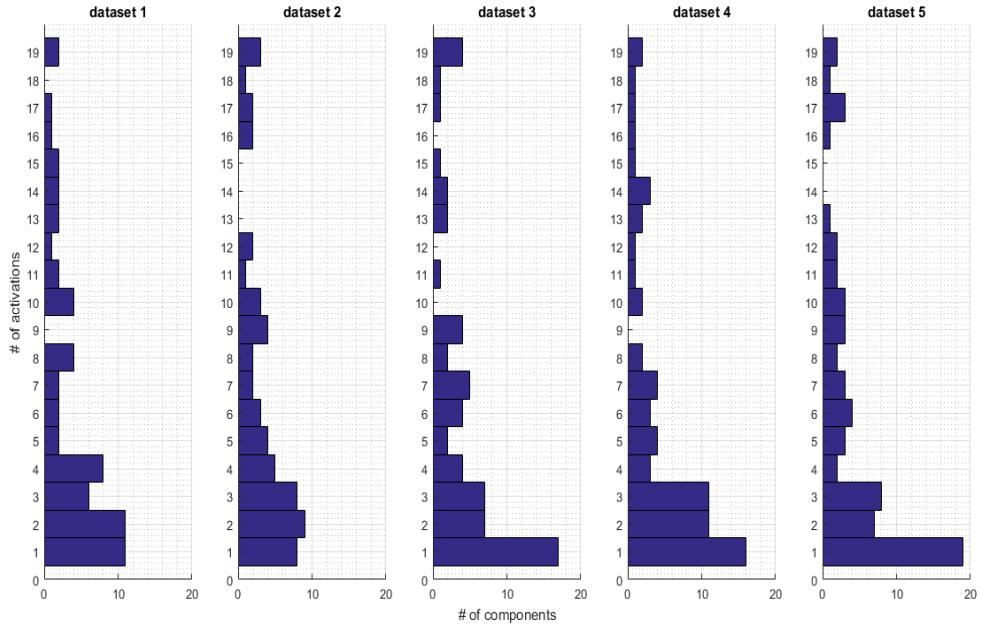
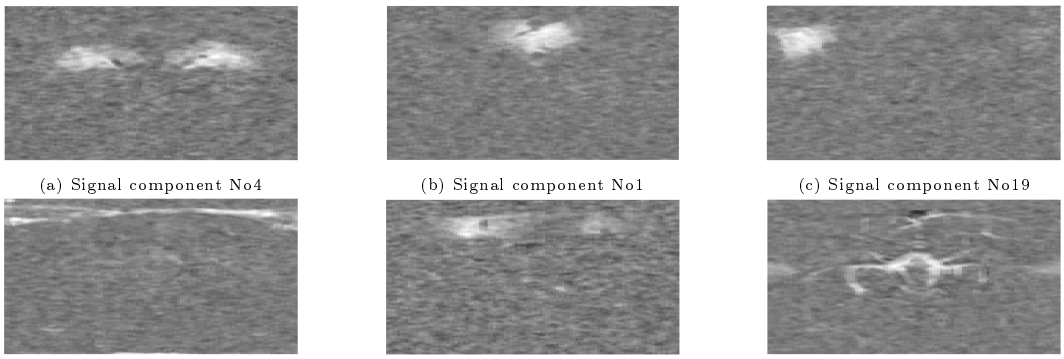


Figure 5.27: Histogram presentation for the five functional ultrasound datasets, depicting the number of activated components per total number of activations (number of activated time frames) during the dynamic analysis with a moving time frame.



(a) Signal component No4 (b) Signal component No1 (c) Signal component No19

(d) Signal component No62 (e) Signal component No5 (f) Signal component No2

Figure 5.28: Presentation of the six spatial signal components (numbered according to the set of 100 reference components of dataset 14_30_03, Appendix Figures 1 - 5) with the largest number of activation frames during the dynamic analysis of the five functional ultrasound datasets.

