BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DATA PROCESSING IN REAL-TIME FMRI NEUROFEEDBACK

ZPRACOVÁNÍ DAT PRO FMRI NEUROFEEDBACK

BACHELOR'S THESIS
BAKALÁŘSKÁ PRÁCE

AUTHOR — **Bc. Martin Bečička**
AUTOR PRÁCE

SUPERVISOR — **Ing. Martin Lamoš, Ph.D.**
VEDOUCÍ PRÁCE

BRNO 2019

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

Bakalářská práce

bakalářský studijní obor **Biomedicínská technika a bioinformatika**
Ústav biomedicínského inženýrství

Student: Bc. Martin Bečička ***ID:*** 174191
Ročník: 3 ***Akademický rok:*** 2018/19

NÁZEV TÉMATU:

Zpracování dat pro fMRI neurofeedback

POKYNY PRO VYPRACOVÁNÍ:

1) Proveďte literární rešerši metody funkčního zobrazování pomocí magnetické rezonance (fMRI) se zaměřením zejména na možnosti zpracování dat v reálném čase pro využití v rámci neurofeedbacku. 2) Seznamte se s dosavadním řešením zpracování fMRI dat na pracovišti Laboratoře multimodálního a funkčního zobrazování CEITEC MU. 3) Navrhněte postup zpracování dat pro fMRI neurofeedback s ohledem na dosažení lepšího poměru signál-šum. 4) Realizujte navržený postup a ověřte jeho funkčnost - každá nově příchozí fMRI data musejí být zpracována v čase pod jednu sekundu. 5) Proveďte porovnání s dosavadním řešením vzhledem ke kvalitě extrakce relevantního BOLD signálu z oblasti zájmu. 6) Diskutujte dosažené výsledky a zhodnoťte účinnost a využitelnost řešení.

DOPORUČENÁ LITERATURA:

[1] PARET, Christian, Rosemarie KLUETSCH, Jenny ZAEHRINGER, Matthias RUF, Traute DEMIRAKCA, Martin BOHUS, Gabriele ENDE and Christian SCHMAHL, 2016. Alterations of amygdala-prefrontal connectivity with real-time fMRI neurofeedback in BPD patients. Social Cognitive and Affective Neuroscience. 11(6), 952–960. ISSN 17495024.

[2] YOUNG, Kymberly D., Vadim ZOTEV, Raquel PHILLIPS, Masaya MISAKI, Han YUAN, Wayne C. DREVETS and Jerzy BODURKA, 2014. Real-Time fMRI Neurofeedback Training of Amygdala Activity in Patients with Major Depressive Disorder. PLoS ONE. 9(2), e88785. ISSN 1932-6203.

Termín zadání: 4.2.2019 ***Termín odevzdání:*** 24.5.2019

Vedoucí práce: Ing. Martin Lamoš, Ph.D.
Konzultant:

prof. Ing. Ivo Provazník, Ph.D.
předseda oborové rady

ABSTRACT

The presented thesis deals with real-time digital filtering of fMRI neurofeedback data. It analyzes currently used solution at CEITEC MU chiefly in respect to finding ways to shorten the delay at the beginning of each neurofeedback block which is introduced by digital filtering. Current solution uses extended Kalman filter mainly for its real-time and smoothing properties. Analysis of 150 individual neurofeedback blocks yielded true learning period of Kalman filter which has been found to be significantly shorter than is set in the current solution. Different options to further reduce the transient period have been explored and short moving average filter has been chosen as an optimal trade-off between transient period, filter delay and its smoothing properties.

KEYWORDS

Functional MRI neurofeedback, extended Kalman filter, real-time digital filtering, moving average filter

ABSTRAKT

Tato práce se zabývá digitálním filtrováním dat získaných z fMRI neurofeedbacku v reálném čase. Práce analyzuje dosavadní řešení používané v CEITEC MU, se zaměřením na zkrácení prodlení na začátku každého neurofeedback bloku, které je způsobeno digitálním filtrováním. Dosavadní řešení používá, hlavně pro jeho online a vyhlazovací vlastnosti, nelineární Kalmánův filtr. Analýzou 150 průběhu fMRI neurofeedback sezení byla zjištěna dolní hranice, kterou nelineární Kalmánův filtr potřebuje k naučení. Počet potřebných vzorků je významně menši než je nastaveno v dosavadním řešení. Další možnosti zkrácení prodlení byly prozkoumány a klouzavý průměrovací filtr byl vybrán jako optimální kompromis mezi dobou prodlení, zpoždění filtru a jeho vyhlazovacími vlastnostmi.

KLÍČOVÁ SLOVA

Funkční MRI neurofeedback, nelineární Kalmánův filter, real-time digitální filtrování, průměrovací filtr

BEČIČKA, Martin. *Data processing in real-time fMRI neurofeedback*. Brno, 2019, 47 p. Bachelor's Thesis. Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Biomedical Engineering. Advised by Ing. Martin Lamoš, Ph.D.

DECLARATION

I declare that I have written the Bachelor's Thesis titled "Data processing in real-time fMRI neurofeedback" independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the thesis and listed in the comprehensive bibliography at the end of the thesis.

As the author I furthermore declare that, with respect to the creation of this Bachelor's Thesis, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll., Section 2, Head VI, Part 4.

Brno ....

author's signature

ACKNOWLEDGEMENT

I would like to thank my advisor, Ing. Martin Lamoš, Ph.D., for consulting, patience and contributions which had lead this thesis to its end.

Brno . . . .

author's signature

Contents

Introduction **9**

1 Functional MRI **10**
1.1 Working principle . . . 10
1.2 The hemodynamic response . . . 11
1.3 Characteristics of BOLD signal . . . 13
1.4 Temporal and spatial properties . . . 14

2 Post-hoc analysis of fMRI data **15**
2.1 Preprocessing of fMRI data . . . 15
2.1.1 Slice-timing correction . . . 15
2.1.2 Realignment . . . 15
2.1.3 Coregistration . . . 16
2.1.4 Normalization . . . 16
2.1.5 Spatial smoothing . . . 17
2.1.6 Temporal filtering . . . 17
2.1.7 Outlier removal . . . 17
2.2 Statistical analysis . . . 18
2.2.1 T-statistics . . . 18
2.2.2 General linear model . . . 18

3 Real-time fMRI neurofeedback **21**
3.1 Current neurofeedback solution . . . 22

4 Filtering of temporal data **26**
4.1 Finite impulse response filters . . . 26
4.1.1 Moving average filter . . . 27
4.2 Infinite impulse response filters . . . 27
4.3 Kalman Filter . . . 29
4.4 Adaptive filters . . . 31

5 Proposed approach for online data filtering **32**
5.1 Analysis of extended Kalman filter . . . 32
5.2 Temporary filter . . . 35
5.3 Spike detection . . . 36

6 Results **38**

7 Conclusion **41**

Bibliography **43**

List of symbols, physical constants and abbreviations **47**

List of Figures

1.1 BOLD signal and its indirect relationship to neuronal activity 11
1.2 Changes in concentrations of oxyhemoglobin and deoxyhemoglobin after a stimulus 12
1.3 Hemodynamic response function 13
2.1 General linear model for a single voxel 19
2.2 General linear model and information about experimental design . . . 20
3.1 Flow diagram of rtfMRI neurofeedback 21
3.2 Data processing for fMRI neurofeedback 23
3.3 Functional MRI neurofeedback signal before and after extended Kalman filtering 24
3.4 Output from current neurofeedback solution 25
4.1 Inner workings of FIR digital filter 26
4.2 Inner workings of IIR digital filter 28
4.3 Block diagram of a generic adaptive filter 31
5.1 Single-sided amplitude spectrum of a typical fMRI neurofeedack signal before and after Kalman filtering 33
5.2 Deviation of real-time Kalman filters compared to its offline counterparts 34
5.3 Magnitude and phase response of length 3 moving average filter . . . 36
5.4 Spike detection and removal algorithm in action 37
6.1 Outcome of the proposed approach for online data filtering 38
6.2 Comparison of filtering properties of EKF and MA 39
6.3 Amplitude spectrum of neurofeedback signal before and after EK and MA filtering 40

Introduction

The human brain is the center of the nervous system. As such it is responsible for many voluntary and involuntary actions as well as higher order brain functions that are largely unique to humans.[1]

Functional magnetic resonance imaging (fMRI) uses localized changes in blood flow to non-invasively analyze the functioning of the human brain. Recent advancements in the functional magnetic resonance imaging technology and data processing has enabled observations of these changes in real-time.[1]

Neurofeedback is a process which indicates to the participant current information about the functioning of his brain. This information is obtained through sensors or devices. Recently, there has been a growing interest in research using real-time fMRI neurofeedback and its possible therapeutic properties. A growing body of evidence suggests it is possible to train individuals to upregulate or downregulate certain brain areas. Implications of these claims are projected in the following studies. Two such studies implicated amygdala dysregulation as the cause of mood disorders e.g. major depressive disorder [2], [3]. Results suggest that real-time fMRI neurofeedback may be an effective way to treat these disorders. Yet another study suggests neurofeedback may not only be used in correcting abnormally functioning brain areas but can also improve brain efficiency above standard levels.[4]

As the possibility of applying rtfMRI for neurofeedback has only recently started to be properly explored, there is a lack of consensus among the scientific community about the most effective ways to process the rtfMRI data. The focus of this thesis is real-time data filtering and improvement of delay introduced by transient learning period of a currently used solution at the Multimodal and Functional Imaging Laboratory (MAFIL) which is part of Central European Institute of Technology project (CEITEC) located at Masaryk University.

1 Functional MRI

Functional magnetic resonance (fMRI) as its name implies is based on MRI. But unlike MRI, which can show us detailed structural images of our bodies, fMRI can show us images of metabolic activity happening inside these structures such as short-term physiological changes which correlate with neuronal activity. This allows the study of workings of the human brain. Studies utilizing fMRI aim to localize, find out which or determine the degree to which certain areas in the brain are active while performing a certain task. Though only one of the methods used in the field of non-invasive functional neuroimaging, fMRI has quickly become a gold standard soon after its conception in the early 1990s.[5][6]

1.1 Working principle

Functional MRI obtains information about neuronal activity through indirect means, specifically by detecting changes associated with blood flow. Different techniques are utilized to obtain functional images of the brain such as blood oxygen level dependent (BOLD) contrast, arterial spin labeling or perfusion MRI. As the rest of this thesis is concerned with the processing of images obtained through BOLD contrast and since it is also the most widespread fMRI technique, it will be explained more in-depth.[5]

BOLD imaging as its name suggests depends on the level of oxygen in the blood. While in transport oxygen molecules are bound to hemoglobin which acts as a carrier. When bound the resulting compound is called oxyhemoglobin. Hemoglobin without bound oxygen is called deoxyhemoglobin. Oxyhemoglobin and deoxyhemoglobin have different magnetic properties and so behave differently when exposed to an external magnetic field. Oxyhemoglobin has all of its electrons paired and thus has zero magnetic moment. It is weakly diamagnetic with slight repulsion force from magnetic fields. On the other hand, fully deoxygenated hemoglobin is paramagnetic because it has four unpaired electrons and thus a significant magnetic moment [5]. In the study conducted by Pauling and Coryell in 1936, it was discovered that the magnetic susceptibility of deoxygenated blood is about 20% greater than of oxygenated blood. In other words, deoxyhemoglobin acts as an endogenous paramagnetic contrast agent.[5]

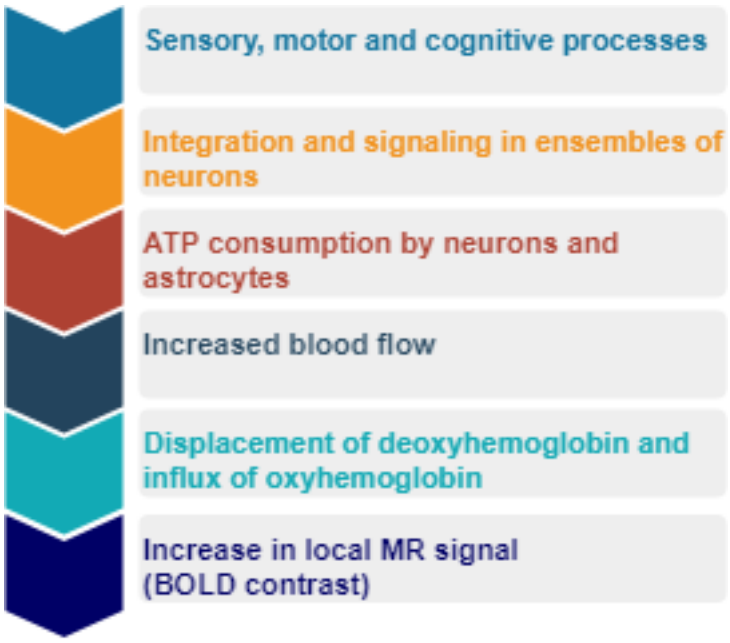

Fig. 1.1: Indirect relationship between BOLD signal and information processing in cognitive, motor and sensory processes. Adapted from [5].

