Biophysical modeling of hemodynamic-based neuroimaging techniques

by

Louis Gagnon

Submitted to the Harvard-MIT Division of Health Sciences and Technology
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Medical Engineering and Medical Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2013

Author ....
Harvard-MIT Division of Health Sciences and Technology
July 23, 2013

Certified by ....
David A. Boas, Ph.D.
Professor of Radiology, Harvard Medical School
Thesis Supervisor

Accepted by ....
Emery N. Brown, M.D., Ph.D.
Director, Harvard-MIT Program in Health Sciences and Technology
Professor of Computational Neuroscience and Health Sciences and Technology

Biophysical modeling of hemodynamic-based neuroimaging techniques

by

Louis Gagnon

Submitted to the Harvard-MIT Division of Health Sciences and Technology
on July 23, 2013, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Medical Engineering and Medical Physics

Abstract

Two different hemodynamic-based neuroimaging techniques were studied in this work.

Near-Infrared Spectroscopy (NIRS) is a promising technique to measure cerebral hemodynamics in a clinical setting due to its potential for continuous monitoring. However, the presence of strong systemic interference in the signal significantly limits our ability to recover the hemodynamic response without averaging tens of trials. Developing a new methodology to clean the NIRS signal from systemic interference and isolate the cortical signal would therefore significantly increase our ability to recover the hemodynamic response opening the door for clinical NIRS studies such as epilepsy. Toward this goal, a new method based on multi-distance measurements and state-space modeling was developed and further optimized to remove systemic physiological oscillations contaminating the NIRS signal. Furthermore, the cortical and pial contributions to the NIRS signal were quantified using a new multimodal regression analysis.

Functional Magnetic Resonance Imaging (fMRI) based on the Blood Oxygenation Level Dependent (BOLD) response has become the method of choice for exploring brain function, and yet the physiological basis of this technique is still poorly understood. Despite the effort, a detailed and validated model relating the signal measured to the physiological changes occurring in the cortical tissue is still lacking. Modeling the BOLD signal is challenging because of the difficulty to take into account the complex morphology of the cortical microvasculature, the distribution of oxygen in those microvessels and its dynamics during neuronal activation. Here, we overcome this difficulty by performing Monte Carlo simulations over real microvascular networks and oxygen distributions measured in vivo on rodents, at rest and during forepaw stimulation, using two-photon microscopy. Our model reveals for the first time the specific contribution of individual vascular compartment to the BOLD signal, for different field strengths and different cortical orientations. Our model makes a new prediction: the amplitude of the BOLD signal produced by a given physiological change during neuronal activation depends on the spatial orientation of the cortical region in the MRI scanner. This occurs because veins are preferentially oriented either perpendicular or parallel to the cortical surface in the gray matter.

Thesis Supervisor: David A. Boas, Ph.D.
Title: Professor of Radiology, Harvard Medical School

Dedication

I dedicate my thesis to my mother Jocelyne and to my friend Nicolas, for their terrific example of courage during the last few years. You are an extraordinary source of motivation for me. I am sorry for not having been closer to you during the hard times.

Acknowledgments

This thesis would not have been possible without the help and the guidance of many people. At the top of the list is obviously David, my advisor. David has been actively involved in my research from the beginning and I feel very fortunate about that. He was key in the implementation of most of the work done during my PhD. Thank you for allowing me to graduate in four years and for allowing me to switch project in the middle of my graduate studies. Finally, thank you for all the parties and diners at your place.

I have been lucky to be mentored by extraordinary people who constantly encouraged me and made me a better scientist through insightful discussions. Special thanks to Elfar Adalsteinsson, Andrew Berger, Div Bolar, Emery Brown, Rick Buxton, Anna Devor, Doug Greve, Randy Gollub, Frederic Lesage, Maria Mody, Leif Ostergaard, Jonathan Polimeni, Patrick Purdon and Bruce Rosen.

I want to thank all the past and present members of the Optics Division at the Martinos Center for being such great colleagues. Thank you for sharing workspace, a coffee machine, computer codes, data sets and instrumentations with me. Thank you for providing feedback on my work and generating new ideas for my research.

I want to thank all the graduate students at the Martinos Center with whom I have been interacting over the last few years: Audrey, Bo, Christin, Dan, and JP. I also want to thank my HST colleagues for making grad school so unique. Special thanks to Alex, Dan, Gaurav, Raj and Tim for making ICM such a great experience.