5.3.2 Observations and Conclusions

The dynamic analysis is performed in a moving time frame on the ICA input dataset. In the present work, a window of length $\sim 42sec$ was chosen due to the long running time of the algorithm, in order to identify the same components using a shorter time frame. In this particular analysis step, we are mainly interested in investigating the capabilities of the ICA in identifying principal signal components in subsets of the data set, and the order of activation of the various components/regions during the experiment.

Using the reference frames to match the ICA results, a set of 100 principal components is indicated for each experiment. Out of these components, a number (about 30 to 40) did not appear at all in the dynamic analysis. This number of active components gives us a rough idea of the size of the brain's actual set of activation regions that can be identified using the ICA of fUS data. Due to the error involved in the matching algorithm, this number is less than 60 - 70 and also depends heavily on the size of the time frame. However, it is encouraging that the number of active components does not vary widely between the independent datasets.

The size of the time frame is a crucial parameter of the dynamic analysis, because it defines the scale of the transient events we can observe in the brain. Table 5.2 shows the results regarding the activeted regions of the brain, obtained from the dynamic analysis. From these results, the five most active components are as follows (the numbering was done based on the reference frames of the 14_30_03 dataset): 4, 1, 19, 62, 5, 2 (Figure 5.28).

Based on the anatomical atlas[1] for the imaging of the mouse brain, we observe that components 4, 1, 19, 5 correspond to distinct regions of the mouse neural network. In particular, component #4, which shows the greatest temporal activation, roughly corresponds to the anatomical region of the hippocampus, which according to literature

[1] https://portal.brain-map.org/

plays a central role in the cognitive function of the brain. Component #1 corresponds to the posterior splenial cortex of the brain, which is part of the Default Mode Network (DMN). It has to do with the animal's spatial perception of its surroundings. Component #19 is part of the somatosensory area of the brain and is related to the processing of the stimulus in the mouse's whisker, which is a basic source of spatial information and allows the animal to perceive and react to features of his environment. Component #5 corresponds to the posterior parietal cortex areas, which - in the present work - are related to eye movement and spatial perception. Their function is closely related to the function of the hippocampus.

Components #62 and #2, at first glance, do not correspond to distinct anatomical regions of the brain. However, we observe from the image of the total brain perfusion that they exactly trace the vascular system in the center the posterior regions of the cerebral cortex we are studying. This allows us to conclude that the processing of functional ultrasound data using the ICA, results in activation centers both in the form of distinct neuroanatomical regions and as parts of the brain perfusion network. This element is of great interest as it gives the opportunity for further research on the results of the ICA, taking into account that we are given the possibility to monitor the brain function from the point of view of two different biological systems: the nervous and the circulatory.

Base on the histograms for each dataset in Figure 5.27 and on the tables that follow, we notice that the five datasets show quantitatively similar behavior during the dynamic analysis of their data. They also present a qualitative similarity in terms of the components that are activated with greater frequency. But it is much more difficult to draw firm conclusions as the frequency of activation decreases, for the relationship between the sets ceases to be so clear. In Chapter 6 we will address some possible future research directions regarding the dynamic analysis of functional ultrasound data.

Total $\#frames = 19$	19 frames on	18 frames on	17 frames on	16 frames on
Dataset 14_14_53	1, 4	-	19	66
Dataset 14_30_03	1, 4, 43	19	5, 56	2, 62
Dataset 14_39_20	1, 4, 28, 91	62	19	-
Dataset 14_43_46	4, 5	62	19	1
Dataset 14_48_15	12, 19	4	1, 2, 6	66

(a) Components with 16 to 19 activated time frames

Total $\#frames = 19$	15 frames on	14 frames on	13 frames on	12 frames on
Dataset 14_14_53	84, 100	2, 62	6, 24	91
Dataset 14_30_03	-	-	-	6, 90
Dataset 14_39_20	2	23, 100	6, 51	-
Dataset 14_43_46	56	2, 41, 100	6, 34	84
Dataset 14_48_15	-	-	62	5, 45