Here is where the strong external magnetic field comes into play. Significant paramagnetic properties of deoxygenated hemoglobin introduce local magnetic field distortions affecting the homogeneity of the field and so changing the relaxation time of nearby hydrogen nuclei. MR signal intensity is reduced. The opposite is also true if deoxyhemoglobin is replaced by oxyhemoglobin, for example by increased blood flow, the local MR signal increases.[7]

1.2 The hemodynamic response

The response of the vascular system to stimuli is known as a hemodynamic response. The hemodynamic response is triggered by information processing which requires activation of ensembles of neurons increasing their metabolic requirements. Neurons respond to an increase in demand by using more adenosine triphosphate, ATP for short. The energy in the form of ATP is used for example in the creation and propagation of action potentials, the process of releasing neurotransmitters, their reception and uptake back into neurons or glial cells. Since the supplies of ATP in the brain are rather limited more ATP needs to be constantly generated from glucose by aerobic or anaerobic glycolysis. Aerobic glycolysis is by far the more efficient one and requires oxygen as well as glucose to be present.[5]

Both nutrients need to be continuously supplied through the vascular system. In case of increased demand for nutrients activated neurons release vasoactive substances

that cause dilation of nearby blood vessels. The increase in diameter reduces vessels resistance to flow. Nevertheless, these are fairly local changes and other steps are necessary to control cerebral blood flow. Higher-resistance arterioles located upstream need to be signaled as well. So there is a need for coordination between local blood flow control mechanisms and upstream control mechanisms. As a result of this blood flow to the activated area is increased, this constitutes the hemodynamic response.[5]

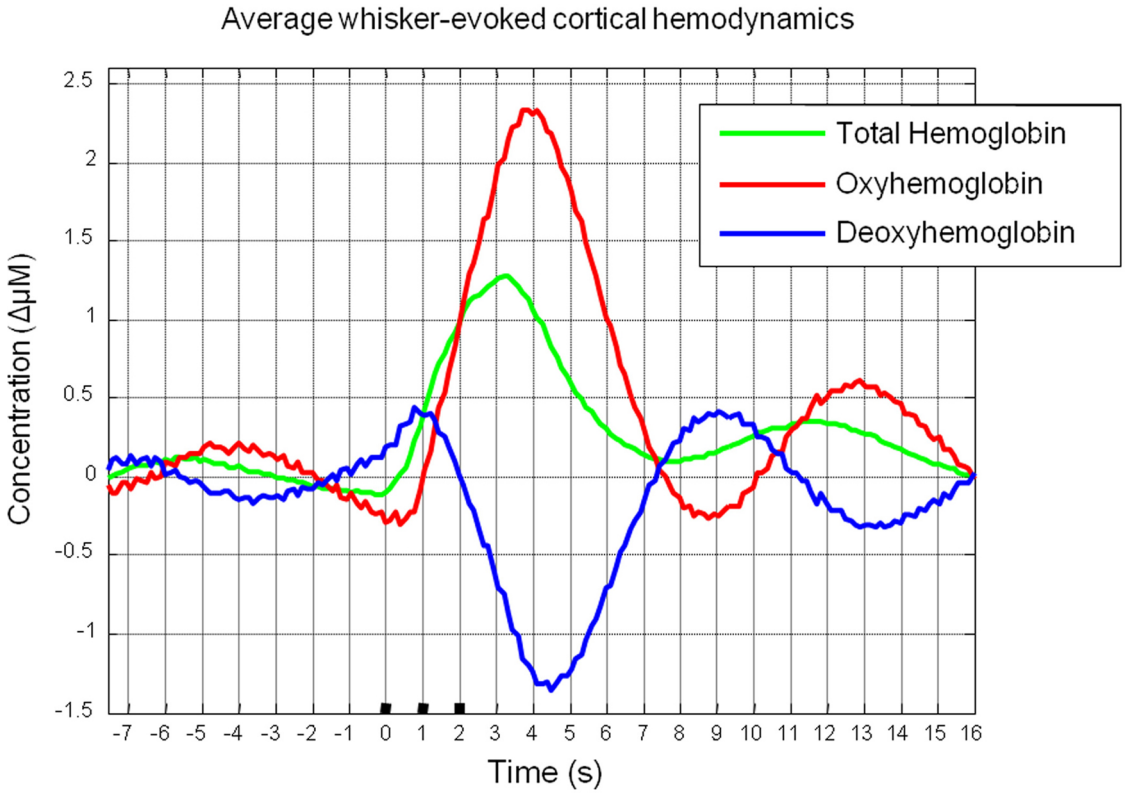

Fig. 1.2: Changes in concentrations of oxyhemoglobin and deoxyhemoglobin after a single brief stimuli commencing at t = 0 and ending at t = 2.[8]

It is important to note a few limitations arising from the indirect relationship between BOLD signal and neuronal activity. Through experiments, it was observed that a hemodynamic response also occurs up to a few millimeters from the neural activity and the subsequent epicenter of the hemodynamic response. To reword the finding the hemodynamic response occurs over a larger area than that in which increased neural activity was really located. This highlights two important limitations of fMRI. The measured signal will copy the architecture of the vascular system. So activation of certain areas can be caused by venous drainage, then local activation. Second, the smallest diameter of a unit able to individually regulate blood flow is in the order of mm thus the resolution of fMRI is ultimately limited by the aforementioned architecture.[5]

1.3 Characteristics of BOLD signal

Increase in activity causes an increase in the local concentration of deoxyhemoglobin and as such reduction in fMRI signal should be observed. Nevertheless, in other experiments conducted by Ogawa and colleagues (1990), it was observed that the fMRI signal actually increases with neuronal activation. In the figure 1.3 a change in BOLD signal following a single brief stimulus can be observed.

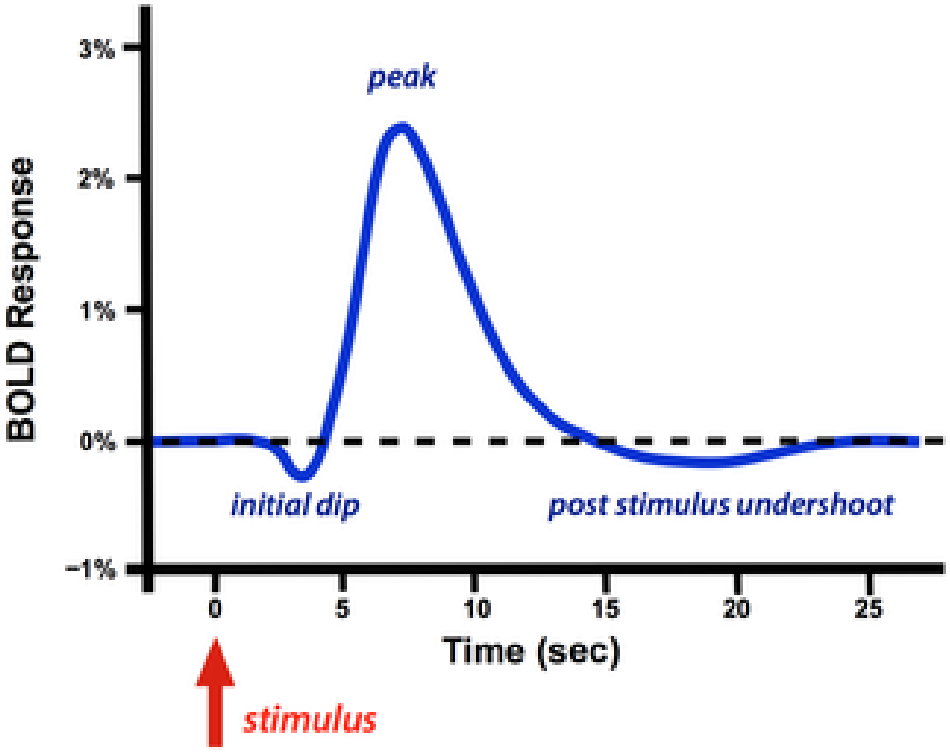

Fig. 1.3: Hemodynamic response function to a single brief stimuli.[9]

The initial dip may not be observed in all measurements and is more frequently observed in high fields (e.g., > 4 T) scanners. The mechanism of origin is still an object of discussion. One of two most influential hypothesis advocated by Menon and colleagues states that the cause is the local increase in deoxyhemoglobin following neuronal activation.[10][11]

Next, the signal increases above baseline levels generally in 2 - 5 s and peaks in 5-15 s after the onset of the stimulus. The vascular system is signaled and there is an increase in the flow of oxygenated blood to the area. The inflow of oxyhemoglobin is greater than the extraction and accumulated deoxyhemoglobin gets washed away at an increased rate. Both of these mechanisms cause an increase in the ratio of oxyhemoglobin to deoxyhemoglobin.[5]

Finally, the signal decreases even slightly under baseline levels. As with the initial dip, the underlying mechanism is still disputed but the most predominant one being delayed vascular compliance. First decrease in diameters of venular, then arterioral and later capillary. In other words, cerebral blood flow decreases more rapidly than blood volume causing a greater increase in the amount of oxyhemoglobin than during baseline levels. So the MR signal is also below baseline levels.[12]

1.4 Temporal and spatial properties

Spatial resolution is determined by the voxel size and is closely related to the experiment needs. "A typical voxel size in human fMRI is $3 \times 3 \times 3$ mm^3, which spans multiple cortical columns and usually covers the entire cortical thickness with its associated vasculature."[13] Studies that aim to explore the entire brain might use voxels as big as 4 - 5 mm^3. On the other hand studies concerning themselves only with a certain part of the brain can collect signal only from that part and gain higher resolution by using voxels 1 mm^3 in size.[5]

The signal measured depends upon the change in the entire voxel thus if smaller voxels e.g., two times smaller are used the signal will also be two times smaller. If conditions are perfect, the perceived change in BOLD signal following a stimulus is only about 2 - 3 %. Halving the signal means a considerable increase in SNR.[5]

Another complication arising from using smaller voxel sizes is the increase in acquisition time. If voxels half as large are used, acquisition times can double or even quadruple depending on a scanner and pulse sequence used.[5]

Temporal resolution is determined by the time to obtain one image of the desired region, also known as repetition time (TR). Depending upon the experiment needs the TR may range from 0.5 to 4 s. As was observed by Kannurpatti and colleagues the expected frequency range for the BOLD signal is approximately 0.01-0.125Hz.[14] So the phenomena in question last about 10 seconds and more thus there is not much gain in shortening TR even more. The gain of information about a signal quickly falls off at TR shorter than 0.5 - 1 s. Therefore longer TRs are generally used and the rest of the signal may be interpolated without significant loss of information. Shorter TRs also require smaller flip angles to be used thus reducing MR signal and may reduce the spatial resolution. Since a scanner can only obtain a set amount of slices for a set amount of time, halving the TR time may mean half the amount of slices per given time.[5]

2 Post-hoc analysis of fMRI data

Offline analysis, otherwise known as post-hoc analysis, serves to analyze already collected data. As such, it is time-insensitive and complicated algorithms can be run to ensure the best quality of the results. Post-hoc analysis can be split into data preprocessing and statistical analysis.

2.1 Preprocessing of fMRI data

Raw fMRI data needs to be further processed after image reconstruction and before they can be statistically analyzed. This step is known as preprocessing. One of the main reasons for preprocessing is low SNR, the BOLD effect changes the acquired signal by only about 2-3%. Another reason is that during the acquisition of fMRI data subjects can move slightly, heartbeat and respiration cause physiological oscillations, inhomogeneities in magnetic field cause distortions and there may be a difference in time acquisition of individual slices. These variabilities can be so severe they can entirely obscure the useful signal. Preprocessing methods focus on improving SNR, realigning, trying to minimize the impact of acquisition techniques and last but not least on mapping acquired images to standardized space.[5][15]

2.1.1 Slice-timing correction

Spatial gradients in fMRI scanners limit the excitation pulse to a single slice (and a single voxel). Most pulse sequences first collect all the odd slices and only then collects all the even slices. This is known as interleaved slice acquisition and is frequently implemented to minimize the influence of excitation pulses upon adjacent slices. For example, with TR of 3 s and 12 slices per imaging area, slice 2 would be acquired about 1.5 s later. This would lead to big timing differences between slides and it would impact statistical analysis which uses relationships between experimental hypotheses and measured data like GLM. Temporal interpolation is often used to correct for the slice-timing differences. Nearby time points are used to estimate the value of a voxel at a time that was not originally measured.[5]

2.1.2 Realignment

During statistical analysis, it is assumed that single voxel represents an unchanging part of the brain. Some sessions can last for 1-2 hours and subjects, even though instructed not to, may get uncomfortable and change their position slightly. Even if the subject did not move consciously, heartbeat and respiration can cause small

periodical changes in position as well. In the process of realignment, only unimodal images are realigned, such as functional images to other functional images or structural to other structural images.[5]

For the correction of motion rigid body transformations with 6 parameters. Three parameters for translation in x, y, z axes and the resting three parameters for rotation around x, y, z axes. Sometimes called Euclidian transformations because they preserve the Euclidian distance between points. So it is assumed that the two images to be realigned can be superimposed on top of one another. Three-dimensional translation and three-dimensional rotation yield a total of 6 parameters to be calculated.[5]

An important step in the computation is the way we determine the amount of dissimilarity or similarity when images are realigned. This similarity measure is called the cost function. In simplest cases, it is the absolute difference of intensities or squared difference of intensities also called the quadratic cost function. In the process of realignment, the goal is to find rigid body transformation with the smallest value of the cost function.[5]

2.1.3 Coregistration

This step is often used to align or map two images obtained through different modalities, such as structural images to functional images. Because TR is generally set low images obtained through fMRI are usually blurry and have low spatial resolution making the identification of anatomical structures difficult. It is therefore advantageous to first obtain T_1 weighted structural images and then in preprocessing coregister them to functional images.[16]

The rigid body transformation may be used again with an appropriate cost function. Because images obtained from different modalities can have different intensity dependencies dynamic a cost function called mutual information may be used. Mutual information cost function minimizes mutual entropy and expresses the degree of uncertainty with to which we can predict changes in the first image if changes in the second image are known. Yet another reason for coregistration is a geometrical distortion of images. Different pulse sequences may introduce stretching along with some or all axes. Here, it is necessary to use affine transformations and add three more parameters reflecting the amount of scaling along x, y and z axes.[16]