I want to thank Professors Elfar Adalsteinsson, Sangeeta Bhatia, Jeff Drazen, Elazer Edelman, Rick Mitchell, Al Oppenheim, Robert Padera, Guido Sandri, Tim Swager and George Verghese for being such great teachers.

Thanks to my girlfriend, my family and all my friends for always supporting me and making me so happy. Special thanks to my roommate Aziz for entertaining me over the last years with various discussions on topics ranging from hockey to the cosmological constant problem.

Finally, I want to acknowledge financial support from the Fonds Quebecois de la Recherche sur la Nature et les Technologies (FQRNT), the IDEA-squared program at MIT and the Advanced Multimodal Neuroimaging Training Program at the Martinos Center. My research was also supported by NIH grants P41-RR14075, R90-DA023427, R01-EB000790 and R01-EB006385.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Why measuring the hemodynamic response of the brain?

The human brain is a very complex organ composed of billions of neurons. Detecting which neurons are involved in a specific cognitive process is a challenging task. In fact, the electromagnetic fields produced by the neurons when they discharge are very tiny. To be able to detect them, several neurons from a given location have to discharge synchronously. Moreover, the measurements of the electromagnetic fields produced by neurons can only be performed non-invasively at the surface of the head. This means that complex mathematical models have to be used to solve the inverse problem and to reconstruct the activation located in the brain. Given the high level of noise in electrophysiological recordings, this is a challenging task.

To overcome this drawback, researchers often measure the regional changes in blood oxygenation following neuronal activation, which are much easier to detect. These changes in blood oxygenation are the results of two competing mechanisms.

1. The first one is that increased firing rate requires more oxygen to be consumed

in order to maintain the ionic concentrations across the cell membrane. This decreases the oxygen saturation of the blood supplying the activated cognitive region.

2. The second mechanism is a feedback mechanism that dilates regionally the blood vessels supplying the activated cognitive region, increasing blood flow to this region of the brain. This mechanism is called neurovascular coupling. The increase in blood flow enhances the arrival of freshly oxygenated blood to the activated cognitive region and therefore increases blood oxygenation.

Generally, the second mechanism wins and the net result is a regional increase in blood oxygenation in the activated cognitive area. Detecting an increase in blood oxygenation during a cognitive task is a good indicator that this region was activated during the cognitive process. Using the same reasoning, we found that blood oxygenation decreases in regions where neuronal inhibition occurred. There are a lot of circumstances for which the above conclusion doesnt hold due to alteration to the neurovascular coupling. These circumstances are still the subject of intense research. Nevertheless, mapping task-evoked variations in oxygenation in the brain is an accurate way of mapping which region of the brain is involved in a specific cognitive process.

1.2 Two complementary ways of measuring task-evoked hemodynamics in the brain

The most widely used method to measure task-evoked blood oxygenation variations is Blood Oxygen Level Dependent functional Magnetic Resonance Imaging (BOLD-fMRI) [82, 102]. The advantages of this method are that it is relatively straightforward to implement on a conventional MRI scanner and it provides functional brain mapping with high spatial resolution (1 mm^3) and with uniform sensitivity across the entire brain. An inconvenient of this technique is that it is really difficult to relate the amplitude of the signal measured to the amplitude of the physiological changes occurring in the cortical tissue [16]. The principal reason is that the amplitude of the signal detected in a given voxel depends not only on the physiological changes but also on baseline physiology in this voxel (blood volume, blood oxygenation, blood flow, etc) [22, 61].

Another complementary technique to measure task-evoked oxygenation variations in the brain is Near-Infrared Spectrocopy (NIRS) [99, 59, 36]. As opposed to BOLD-fMRI, NIRS is a low-cost portable technique that has potential for long term bedside monitoring of brain functions in a clinical setting. NIRS is also less sensitive to motion artifact compared to fMRI which makes it the method of choice for cognitive studies on young children. Finally, NIRS also provide simultaneously changes in blood oxygenation as well as changes in blood volume. The major inconvenient of NIRS is its non-uniform sensitivity to different brain regions, with very high sensitivity to superficial region and no sensitivity to deeper brain regions [7]. This gives rise to two technical problems: (1) It is not possible to measure task-evoked hemodynamic variations in deep regions of the brain and (2) the signal measured is strongly contaminated by systemic physiology (unrelated to the cognitive process studied) occurring in the scalp. Moreover, the spatial resolution of NIRS is limited to 3 cm which introduces quantification error due to large pial vessels irrigating the surface of the cortex.