(b) Components with 12 to 15 activated time frames

Total $\#frames = 19$	11 frames on	10 frames on	9 frames on	8 frames on
Dataset 14_14_53	7, 49	12, 52, 56, 94	-	**5, 43, 92, 98**
Dataset 14_30_03	84	11, 67, 100	3, 41, 73, 88	31, 54
Dataset 14_39_20	24	-	22, 34, 38, 95	41, 77
Dataset 14_43_46	27	52, 55	-	51, 64
Dataset 14_48_15	31, 84	29, 49, 71	24, 82, 96	73, 90

(c) Components with 8 to 11 activated time frames

Total $\#frames = 19$	7 frames on	6 frames on	5 frames on
Dataset 14_14_53	69,95	51, 81	55, 70
Dataset 14_30_03	33, 39	46, 75, 92	12, 55, 65, 91
Dataset 14_39_20	5, 12, 28, 39, 56	53, 64, 69, 89	27, 46
Dataset 14_43_46	3, 13, 39, 44	68, 79, 94	7, 77, 81, 95
Dataset 14_48_15	21, 34, 63	30, 56, 69, 95	72, 94, 97

(d) Components with 5 to 7 activated time frames

Total $\#frames = 19$	4 frames on	3 frames on
Dataset 14_14_53	**13, 20, 27, 32, 39, 53, 93, 97**	3, 17, 33, 48, 60, 87
Dataset 14_30_03	**32, 47, 72, 79, 94**	23, 57, 69, 77, 78, 96, 97, 99
Dataset 14_39_20	3, 25, 78, 90	13, 14, 31, 47, 60, 65, 72
Dataset 14_43_46	20, 21, 82	15, 22, 30, 49, 54, 65, 69, 91, 92, 97
Dataset 14_48_15	22, 89	3, 14, 39, 53, 59, 60, 78, 100

(e) Components with 3 to 4 activated time frames

Total $\#frames = 19$	2 frames on
Dataset 14_14_53	11, 14, 26, 30, 35, 38, 41, 54, 79, 90, 99
Dataset 14_30_03	7, 10, 21, 28, 48, 49, 68, 81, 93
Dataset 14_39_20	26, 33, 54, 57, 79, 82, 94
Dataset 14_43_46	12, 16, 17, 18, 24, 43, 53, 74, 76, 83, 96
Dataset 14_48_15	32, 33, 41, 42, 66, 75, 98

(f) Components with 2 activated time frames

Total $\#frames = 19$	1 frame on
Dataset 14_14_53	**21, 25, 45, 47, 50, 59, 64, 65, 76, 78, 80**
Dataset 14_30_03	**13, 20, 22, 24, 51, 63, 64, 89**
Dataset 14_39_20	**10, 15, 20, 21, 37, 40, 48, 52, 58, 59, 68, 74, 80, 81, 85, 91, 99**
Dataset 14_43_46	**8, 29, 31, 32, 35, 38, 40, 48, 57, 59, 72, 75, 85, 88, 90, 99**
Dataset 14_48_15	**8, 9, 11, 13, 23, 35, 36, 43, 44, 52, 55, 64, 70, 74, 80, 86, 87, 92, 99**

(g) Components with 1 activated time frame

Table 5.2: Presentation of the spatial signal components per number of activation frames, during the dynamic analysis experiment of the functional ultrasound data. Component numbering follows the reference set of 100 signal components of dataset 14_30_03, presented in Appendix A.

Bibliography

[1] N. K. Logothetis, «What we can do and what we cannot do with fmri», *Nature*, vol. 453, no. 7197, pp. 869–878, 2008.

[2] E. Macé, 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, 2011.

[3] T. Deffieux, C. Demene, M. Pernot, and M. Tanter, «Functional ultrasound neuroimaging: A review of the preclinical and clinical state of the art», *Current opinion in neurobiology*, vol. 50, pp. 128–135, 2018.

[4] H. W. Kuhn, «The hungarian method for the assignment problem», *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.