2.1.4 Normalization

So far intra-subject images were realigned and coregistered. Most studies, however, need to analyze data from different subjects. The human brain is remarkably anatomically variable. Subjects brain in a single study can differ in size by 30%. Brains can also differ in shape, some may be longer and thinner while others are

shorter. This inter-subject variation makes it difficult to compare data acquired from different individuals. By normalization, images are mapped to a standardized space most frequently to MNI space. The brain template of MNI was created by using a large amount of MRI scans and was defined by the Montreal Neurological Institute.[5]

Normalization is a special form of coregistration. It uses affine transformations to stretch, squeeze and warp each brain to match the MNI template as best as it can. Therefore it is non-Euclidian transform since it does not preserve the Euclidean distance between two points.[5]

2.1.5 Spatial smoothing

Spatial smoothing employs spatial filters to remove noise. The most common filter is a 3D Gaussian spatial filter. Which convolutes images with 3D Gaussian kernel which acts as a low-pass filter. The resulting image becomes more blurry but the signal-to-noise ratio (SNR) grows larger. Ultimately higher SNR means less false positives, which improves the accuracy of the analysis.[5]

2.1.6 Temporal filtering

Finally, data may be converted by Fourier transform to the frequency domain and filters can extract most relevant data or suppress undesirable bandwidths. One such filter suppresses breathing artifacts. Respirometers may determine breathing frequency and design either a low pass or band-stop filters.[5]

Prerequisites for temporal filtering are that the TR must be such so that there would be no aliasing of the desired frequencies and experiment design needs to be taken into consideration. Particularly the speed of stimuli presentation. Fast event-related experiments may present stimuli every few seconds thus blending in with respiration. If Normal respiration rate is around 0.25 Hz so sampling rate of less than 0.5 Hz or TR smaller than 2 s would have to be used.[5]

2.1.7 Outlier removal

Non-linear artifacts are characterized by heavy-tail distributions (distributions do not exponentially decay to zero). They tend to express themselves as abrupt intense changes in signal intensity, known as spikes or outliers. Causes of non-linear artifacts are usually inhomogeneities in the magnetic field, respiration, movement, and heartbeat artifacts.[17]

A simple approach for outliers removal would be to specify a certain threshold to define outliers. E.g. FSL, a library of analysis tools for fMRI, uses the 75th percentile

+ 1.5 times the interquartile range as a threshold [18]. Another fMRI analyzing tool uses 4 to 8 standard deviation threshold for the construction of an outlier map[19].

After outliers have been identified data containing them can be either rejected or replaced with new values. Rejecting them in fMRI could cause abrupt changes in the feedback signal. The better option would be to replace them to e.g. fit a smooth curve.

2.2 Statistical analysis

After preprocessing, data can finally be analyzed. Statistical analysis methods are used to asses whether the differences between data sets are significant enough or are a mere result of random variables. One of the most widely used methods is a general linear model, or GLM for short.

2.2.1 T-statistics

The simplest approach is to subtract the mean of the signal acquired in resting periods from the mean of the signal acquired in times of neuronal activity. This approach sets high requirements on SNR and is most noise sensitive. T-statistics go beyond this simple approach by also taking the amount of variability in data into consideration. The results of the t-statistics measure if the difference between two data sets is statistically significant.[15]

2.2.2 General linear model

GLM entails a generalized multiple linear regression analysis. It assumes the data is composed of a linear combination of model factors and uncorrelated noise, or error term. GLM aims to find parameters β for design matrix G, such that they best explain data Y and thus minimize the residual term ϵ [5]. The equation may be expressed as:

$$Y = \beta * G + \epsilon \tag{2.1}$$

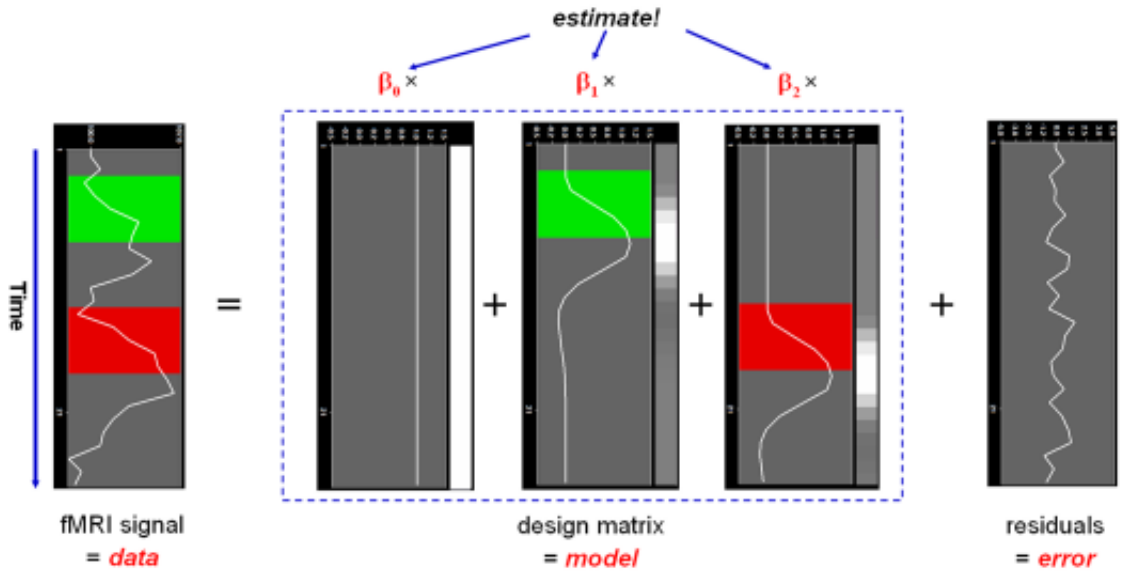

Fig. 2.1: General linear model for a single voxel.[20]

The equation is graphically represented in figure 2.1. In equation 2.1 the term Y represents a 2D data matrix that represents data from fMRI. Each voxel occupies one column and rows represent points in time. Matrix G, or design matrix... Vector β, or the parameter matrix sets how much each regressor contributes to each voxel. The last term is the residual or error term. Its dimensions are identical to a data matrix Y thus each entry represents the addition of error to one voxel at a certain time point. The contribution of error term decreases if the matrix-vector multiplication β*G explains data matrix Y well.[20]

As can be seen in figure 2.2 addition of information about experimental design to a design matrix should minimize the residual term ϵ.[20]

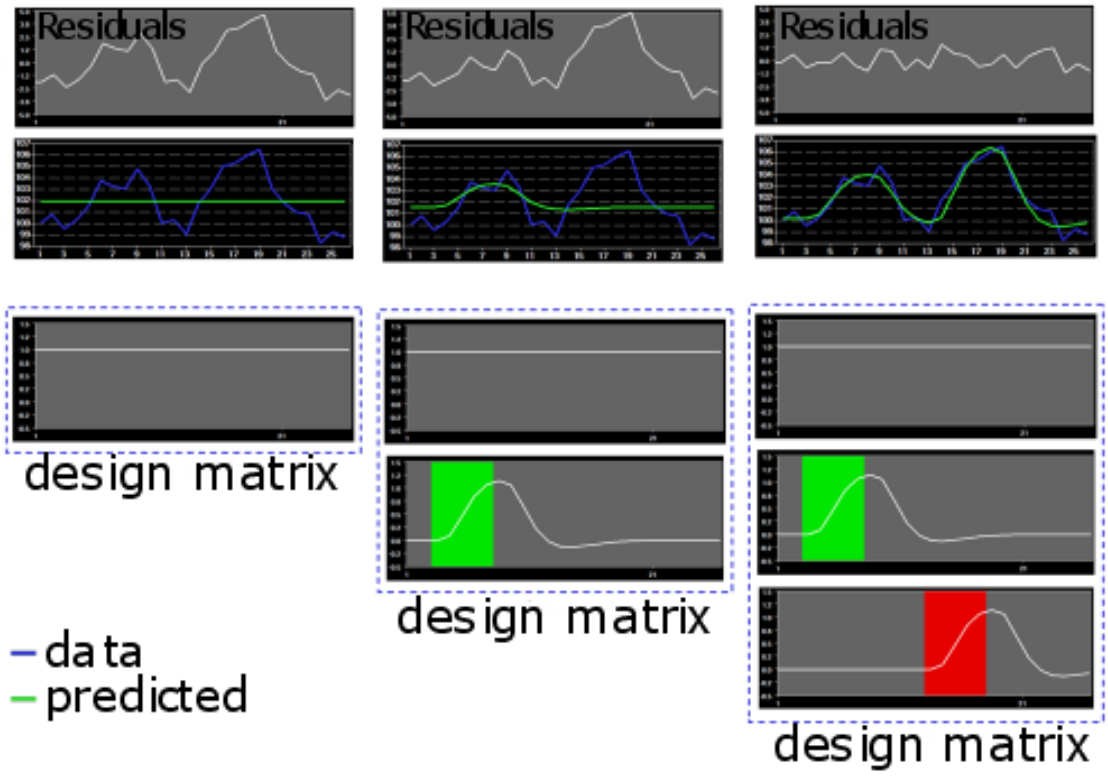

Fig. 2.2: Improvement of prediction with addition of information about experimental design. Adapted from [20].

If the residual term remains considerable the hypothesis of an experiment may be wrong. Significance tests are used to measure whether the model significantly well describes the variance in a time course of a voxel or in other words measures how well the model fits the data. Example of significance tests are t-statistics, F-statistics, z-statistics.

3 Real-time fMRI neurofeedback

Neurofeedback is a real-time process in which information about an individual's brain activity is communicated back to the individual. Therefore creating a feedback loop that allows for increased self-awareness and self-regulation. Neurofeedback training aims to volitionally up-regulate or down-regulate brain activity and has been successfully applied to treat conditions such as ADHD.[21]

Real-time fMRI (rtfMRI) allows for analysis within a single repetition time. The schematic representation of rtfMRI neurofeedback may be seen in figure 3.1 below. Compared to other modalities such as EEG neurofeedback, which offers superb temporal resolution but poor spatial resolution, fMRI offers better spatial resolution at the cost of decreased temporal resolution. Another difference is the inherent feedback delay of fMRI which in some cases could make the neurofeedback training hard or impossible. That is also why the subject is usually informed about the delay, which is dependent upon image acquisition time, processing time and the inherent BOLD response, which takes about 6 seconds to peak.[22]

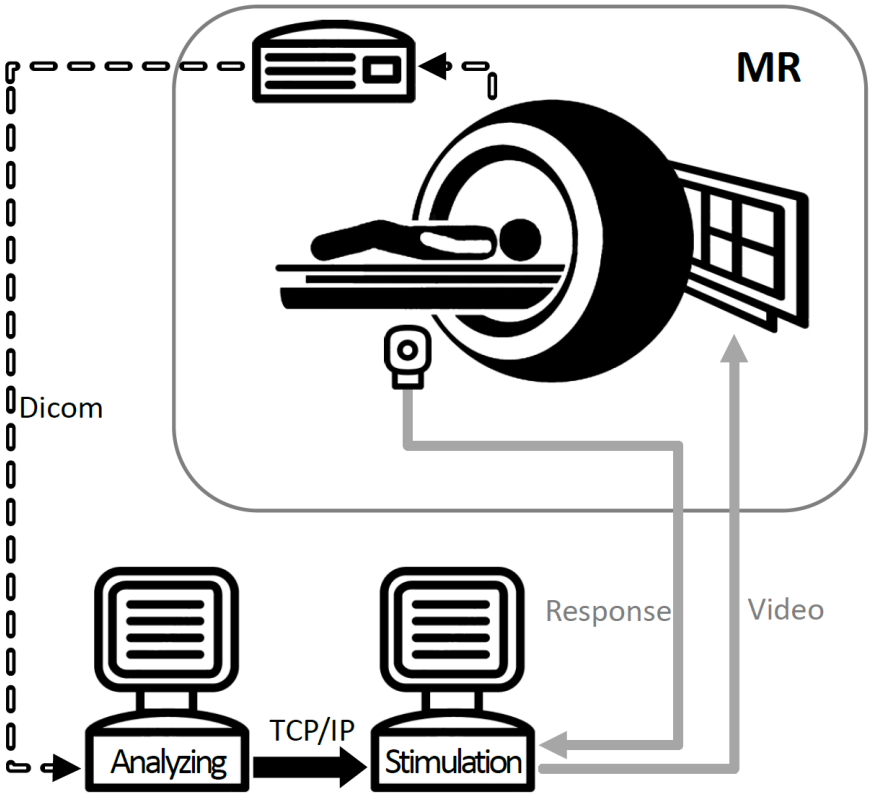

Fig. 3.1: Flow diagram of rtfMRI neurofeedback

Data acquired by MRI are converted to DICOM format before being sent further. DICOM is the international standard for transmitting and storing medical imaging data.[23] The DICOM data are received and then user defined preprocessing and analysis ensues. The feedback loop closes when the results are communicated to the subject through neurofeedback software. Neurofeedback software may display to the subject information about the task at hand, about the levels of local neuron activation at sites of interest and also receive information from the subject. The [21] Degree of cerebral activation at ROIs is usually communicated visually. The form varies from emoticons to fluctuating level of a thermometer or a simple graph.

3.1 Current neurofeedback solution

Over the last decade interest in rtfMRI neurofeedback has grown. There still exists little consensus among experts about the quality of different signal processing strategies. Nevertheless, the pressure on the quality of signal processing strategies might be even higher as the big part of the processing needs to happen in real-time. Neurofeedback also puts certain demands on the preprocessing part. "Particularly, artifacts and sudden signal changes need to be considered since they may lead to learning of nuisance signals."[17]

A diagram showing data processing steps for fMRI neurofeedback currently implemented at CF MAFIL CEITEC can be seen in the figure above. Each step will now be briefly described.