1.3 Why bother about quantitative hemodynamic measurements?

One could say that mapping brain function is a yes/no question. A brain region is either involved or not involved during a cognitive process. Well, that is not the entire story. The brain is never at rest. Information is continuously processed everywhere in the brain. Most cognitive processes will involve a lot of regions and one would get that all the brain gets activated during any cognitive task. It is much more interesting to studies how much each region is involved during a cognitive process. This by itself justifies the need for quantitative hemodynamic measurements.

Another point is that most cognitive studies involve multiple subjects. It becomes really hard to study cognitive processes in different subjects if the measurements depend on baseline physiology (i.e. whether the subject had coffee or not that morning). The same reasoning applies to clinical studies, where disease (or medication) can alter physiology. The lack of quantification in hemodynamic-based brain imaging largely limits the utilization of these techniques in the clinic.

Finally, quantitative hemodynamic imaging of the brain provides additional physiological information about the brain health. It is well known that oxygen supply is critical to brain functions [123, 112]. Measuring hemodynamic parameters in the brain in disease states or during healthy aging is an important research area.

1.4 Alternatives for quantitative measurements

Positron Emission Tomography (PET) is a very quantitative technique [39] that is already well established in the clinic [112]. However, PET doesnt allow measuring physiological changes occurring within seconds since measurements must be collected for several minutes before the image can be reconstructed. For the purpose of imaging brain functions during short cognitive processes, PET is not the method of choice. Moreover, the method requires the injection of radioactive contrast agents, which makes the method not suitable for repeated measurements.

Alternatively, there are a lot of fMRI sequences specifically designed for measuring quantitative physiology in the brain [57, 33, 9, 17]. These methods are really promising in the near future. At the moment, the signal-to-noise ratio (SNR) with these methods is much lower compared to BOLD-fMRI and this remains the main issue today. Moreover, these sequences must be validated against well-establish modalities (like PET) since some of them rely on a number of assumptions about cerebral physiology. Nevertheless, these new sequences constitute the future of brain imaging.

1.5 Importance of BOLD

The above section should have convinced you that BOLD-fMRI is still the method of choice in several applications [89]. The method is well established and high SNR is achieved. Variations of the order of 0.1 % in the signal measured are routinely detected. Such sensitivity is desirable to study resting-state brain functions, a new avenue in cognitive neuroscience. Moreover, BOLD-fMRI contains all the physiological information we want, but in a very convoluted way [16]. We are very good at measuring this signal. We only need to underpin the information contained in it.

These are the main reasons why a physiological interpretation of the BOLD signal is still relevant today, twenty years after the first BOLD-fMRI measurements. Although the technique has driven a revolution in neuroscience, the physiological basis of the signal is still poorly understood, preventing the technique to be used at its full power.

1.6 Importance of NIRS

As mentioned before, NIRS is a portable technique and allows for continuous monitoring. This opens the doors for specific clinical applications. A first example is pain monitoring during surgical procedure [3]. As you might imagine, a quantitative interpretation of the signal is imperative in this case. Another potential clinical application is hemodynamic monitoring during epilepsy [115]. It is very difficult to predict when patient will have seizure so continuous monitoring is a key advantage of NIRS here. Moreover, seizures usually occur with a lot of motion of the subjects head, requiring a method robust to motion artifact (another advantage of NIRS).

Quantitative NIRS in this case would allow monitoring oxygenation of brain tissue during seizures and indicate whether hypoxia is developing during the seizure.

1.7 Goal of the thesis

With these motivations established, the goals of this thesis are twofold. The first goal is **to improve the quantification of the NIRS signal by isolating the cortical signal**. The hypotheses are that state-space analysis combined short optode separation will help removing contamination from superficial tissue and that multimodal NIRS with BOLD-fMRI will allow to quantify the contamination by pial veins washout.

The second goal is **to develop a validated framework that can predict BOLD-fMRI signals from microscopic measurements of the underlying physiology**. The hypothesis is that combining Monte Carlo simulations with two-photon microscopy measurements of vascular morphology and pO2 during functional stimulation will predict the fMRI response measured in vivo.

1.8 Overview of the thesis

This thesis consists of two parts. In the first part, we focus on near-infrared spectroscopy (NIRS). Chapter 2 introduces a state-space framework to remove superficial physiological interference in NIRS data using short separation optode together with Kalman filtering. In Chapter 3, we use this framework to study how the systemic physiology spreads over the surface of the human head. We then improve our method in Chapter 4 by modifying the optical probe to get short separation measurements both at the source and the detector location. In Chapter 5, we combined NIRS with BOLD-fMRI to quantify the contribution of pial vein washout in the NIRS signal.