First, a structural image of the whole brain is acquired. This serves not only as an anatomic template to which all functional images are aligned to but also for localizing ROIs. ROIs are usually defined as a binary mask in MNI space. The structural image is normalized to MNI space which yields a transformation matrix. Masks can then be transformed by the inverse transform matrix to the space of the individual. Next, the first functional image can be co-registered to the structural scan.

In neurofeedback studies, shorter TRs are used so the possible slice-timing errors decrease in magnitude as well. That is why many neurofeedback studies omit slice-timing corrections with negligible consequences.

Next steps are performed in real-time while the next functional image is being acquired. Following functional scans do not need to be co-registered to the structural scan but instead, they can be realigned to the first functional image. Realignment is generally faster as two images from the same modality are aligned compared to co-registration where two images obtained through different modalities are aligned.

After realignment individual's ROIs masks are multiplied with functional images extracting only the signal from these areas. The resulting spatial signal is averaged

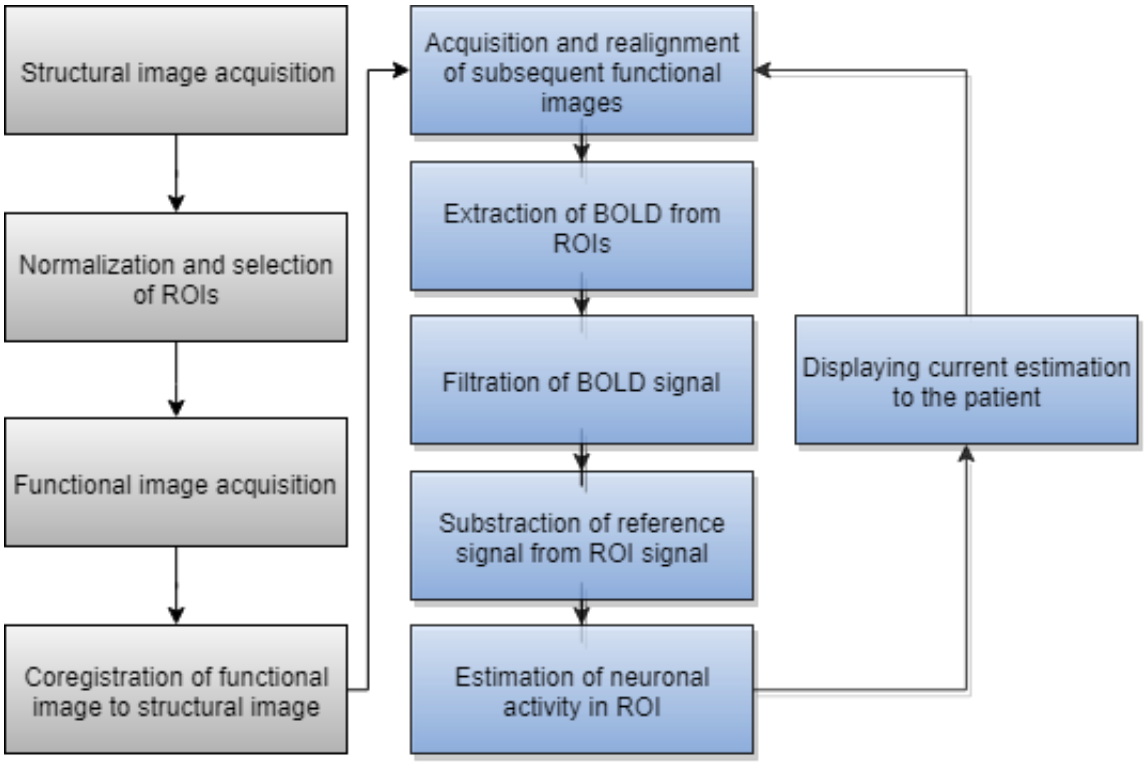

Fig. 3.2: Data processing for fMRI neurofeedback. Blocks with gray background are done before the neurofeedback part commences. Blocks with blue background are done in real-time.

and a single metric of activity from ROI is obtained.

As the data can be noisy and contain artifacts use of temporal filtering can improve the overall SNR. The nonlinear quadratic estimation algorithm is also known as extended Kalman filtering steps in to fill this role. Kalman filter uses a series of measurements that contain inaccuracies and statistical noise to produce an estimate that is more accurate, than those based on signal measurements alone. As it is also a prediction algorithm it is ideal to use in real-time applications where other filter types could introduce greater signal delay. More about Kalman filter may be found in chapter 5.

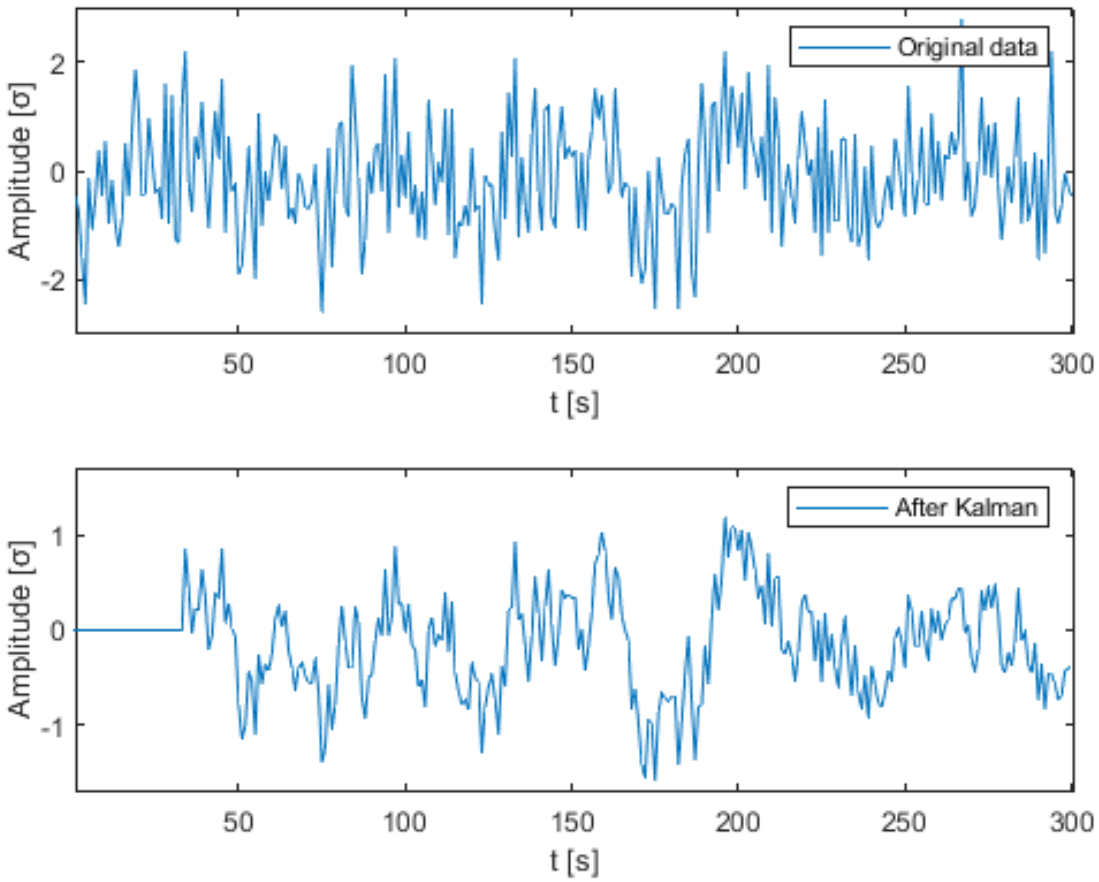

Fig. 3.3: Unfiltered signal and signal after Kalman filtering. Note the difference in amplitude.

The output from Kalman filter may be offset by a constant or by a linear trend further into the session. A trend is a systematic decrease or increase in the independent variable. It is characterized by its slow drift like properties. In fMRI low-frequency drifts of 0 - 0.015 Hz has been frequently observed. The likely cause of origin is scanner instabilities[24]. Drifts may seriously influence results of statistical analyses other analyses may also be adversely affected e.g event-related averaging. Also by removing the systematic low-frequency trends from data, it is easier to focus on the fluctuations in the data which actually carry useful information. Therefore any linear trends are removed after Kalman filtering.

Signals from target ROIs are then scaled by their respective averages and subtracted from control ROI signal. Control ROI is a region which should not react to the stimulus being presented to the subject and should remain relatively stable throughout the experiment. Also if the subject moves or if there is measured activity which concerns both the control region and target region the effect of this activity

can be mitigated as it is clear it is not relevant feedback signal.

Last, estimated activity is communicated to the individual usually as either visual or sound cue.

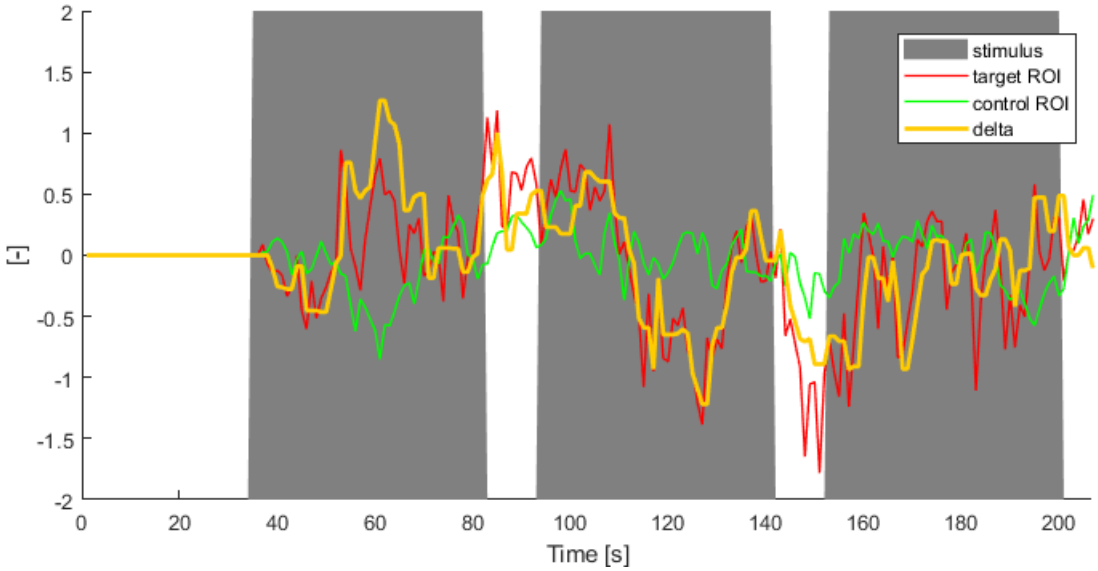

Fig. 3.4: Output from current neurofeedback solution. Final neurofeedback signal delta can be seen in orange.

In the chart 3.4 a typical neurofeedback session in progress can be seen. Grey chart areas signalize that the patient is currently being stimulated. These are intertwined with short periods of rest. A typical session consists of 4-5 neurofeedback blocks each lasting about 8-10 minutes. It can be seen that the first 35 samples of the block, or 35-70 seconds when the usual TR of 1 - 2 seconds is chosen, the output is zero. And it is done so that the extended Kalman filter would have enough time to learn the signal and would output correct values. As the subject can undergo about 3-5 of these blocks in one session depending on TR it amounts roughly to 1.5 - 6 additional minutes spent in MR. During long sessions subjects can become tired and or start moving slightly, not to mention financial costs. The aim of this thesis is to reduce the amount of idle time before every block.

4 Filtering of temporal data

Filters have a wide variety of uses. Most commonly they are used to either increase or decrease amplitude at certain frequencies. There are two major types of digital filters. Finite impulse response filters (FIR) and infinite impulse response filters (IIR). According to the frequency, they will leave unattenuated, they can be further sub-divided into low-pass, high-pass, and band-pass filters.