In the second part of the thesis, we move to microscopic optical imaging (multiphoton microscopy and optical coherence tomography (OCT)) and try to model the BOLD-fMRI signal with the best possible accuracy from the microscopic measurements. In Chapter 6, we introduce a method to model the BOLD-fMRI signal with Monte Carlo simulations over the microscopic measurements. Finally in Chapter 7, we study how the use of OCT measurements can constrain the flow reconstruction in our modeling framework.

Chapter 2

State-space modeling for NIRS

This section was publisehd in:

Gagnon, L., Perdue, K., Greve, D.N., Goldenholz, D., Kaskhedikar, G. and Boas, D.A. (2011). "Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling." *NeuroImage* 56(3): 1362-1371.

In the present study, we combined small separation measurements and state-space modeling for the estimation of the hemodynamic response and simultaneous global interference cancellation. We developed both a static and a dynamic estimator. We evaluated the performance of our algorithms using baseline data taken from 6 human subjects at rest and by adding a synthetic hemodynamic response over the baseline measurements. We finally compared our new methods with the adaptive filter [147] and the standard method using no small SD separation measurement.

2.1 Introduction

Diffuse optical imaging (DOI) is an experimental technique that uses near-infrared spectroscopy (NIRS) to image biological tissue [138, 99, 47, 59, 63]. The dominant chromophores in this spectrum are the two forms of hemoglobin: oxygenated hemoglobin (HbO) and reduced hemoglobin (HbR). In the past 15 years, this technique has been used for the noninvasive measurement of the hemodynamic changes associated with evoked brain activity [138, 63].

Compared with other existing functional imaging methods e.g., functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET), Electroencephalography (EEG), and Magnetoencephalography (MEG), the advantages of DOI for studying brain function include good temporal resolution of the hemodynamic response, measurement of both HbO and HbR, nonionizing radiation, portability, and low cost. Disadvantages include modest spatial resolution and limited penetration depth.

The sensitivity of NIRS to evoked brain activity is also reduced by systemic physiological interference arising from cardiac activity, respiration, and other homeostatic processes [98, 133, 105, 29]. These sources of interference are called global interference or systemic interference. Part of the interference occurs both in the superficial layers of the head (scalp and skull) and in the brain tissue itself. However, the back-reflection geometry of the measurement makes NIRS significantly more sensitive to the superficial layers. As such, the NIRS signal is often dominated by systemic interference occurring in the skin and the skull.

Different methods have been used in the literature to remove the systemic interference from DOI measurements. Low pass filtering is widely used in the literature, as it is highly effective at removing cardiac oscillations [40, 74]. However, there is a significant overlap between the frequency spectrum of the hemodynamic response to brain activity and the spectrum of other physiological variations such as respiration, spontaneous low frequency oscillations and very low frequency oscillations. Frequency-based removal of these sources of interference can therefore result in large distortion and inaccurate timing for the recovered brain activity signal. As such, more powerful methods for global noise reduction have been developed. These include adaptive average waveform subtraction [49], subtraction of another NIRS source-detector (SD) channel performed over a non-activated region of the brain [40], principal component analysis [145, 41] and finally wavelet filtering [84, 94, 73, 85].

A recent development for removing global interference from NIRS measurements is to use additional optodes in the activated region with small SD separations that are sensitive to superficial layers only [117, 147, 148, 146, 136, 141, 51]. Making the assumption that the signal collected in the superficial layers is dominated by systemic physiology which is also dominant in the longer SD separation NIRS channel, those additional measurements can be used as regressors to filter systemic interference from the longer SD separations. Saager et al [116] used additional optodes and a linear minimum mean square estimator (LMMSE) to partially remove the systemic interference in the signal. In a second step, the evoked hemodynamic response was estimated using a traditional block-average method over the different trials. The algorithm was further refined by Zhang et al [147, 148, 146] to consider the non-stationary behavior of the systemic interference. They used an adaptive filtering technique together with additional small separation measurements to filter the systemic interference from the raw signal and then performed the block-average technique to estimate the hemodynamic response in a second step.