4.1 Finite impulse response filters

FIR internal structure can be seen in figure 4.1. As its name implies when it's applied on an impulse function its output will settle on zero in finite time.[25]

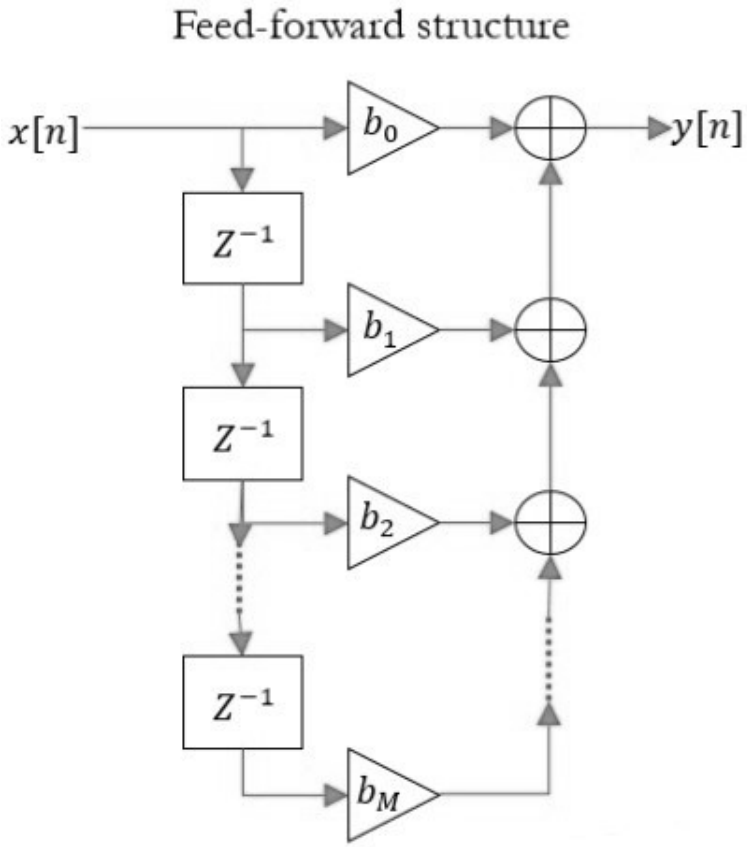

Fig. 4.1: Inner workings of FIR digital filter.[26]

As may be seen in FIR flow graph above, filters output is composed of up to x(N-M) previous inputs. Every input is multiplied by a constant b and then summed up with the rest. As the output is produced only from previous inputs, it takes M + 1 zero input samples before the filter's output is zero.[25]

A major advantage of FIR filters is that they may provide a linear phase response. Since phase is a linear function of frequency then delay is same across the frequency spectrum. Filters with non-linear phase cause phase distortions. To design a linear phase FIR filter, b coefficients must be symmetrical in the time domain.[25]

As there is no internal feedback loop, a bounded input will always produce a bounded output. That is why FIR filters are inherently stable and there is no need to check their stability.[25]

On the other hand, FIR filters need to be of a higher order than IIR filters to achieve comparable frequency response. The order of a filter is determined by the number of delay lines. Higher order filters are more computationally expensive.[26]

4.1.1 Moving average filter

Moving average filter is a simple low-pass FIR filter. It takes L input points computes their average and outputs that as a result. It is a commonly used filtering method for data with a noisy high-frequency component. If the length of the filter is increased the output gets progressively smoother. The equation for a simple moving average filter where x denotes input can be seen below.

$$y[n] = \frac{1}{L} \sum_{k=0}^{\infty} x[n-k] \tag{4.1}$$

4.2 Infinite impulse response filters

IIR, as its name suggest, is a filter which output never becomes exactly zero when an impulse function is presented to the input.[25]

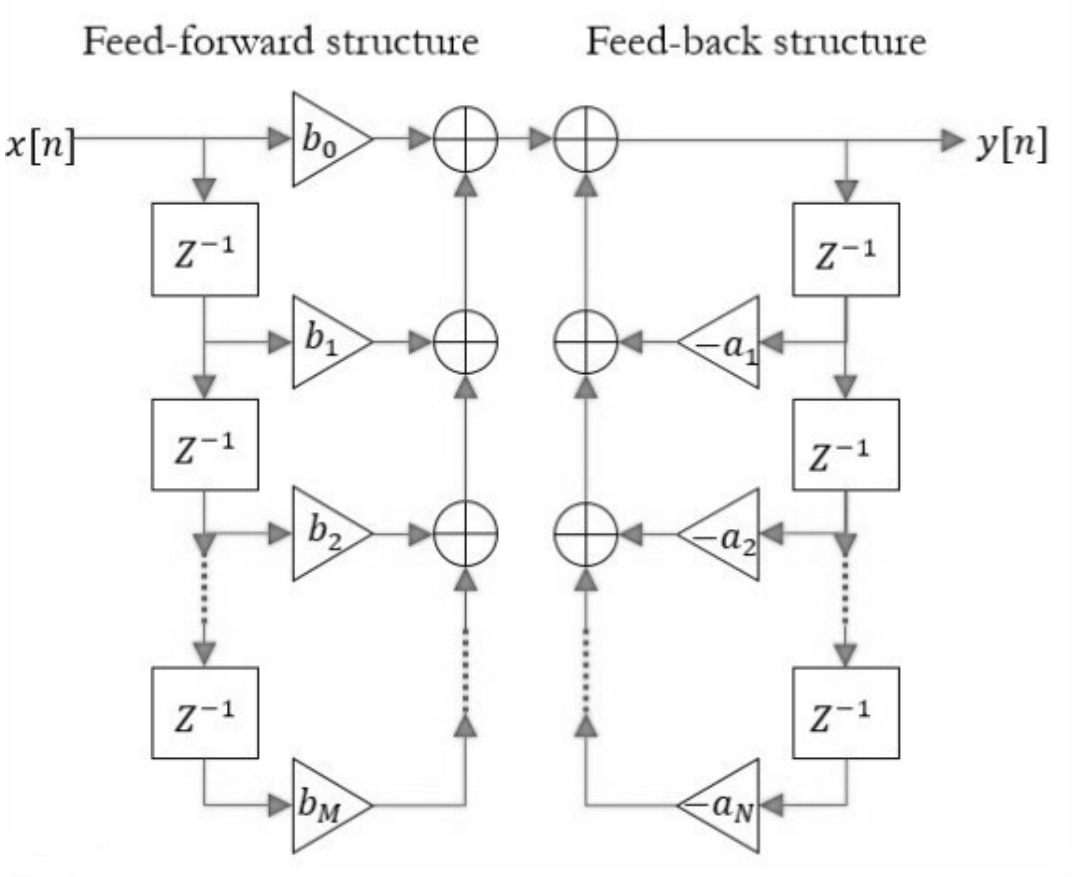

Fig. 4.2: Inner workings of IIR digital filter.[26]

As can be seen in figure 4.2, IIR shares a feed-forward structure part with FIR. The second part constitutes a feedback loop. Current output depends not only on current and past inputs but also on past outputs.[25]

As there is a feedback loop, past numeric errors may be remembered by the filter and propagate to the next cycle. FIR filters, on the other hand, do not remember previous outputs and its past numeric errors do not propagate further. In IIR filters bounded inputs may produce unbounded outputs if not designed properly. Therefore IIR filters should be tested for stability.[25]

IIR filters have non-linear phase responses and cause phase distortions. Each frequency component experiences different phase delay.[25]

On the other hand, IIR filters achieve desired filter characteristics with lower order than their finite counterparts. They are less computationally heavy and introduce less average signal delay.[25]

4.3 Kalman Filter

Linear Kalman filter is an adaptive recursive estimation algorithm. It takes multiple measurements to "learn" the characteristics of the signal and is then able to provide an estimate of the output of the system to a certain input that tends to be more accurate than a single measurement of that same output. That is achieved by joining probability distributions of the measurement and the output estimate. The equations necessary for the implementation of the Kalman filter may be divided into two phases: the prediction phase and update phase. The Prediction phase which tries to predict the state of a system at a current time point, produced from previous time step estimates. Update phase takes the current measurement of a system into consideration and combines them with results from the prediction phase. Equations for both phases may be seen below. Variables with a hat over them indicate the variable is an estimate.[17], [27]

Prediction phase:

$$\hat{x}_{k|k-1} = F_k * \hat{x}_{k-1|k-1} + B_k * u_k \tag{4.2}$$

Where $\hat{x}_{k|k-1}$ is the "new" a priori estimate. F_k is the state-transition model. $\hat{x}_{k-1|k-1}$ is the previous estimate as calculated in the previous step by equation 4.5. B_k is the control matrix applied to control vector u_k. Control vector may contain information about actions taken by the system in the previous step. Such as if autonomous car decided it was going to fast in the previous step and started to break. The equation could be reformulated like this, the new a priory estimate is obtained from the previous estimate combined with known prediction model plus the influence of external variables.[27]

$$P_{k|k-1} = F_k * \hat{P}_{k-1|k-1} * F_k^T + Q_k \tag{4.3}$$

Where $P_{k|k-1}$ is a state covariance matrix (uncertainty in the $\hat{x}_{k|k-1}$ estimated state). $\hat{P}_{k-1|k-1}$ is the covariance matrix of the previous estimate $\hat{x}_{k|k-1}$. Q_k is the process noise covariance matrix (uncertainty from the environment, not modeled by theoretical equations represented by F_k). In other words, the uncertainty of the new a priory estimate equals the previous uncertainty estimate plus uncertainty from the environment.[28]

Update phase:

$$K_k = \frac{P_{k|k-1} * H_k^T}{P_{k|k-1} * H_k^T * H_k + R_k} \tag{4.4}$$

Where K_k is the Kalman gain matrix. H_k is the observation model, mapping true state space into observed space - sensors might have different scale and units than

the state space used in the rest of the calculation. Finally, R_k is the measurement covariance matrix (uncertainty measure of the measurement).[28]

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - H_k * \hat{x}_{k|k-1}) \tag{4.5}$$

Where y_k is the measurement matrix - measurements obtained currently from the sensors.

$$P_{k|k} = \hat{P}_{k|k-1} - K_k * H_k * \hat{P}_{k|k-1} \tag{4.6}$$

In the update phase measurements from the sensors, y_k are combined with calculated estimates that are based on previous knowledge about the system. Kalman gain modifies equation 4.5 such that if the predicted estimates have much less uncertainty than the uncertainty of sensor measurements than the Kalman gain will be close to zero so the new estimate will come mostly from the predicted estimates and vice versa.[29]

At the beginning of the algorithm, the posterior estimate comes from sensor measurements. As Kalman filter learns the estimates get progressively more accurate, and the filter relies less on sensor measurements. The learning period is therefore needed and in each neurofeedback session first, few fMRI measurements are used to learn the filter. The results at this stage are basically unfiltered and are unfit for neurofeedback. The number of neurofeedback sessions conducted in one sitting varies but it is not unusual to have 4 and more sessions. This is not time efficient not only it is more financially costly, but it also increases the time the individual cannot move and so his discomfort. This may lead to more motion artifacts.[27]

Linear Kalman filter was extended to remove outliers. Such modification breaks the law of linearity and nonlinear Kalman filter is also known as extended Kalman filter. This was done by setting a threshold for the amount of change that may occur between the two following Kalman steps. The threshold was calculated as 0.9 times the standard deviation of the signal to date. When the spike is detected previous output is held as well as previous covariance matrices. More about the mathematical implementation of this filter may be found in [17].

As with any spike detection algorithm either a threshold value needs to be set beforehand or is calculated from the data set. An issue arises if the spike detection needs to happen online, there is a low sampling rate and reliable results need to be obtained as soon as possible.

4.4 Adaptive filters

Adaptive filters change filter coefficients to better adapt to the characteristics of a changing signal. Coefficients are adapted by an optimization algorithm which usually includes minimization of an error criterion.[25]

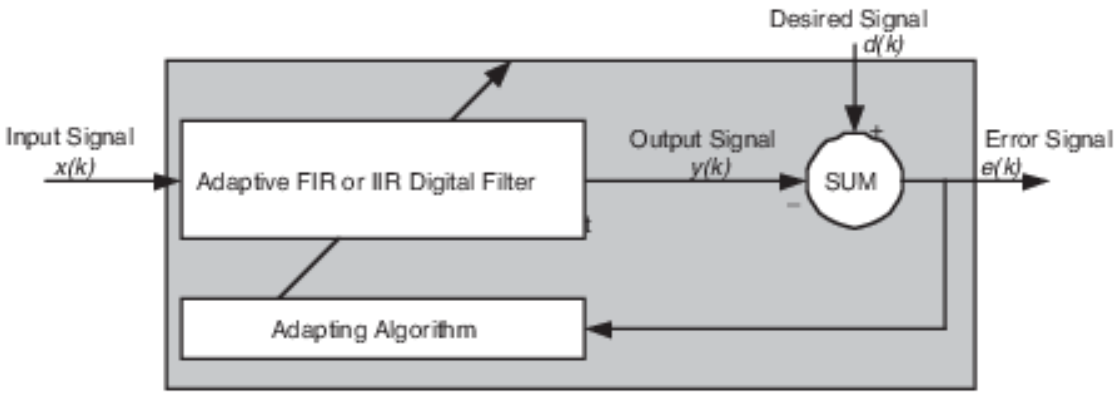

Fig. 4.3: Block diagram of a generic adaptive filter.[29]

As may be seen in figure 4.3, FIR and IIR filters and many others may be modified to be implemented inside an adaptive filter. Adapting algorithms use error criterions such as least means square or root means square. Desired signals are compared to the output signal and their difference is fed to an adapting algorithm. Desired signals in the scope of fMRI are usually the hemodynamic function. The adaptive filter then tries to extract data with a similar structure as HRF.[25][30]

5 Proposed approach for online data filtering

Inherent fMRI NFB delay can make neurofeedback training hard or impossible. It is therefore important not to introduce further delay by introducing filters not suitable for real-time applications. Filtering is necessary since the signal contains high-frequency noise and spikes if both have sufficient amplitude they are mistaken by the subject as useful NFB signal. This can lead to learning of nuisance signals and or can cause confusion. Filtering is made more difficult because the sampling rate is quite low (TR = 1 - 2 seconds) while at the same time it would be optimal to have an accurate output as soon as possible.

Current data filtering solution utilizes an extended Kalman filter (EKF), which was chosen as a compromise between its smoothing properties and its extremely low lag, which is why it is used mainly in real-time applications. Nevertheless, the Kalman filter takes a certain time at the beginning of each neurofeedback block to "learn" the characteristics of the signal. Learning length is set to 35 samples which can translate to a transient period lasting 35 - 70 seconds each neurofeedback block depending on the repetition time. Sessions most often consist of 3-5 neurofeedback blocks thus the learning period significantly lengthens neurofeedback sessions.

Data was acquired from the Multimodal and Functional Imaging Laboratory, MAFIL, of the Central European Institute of Technology (CEITEC). The data set comprises of 50 subjects all of which had 3 neurofeedback sessions. Relevant data had to be extracted and the whole data sets contain 150 individual neurofeedback blocks all of which had the same experimental design. Target ROI was amygdala, which is a usual object of study due to its role in emotional responses and mental states. All data had been linearly detrended and subsequently z-score standardized.

5.1 Analysis of extended Kalman filter

Functional MRI neurofeedback is basically a one sensor measurement. As was explained in section 4.3 Kalman filter is very good at combining multiple sensors measurements, where each has a given measurement uncertainty. In this particular case, only one sensor measurement is obtained and Kalman filter works on narrowing down the real underlying value. Even state transition model matrix is set to $F_k = 1$ and so is observation model matrix $H_k = 1$, therefore all the matrices in equations become one dimensional, scalars. The model can be expanded in case of signals from a various number of target ROIs that need to be merged, nevertheless in this application Kalman filter effectively acts a linear low-pass adaptive filter.

Since the covariance would have two identical variables (covariance with itself), simply variances can be computed. The update rate sets the cut-off frequency of low-pass Kalman filter and can be computed as

$$\lambda = \frac{R_k}{Q_k} \tag{5.1}$$

As these variables directly impact Kalman gain, update rate sets a limit on how quickly can the state variables change. Low-pass filtering behavior can be seen in figure 5.2.

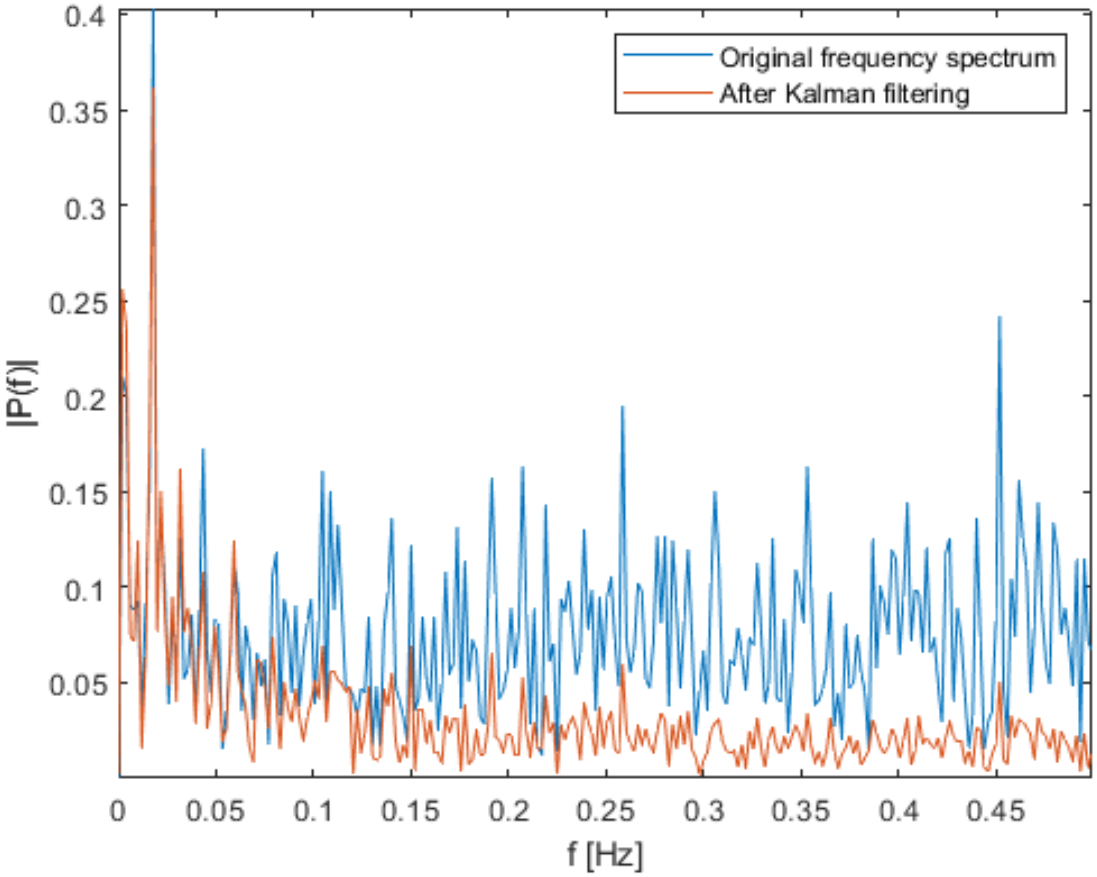

Fig. 5.1: Single-sided amplitude spectrum of a typical fMRI neurofeedack signal before and after Kalman filtering.

Here, the necessity of a 35 sample learning period will be discussed. There are five inputs to the Kalman filter. Previous output, its uncertainty, current measurement, its uncertainty, uncertainty from the environment and threshold for spike detection algorithm. Only current measurement uncertainty, uncertainty from the environment and threshold values can be manipulated directly and all of these inputs have one thing in common, they all stem from standard deviation calculation. That is why to

test real-time Kalman properties there has been constructed an offline Kalman filter where the only difference is that it sees all data points and thus is able to calculate standard deviation more reliably. Both the extended version of Kalman and version without spike detection algorithm has been tested with their respective counterparts. Results can be seen in figure 5.2.

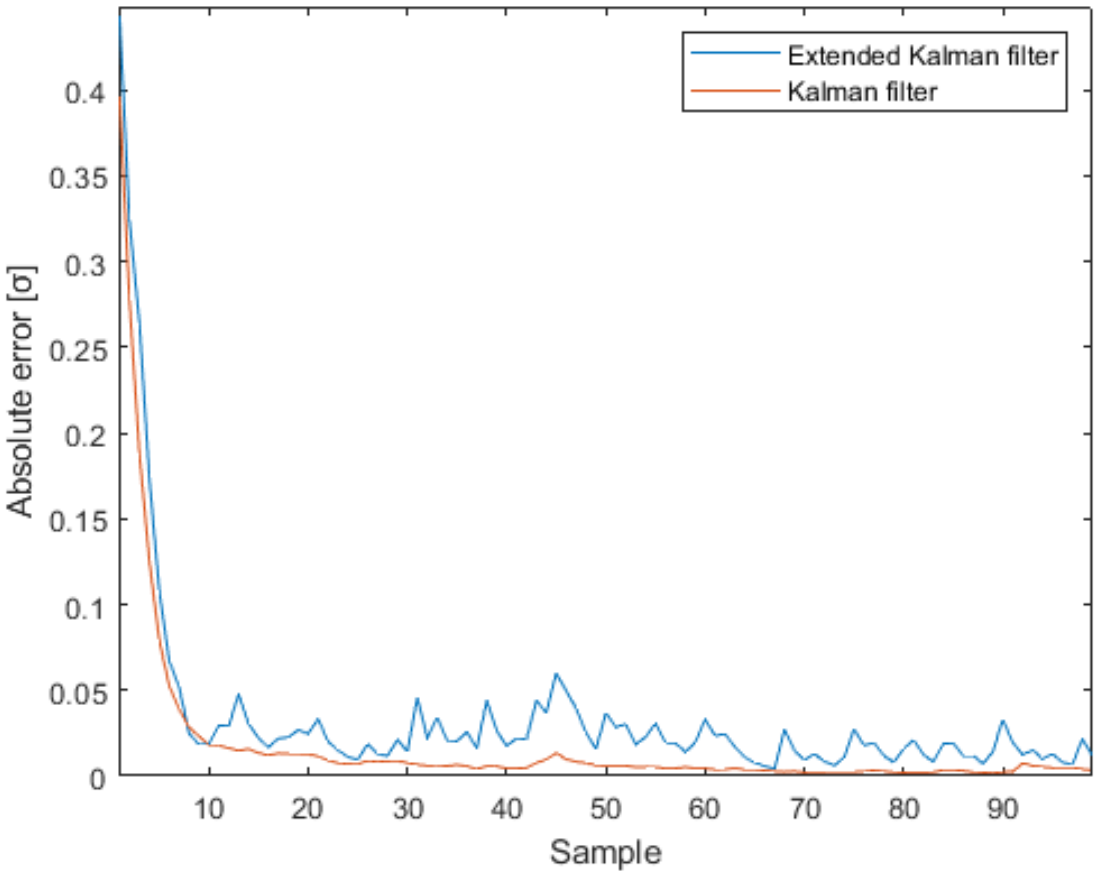

Fig. 5.2: Deviation of real-time Kalman filters compared to its offline counterparts.

As was expected most deviation can be seen in the first few samples. The likely cause is that there are not enough samples to calculate the standard deviation (std) i.e., there are wide confidence intervals. Std estimates vary greatly until about the 10 sample. This is the case for both variations of Kalman. After the first few samples Kalman filter continues to exponentially decade to zero. Nevertheless, EKF has ups and downs, where the downs come close to its de-extended version. This is likely because the deviation from spike detection is superimposed with the Kalman filter deviation.

To make sure the behavior is consistent with respect to Kalman we certainly know is in its taught phase, unlearned Kalman filters were run in 200-500 sample range and compared with Kalmans that were running from the beginning. Same

exponential decay curves were observed and had also reached their plateau phases around sample 10.

To quantify the absolute error of EKF, three intervals were chosen for which the average error was calculated. Average absolute errors were

$$AAE_{10-35} = 0.0232 \qquad AAE_{35-60} = 0.0283 \qquad AAE_{60-85} = 0.0153$$

Higher error in interval 35 - 60 is caused by an outlier near sample 45 and real value most probably lies somewhere between the first and third interval. The average absolute errors of all the intervals are relatively low with little change in absolute average errors between these intervals. Therefore EKF can be considered taught after the tenth sample.

Further decrease in learning time would be possible if, for example, we have used a certain amount of data from previous sessions. However, this could adversely affect the solution's robustness. Using data from previous neurofeedback blocks of the same individual within one session would most likely not strongly affect solution robustness. First neurofeedback block still would have a learning period of at least 10 samples while others would have less. This would be inefficient from the viewpoint of further offline analysis because of their time shift which would have to be accounted for or corrected somehow during that stage.

EKF should still be ideally used and as there is no way of shortening its learning period without affecting its robustness, the next section is concerned with introducing a temporary filter. Which could filter signal before EKF learning period is over and then switch to EKF output.

5.2 Temporary filter

Further decrease in the amount of block delay is difficult since it is connected to lag and smoothing. When one is modified all others are affected as well. There is a well-known trade-off between the degree of smoothing and lag. The more we would like to smooth the signal the more lag is introduced. Offline applications can sometimes work around this issue e.g. symmetric FIR filters can be applied to signal twice in each direction to remove the lag. Nevertheless, this is not applicable to real-time solutions. And when we would decrease the block delay it would adversely affect smoothing.

Adaptive algorithms or estimation algorithms (KF), all need a period of learning or adaptation. FIR or IIR filters of shorter length would be therefore more useful. IIR has a non-linear phase response and could introduce unwanted phase distortions, therefore, low-pass FIR filter would be a better option.

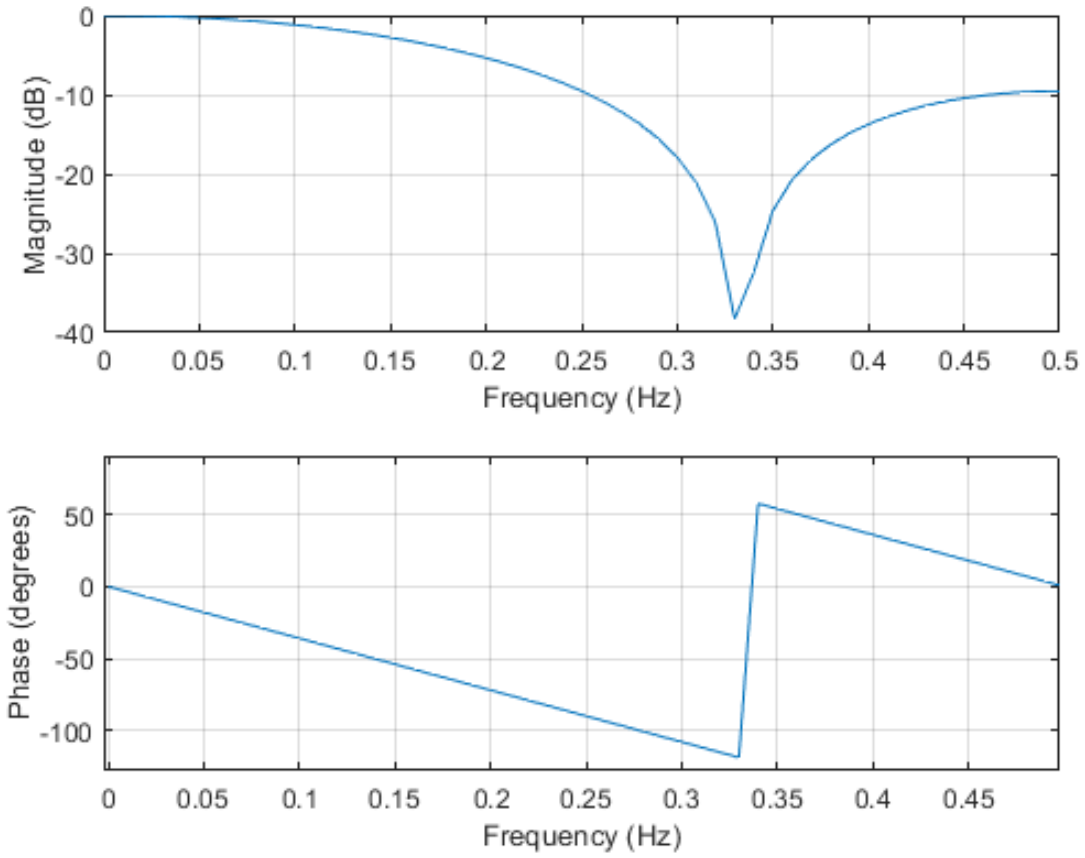

Fig. 5.3: Magnitude and phase response of length 3 moving average filter.

A causal moving average filter is a low-pass filter used mainly for its smoothing properties. The optimal length of a window had been found to be 3, as a tradeoff between smoothing and lag. Frequency response and phase response of such a filter can be seen in graph 5.3. The attenuation at 0.1 Hz is around -1 dB which is within acceptable limits.

5.3 Spike detection

Spikes are an integral part of fMRI and their detection and removal is an important part of the overall SNR improvement process. Examples of spikes in the unfiltered signal may be seen in figure 5.4.

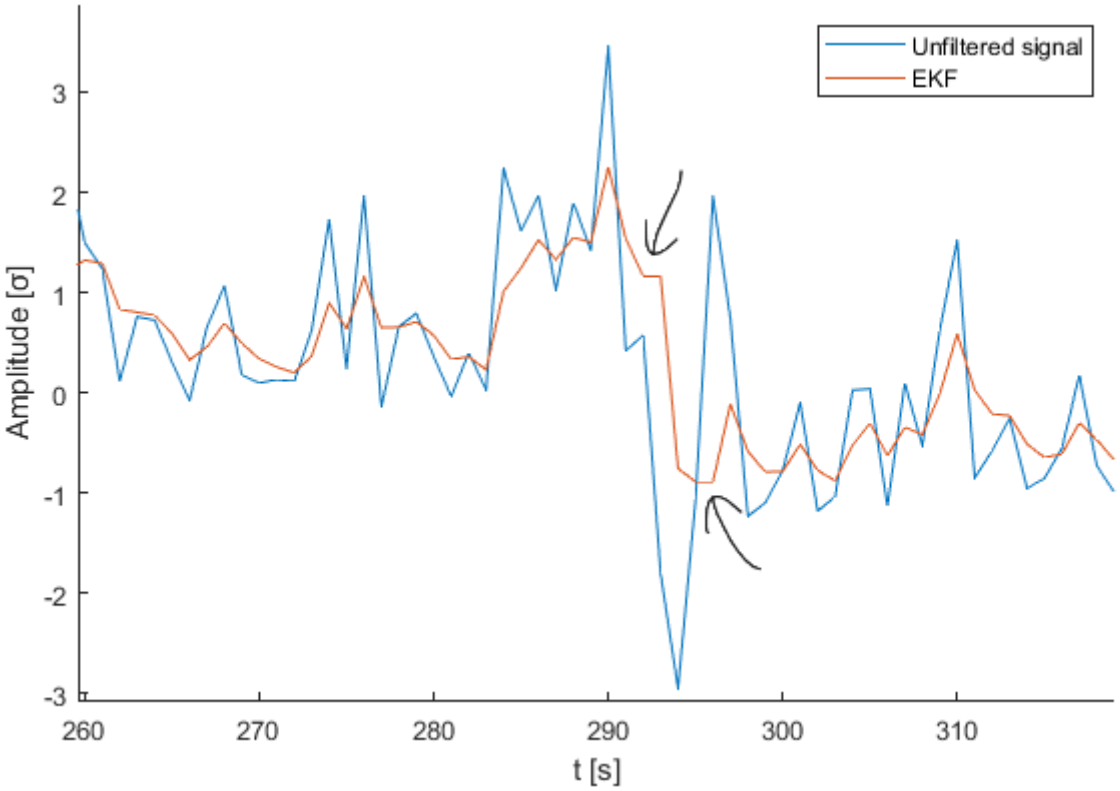

Fig. 5.4: Spike detection algorithm. Arrows indicate where spikes were detected and subsequently removed.

The threshold for spike detection is set at 0.9 of standard deviation, calculated from the unfiltered signal. If the Kalman filter would be updated by a greater amount than that, a spike is detected and previous EKF output value is held. Spike length is set to one in both directions, thus spikes lasting longer than one sample will have only the first data point removed. Counters for positive and negative spike lengths are separate, so in case there are two subsequent spikes in a negative and positive direction both would be removed. Due to its low sampling frequency spikes, longer than one samples are not common.

6 Results

The proposed solution utilizes two different filters, temporary MA filter, and EKF. Temporary MA filter can be switched to anywhere after EKF transient period is over and in the previous chapter had been found to be after sample 10. The outcome of the proposed approach may be seen in figure 6.1

To ensure there is no significant difference when the switch happens Wilcoxon rank-sum test was used. A potential significant difference could potentially confuse the volunteer or it could lead him to learn nuisance signals. The output from both filters were taken at the switch point from the whole data set. Wilcoxon rank-sum test p-value for switch happening at sample 11 was 0.85 well above the significance level α of 0.05 required to reject the null hypothesis.

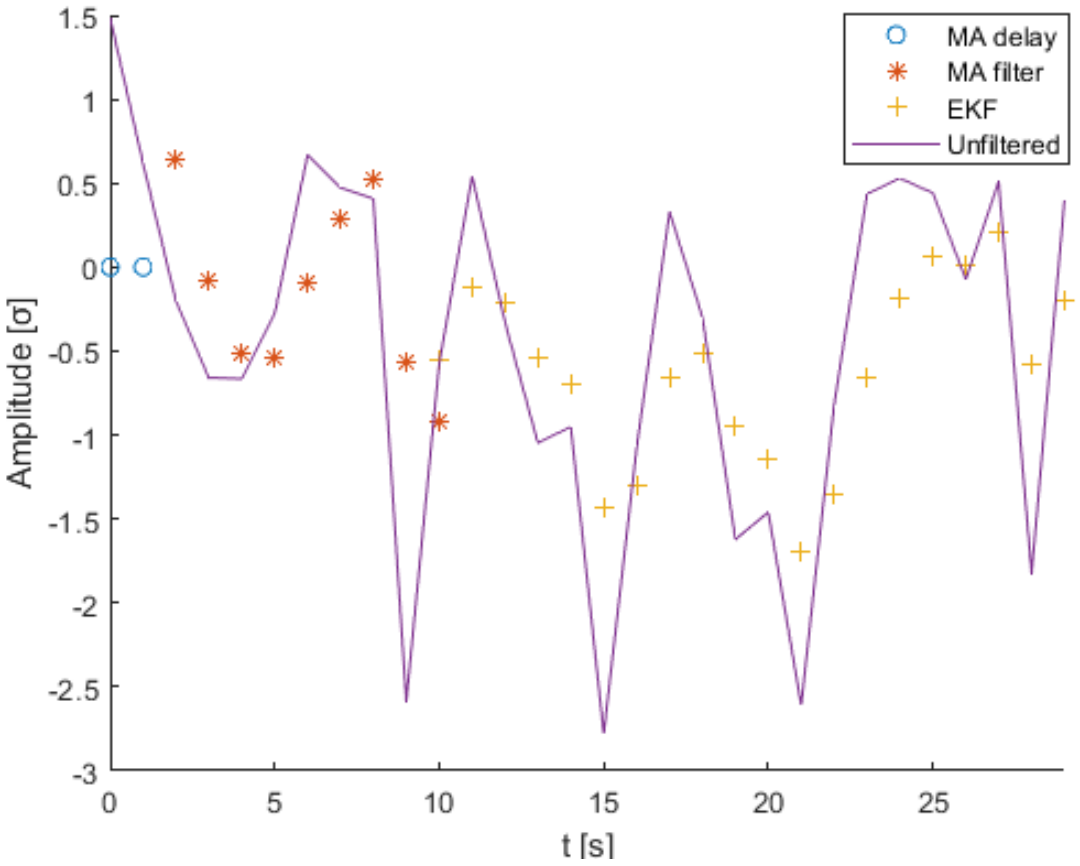

Fig. 6.1: Outcome of the proposed approach for online data filtering.

The course of unfiltered and filtered signals may be seen in 6.2. Although not an ideal solution in terms of filter lag the output of moving average filter closely follows that of an extended Kalman filter. Probably the biggest differences are seen when a spike is detected and removed by EKF for example at sample 35.

The figure also nicely illustrates the real-time properties of both filters. Kalman reacts immediately to a trend reversal and usually peaks at the same time as the unfiltered signal, while the MA can sometimes be seen lagging behind. MA delay can be most easily seen during short spike values such as those around sample 16 and 31. The introduced lag is not large, maximum lag is one sample as can be expected from a length 3 FIR filter.

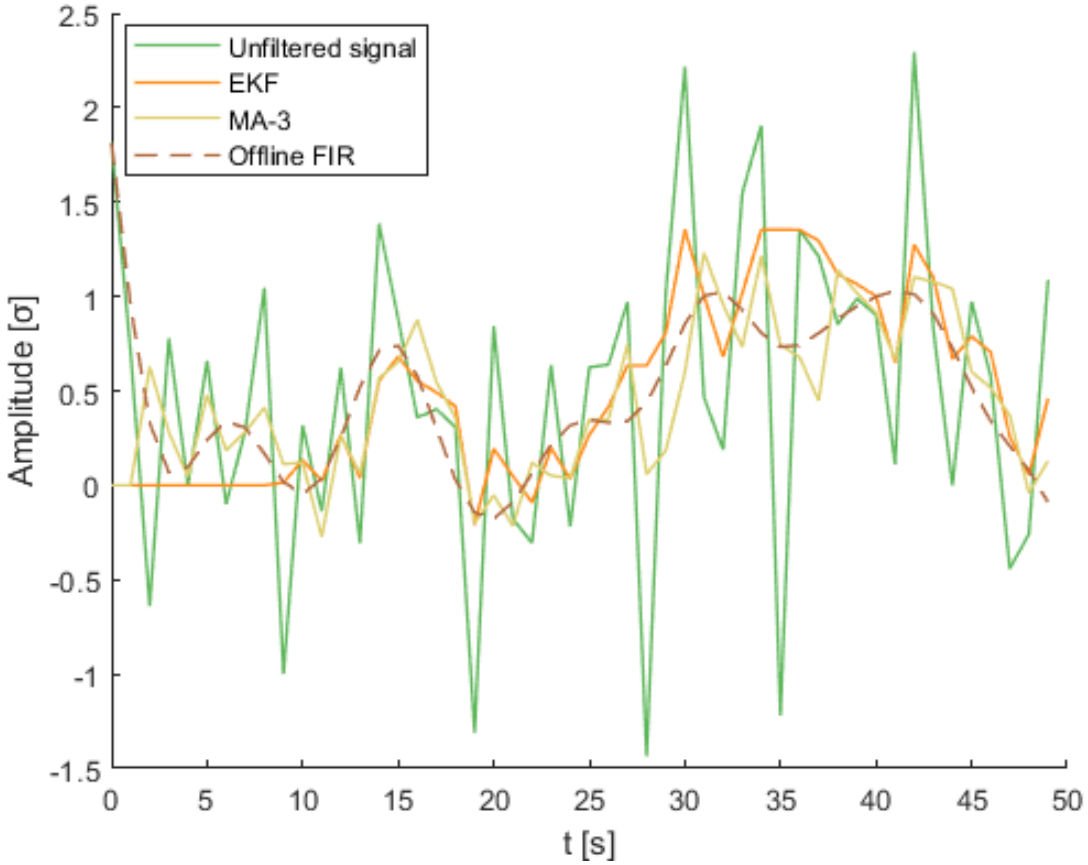

Fig. 6.2: Comparison of filtering properties of EKF and MA.

Latly, the figure contains for comparison an ideal offline smoother. As on offline smoother low-pass FIR of length 30 was used. The signal has been has been filtered twice. Once in both directions to remove filter's group delay. As can be seen at sample 35, although not completely filtering out parts of high-frequency noise at times EKF actually follows the underlying trend better, on account of the spike removing algorithm.

Greater lengths of subsequently allowed spikes were tested and the best results were obtained with maximum spike length of one. Which is consistent with the nature of spike-like artifacts and with low sampling frequencies used in fMRI neurofeedback studies.

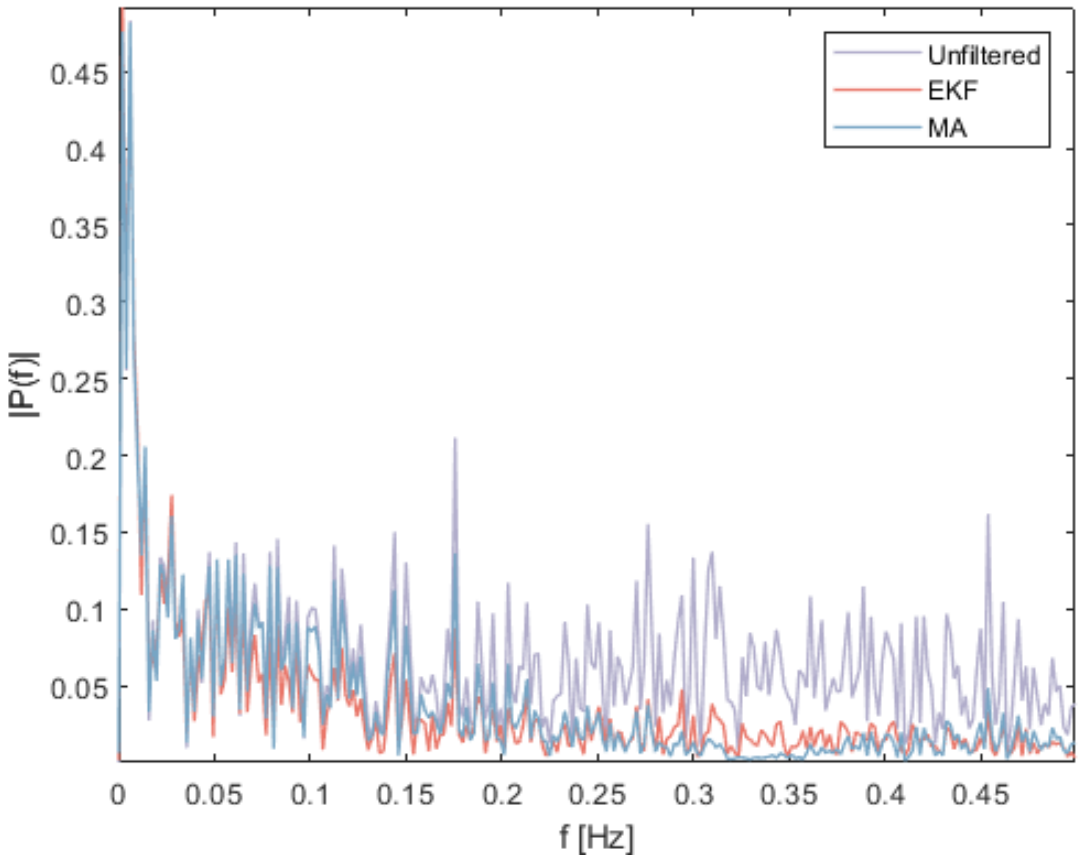

Fig. 6.3: Amplitude spectrum of neurofeedback signal before and after EK and MA filtering.

Comparison of both filters in the frequency domain can be seen in figure 6.3. MA filter seems to have a slow roll-off and low stopband attenuation. A rising arm of stopband attenuation ripple may be seen around 0.4 Hz. EKF also does not possess ideal low-pass filter properties. Amplitude attenuation leaves part of the spectrum not carrying useful information attenuated but not completely stopped. In contrast to the MA filter it also slightly attenuates lower frequencies, which is probably caused by the varying Kalman gain and so varying update steps.

7 Conclusion

The thesis pursues two main goals, both of them are closely related to real-time filtering of fMRI neurofeedback data: to shorten the delay at the beginning of each neurofeedback block introduced by digital filters and to analyze necessary learning period of the currently used filter at CEITEC MU, more specifically that of an extended Kalman filter.

In the theoretical part, the reader is familiarized with the basic working principles behind fMRI. Underlying biological processes involved in the formation of a functional image are explained along with properties and limitations of fMRI. One of such important limitations is the inherent delay between actual neuronal activity and the detection of a BOLD signal which can take up to 6 seconds. Next, the post-hoc data preprocessing methods and statistical tests are described. Chapter 3 builds on top of previous chapters by introducing fMRI neurofeedback and real-time fMRI data processing pipeline while also introducing current solution utilized at CEITEC MU. Real-time data processing is based on offline methods while introducing some new challenges and limitations. The next chapter briefly describes types, advantages, and limitations of various types of digital filters while also containing a more in-depth mathematical description of Kalman filters.

The practical part starts by chapter 5. Issues implicit in real-time filtering in general and others limited to fMRI had been identified and described. Next, In-depth analysis of current solution had revealed that the number of samples needed for Kalman to learn the signal had been overestimated. Further shortening the learning period would be possible but detrimental to solution robustness. Adjustments to be made had been described and temporary filter which can further shorten the transient state had been outlined. As a temporary filter MA of short length had been proposed which switches to EKF when the output is reliable. Although not ideal in terms of filter delay, its smoothing properties were found to be adequate with relation to EKF.

Lastly, the EKF filter and MA filter outputs have been compared. Wilcoxon rank sum statistical test was utilized to ensure volunteers will not see a dramatic change in values when filters are switched. The proposed solution saves about half a minute at the beginning of each neurofeedback block, about 1.5 - 3 minutes depending on the TR and number of NFB blocks in each session. This significantly reduces the time volunteers have to spend in claustrophobic space of MR scanner and reduces financial costs fMRI neurofeedback. These resources can be spend on prolonging the neurofeedback sessions or on further instructing the volunteer about the task at hand which could improve volunteer's learning curve.

Further tests and possible improvements could be made by running whole fMRI

neurofeedback experiments with low order filters such as MA and statistically comparing its neurofeedback effects compared to EKF. Results could indicate an underlying relationship between smoothing, filter delay, and volunteer's learning curves.

Bibliography

1. LINHARTOVÁ, Pavla; LÁTALOVÁ, Adéla; KÓŠA, Barbora; KAŠPÁREK, Tomáš; SCHMAHL, Christian; PARET, Christian. fMRI neurofeedback in emotion regulation: A literature review. *NeuroImage*. 2019, vol. 193, pp. 75–92. ISSN 10538119. Available from DOI: 10.1016/j.neuroimage.2019.03.011.

2. YOUNG, Kymberly D.; ZOTEV, Vadim; PHILLIPS, Raquel; MISAKI, Masaya; YUAN, Han; DREVETS, Wayne C.; BODURKA, Jerzy. Real-Time fMRI Neurofeedback Training of Amygdala Activity in Patients with Major Depressive Disorder. *PLoS ONE*. 2014, vol. 9, no. 2, pp. e88785. ISSN 1932-6203. Available from DOI: 10.1371/journal.pone.0088785.

3. BRÜHL, Annette Beatrix; SCHERPIET, Sigrid; SULZER, James; STÄMPFLI, Philipp; SEIFRITZ, Erich; HERWIG, Uwe. Real-time Neurofeedback Using Functional MRI Could Improve Down-Regulation of Amygdala Activity During Emotional Stimulation: A Proof-of-Concept Study. *Brain Topography*. 2014, vol. 27, no. 1, pp. 138–148. ISSN 0896-0267. Available from DOI: 10.1007/s10548-013-0331-9.

4. SCHARNOWSKI, Frank et al. Manipulating motor performance and memory through real-time fMRI neurofeedback. *Biological Psychology*. 2015, vol. 108, pp. 85–97. ISSN 03010511. Available from DOI: 10.1016/j.biopsycho.2015.03.009.

5. HUETTEL, Scott A; SONG, Allen W; MCCARTHY, Gregory. *Functional magnetic resonance imaging*. 1st ed. Sinauer Associates, 2004. ISBN 0878932887.

6. FARO, Scott H; MOHAMED, Feroze B. *Functional MRI: basic principles and clinical applications*. Springer Verlag, 2006.

7. BUXTON, Richard B. The physics of functional magnetic resonance imaging (fMRI). *Reports on Progress in Physics*. 2013. ISBN 1361-6633 (Electronic)\r0034-4885 (Linking). ISSN 00344885. Available from DOI: 10.1088/0034-4885/76/9/096601.

8. SAKA. Linear superposition of sensory-evoked and ongoing cortical hemodynamics. *Frontiers in Neuroenergetics*. 2010. ISBN 1662-6427 (Electronic) 1662-6427 (Linking). ISSN 16626427. Available from DOI: 10.3389/fnene.2010.00023.

9. ELSTER, Allen D. *Questions and Answers xn–in-f1t MRI*. BOLD and Brain Activity. 2018. Available also from: http://mriquestions.com/does-boldbrain-activity.html.

10. MENON, Ravi S.; OGAWA, Seiji; HU, Xiaoping; STRUPP, John P.; ANDERSON, Peter; UĞURBIL, Kămil. BOLD Based Functional MRI at 4 Tesla Includes a Capillary Bed Contribution: Echo-Planar Imaging Correlates with Previous Optical Imaging Using Intrinsic Signals. *Magnetic Resonance in Medicine*. 1995. ISBN 0740-3194. ISSN 15222594. Available from DOI: 10.1002/mrm.1910330323.

11. HU, Xiaoping; YACOUB, Essa. *NeuroImage*. The story of the initial dip in fMRI. 2012. ISBN 1053-8119. ISSN 10538119. Available from DOI: 10.1016/j.neuroimage.2012.03.005.

12. ZIJL, Peter C.M. van; HUA, Jun; LU, Hanzhang. *NeuroImage*. The BOLD post-stimulus undershoot, one of the most debated issues in fMRI. 2012. ISBN 1053-8119. ISSN 10538119. Available from DOI: 10.1016/j.neuroimage.2012.01.029.

13. GOENSE, Jozien; BOHRAUS, Yvette; LOGOTHETIS, Nikos K. fMRI at High Spatial Resolution: Implications for BOLD-Models. *Frontiers in Computational Neuroscience*. 2016. ISBN 1662-5188. ISSN 1662-5188. Available from DOI: 10.3389/fncom.2016.00066.

14. KANNURPATTI, Sridhar S.; BISWAL, Bharat B.; KIM, Young Ro; ROSEN, Bruce R. Spatio-temporal characteristics of low-frequency BOLD signal fluctuations in isoflurane-anesthetized rat brain. *NeuroImage*. 2008. ISBN 1053-8119. ISSN 10538119. Available from DOI: 10.1016/j.neuroimage.2007.05.061.

15. FMRI TEAM BRNO. *Funkční magnetická rezonance / fMRI Brno* [online] [visited on 2018-11-25]. Available from: http://fmri.mchmi.com/main%7B%5C_%7Dindex.php?strana=16.

16. LINDQUIST, Martin A. The Statistical Analysis of fMRI Data. *Statistical Science*. 2008, vol. 23, no. 4, pp. 439–464. ISSN 0883-4237. Available from DOI: 10.1214/09-STS282.

17. KOUSH, Yury; ZVYAGINTSEV, Mikhail; DYCK, Miriam; MATHIAK, Krystyna A.; MATHIAK, Klaus. Signal quality and Bayesian signal processing in neurofeedback based on real-time fMRI. *NeuroImage*. 2012. ISBN 1095-9572 (Electronic)\r1053-8119 (Linking). ISSN 10538119. Available from DOI: 10.1016/j.neuroimage.2011.07.076.

18. JENKINSON, Mark. *FSLMotionOutliers - FslWiki* [online]. 2018 [visited on 2018-12-17]. Available from: https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLMotionOutliers.

19. BRAINVOYAGER. *VTC inspector - BrainVoyager* [online]. 2018 [visited on 2018-12-17]. Available from: https://support.brainvoyager.com/brainvoyager/available-tools/86-available-plugins/371-vtc-inspector.

20. RAINER GOEBEL. *The General Linear Model (GLM)* [online]. 2014 [visited on 2018-11-25]. Available from: http://www.brainvoyager.com/bvqx/doc/UsersGuide/StatisticalAnalysis/TheGeneralLinearModel.html.

21. DEWIPUTRI, Wan Ilma; AUER, Tibor. *Malaysian Journal of Medical Sciences*. Functional magnetic resonance imaging (fMRI) neurofeedback: Implementations and applications. 2013. ISBN 1394-195X (Print) 1394-195X (Linking). ISSN 1394195X.

22. SORGER, Bettina; KAMP, Tabea; WEISKOPF, Nikolaus; PETERS, Judith Caroline; GOEBEL, Rainer. When the Brain Takes 'BOLD' Steps: Real-Time fMRI Neurofeedback Can Further Enhance the Ability to Gradually Self-regulate Regional Brain Activation. *Neuroscience*. 2018. ISBN 1873-7544 (Electronic)\r0306-4522 (Linking). ISSN 18737544. Available from DOI: 10.1016/j.neuroscience.2016.09.026.

23. DICOM. *DICOM Standard* [online] [visited on 2018-11-25]. Available from: https://www.dicomstandard.org/.

24. SMITH, Anne M.; LEWIS, Bobbi K.; RUTTIMANN, Urs E.; YE, Frank Q.; SINNWELL, Teresa M.; YANG, Yihong; DUYN, Jeff H.; FRANK, Joseph A. Investigation of Low Frequency Drift in fMRI Signal. *NeuroImage*. 1999, vol. 9, no. 5, pp. 526–533. ISSN 10538119. Available from DOI: 10.1006/nimg.1999.0435.

25. BLANDFORD, Dick Kent; PARR, John M. *Introduction to digital signal processing*. Pearson, 2013. ISBN 9780131394063. Available also from: http://www.mypearsonstore.com/bookstore/introduction-to-digital-signal-processing-9780131394063.

26. MATHURANATHAN. *Choosing a filter – FIR or IIR – Understanding the design perspective | GaussianWaves* [online]. 2017 [visited on 2018-12-21]. Available from: https://www.gaussianwaves.com/2017/02/choosing-a-filter-fir-or-iir-understanding-the-design-perspective/.

27. BISHOP, Gary; WELCH, Greg. An Introduction to the Kalman Filter. *Proc of SIGGRAPH, Course*. 2001. ISBN 9780582403857. ISSN 10069313. Available from DOI: 10.1.1.117.6808.

28. ULUSOY, Melda; MATHWORKS. *Understanding Kalman Filters - MATLAB* [online] [visited on 2018-12-17]. Available from: https://www.mathworks.com/videos/series/understanding-kalman-filters.html.

29. MATHWORKS. *Overview of Adaptive Filters and Applications - MATLAB* [online] [visited on 2018-12-21]. Available from: https://www.mathworks.com/help/dsp/ug/overview-of-adaptive-filters-and-applications.html.

30. STEFFENER, J.; LANGE, G.; BLY, B.M.; BISWAL, B.; REISMAN, S. Using an adaptive filter to extract the hemodynamic response from fMRI data. In: *2003 IEEE 29th Annual Proceedings of Bioengineering Conference*. IEEE, pp. 75–76. ISBN 0-7803-7767-2. Available from DOI: 10.1109/NEBC.2003.1215999.

List of symbols, physical constants and abbreviations

MRI	Magnetic resonance imaging
fMRI	Functional magnetic resonance imaging
BOLD	Blood oxygenation level dependent
SNR	Signal-to-noise ratio
FIR	Finite impulse response
IIR	Infinite impulse response
CEITEC	Central European Institute of Technology
GML	General linear model
rtfMRI	Real-time functional magnetic resonance imaging
NFB	Neurofeedback
EKF	Extended Kalman filter
MA	Moving average
MU	Masaryk University