Although these methods greatly reduced global interference in NIRS data, the filtering of the systemic interference and the estimation of the hemodynamic response were performed in two steps, which might not be optimal. Previous studies have shown that the simultaneous estimation of the hemodynamic response and removal of the systemic interference using temporal basis functions [81, 111] or auxiliary systemic measurements [28] was possible using state-space modeling. Moreover, Diamond et al proposed a way to quantify the accuracy of such filtering methods. Real NIRS data collected over the head of human subjects at rest were used to generate realistic noise. A synthetic hemodynamic response was added over the real NIRS baseline time course and the response was then recovered from this noisy data set. The recovered response was then compared with the synthetic one used to generate the time course. This method for evaluating reconstruction algorithms has been reproduced by other groups [84, 94, 85].

In the present study, we combined small separation measurements and state-space modeling for the estimation of the hemodynamic response and simultaneous global interference cancellation. We developed both a static and a dynamic estimator. We evaluated the performance of our algorithms using baseline data taken from 6 human subjects at rest and by adding a synthetic hemodynamic response over the baseline measurements. We finally compared our new methods with the adaptive filter [147] and the standard method using no small SD separation measurement.

2.2 Methods

2.2.1 Experimental data

For this study, 6 healthy adult subjects were recruited. The Massachusetts General Hospital Institutional Review Board approved the study and all subjects gave written informed consent. Subjects were instructed to rest while simultaneous BOLD-fMRI and NIRS data were collected. Three 6-minute long runs were collected for each subject. Only the NIRS data was used in this study. The localization and the geometry of the NIRS probe used are shown in Fig. 4-1 a) and b) respectively. Only the two 1 cm SD separation channels and the 8 closest neighbor (3 cm SD separation) channels were used in the analysis.

Changes in optical density for each SD pair were converted to changes in hemoglobin concentrations using the Beer-Lambert relationship [19, 24, 7] and the SD distances

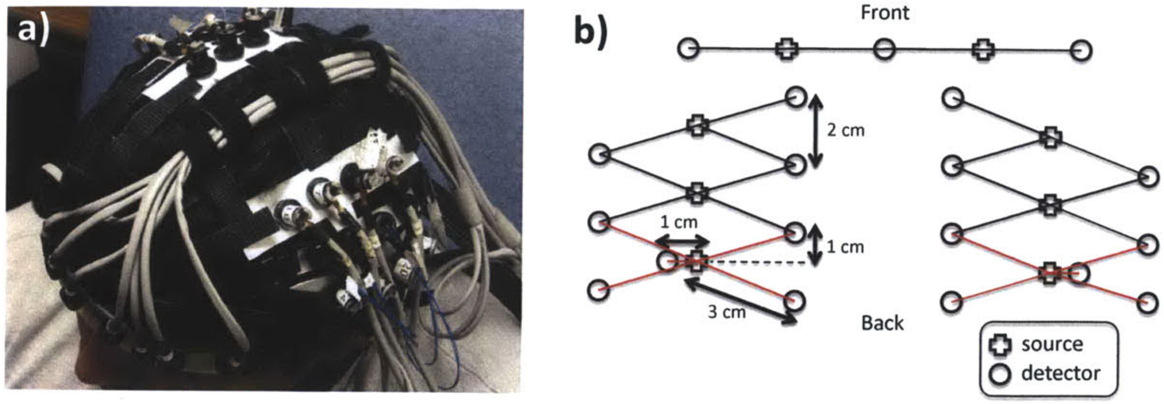

Figure 2-1: a) Position of the probe over the head of the subjects b) Geometry of the optical probe. Two different SD separations were used: 1 cm and 3 cm. The NIRS channels used for the analysis are shown in red.

illustrated in Fig. 4-1 b). A pathlength correction factor of 6 and a partial volume correction factor of 50 were used for all SD pairs [69, 70].

2.2.2 Synthetic hemodynamic response

To compare the performance of our two algorithms with existing algorithms, a synthetic hemodynamic response was generated using a modified version of a three compartment biomechanical model [68, 66, 65]. Each parameter of the model was set to the middle of its physiological range [68] which results in an HbO increase of 15 μM and an HbR decrease of 7 μM. The amplitude of this synthetic response was of the same order as real motor responses on humans using NIRS and those specific pathlength and partial volume correction factors [70]. These synthetic HbO and HbR responses were then added to the unfiltered concentration data with an inter-stimulus interval taken randomly from a uniform distribution (10-35 s) for each individual trial. Over the six-minute data series, we added either 10, 30 or 60 individual evoked responses. The resulting HbO and HbR time courses were then highpass filtered at 0.01 Hz to remove any drifts and lowpass filtered at 1.25 Hz to remove the instrument noise. The filter used was a 3$^{\text{rd}}$ order Butterworth-type filter.

Four different methods were then used to recover the simulated hemodynamic response added to our baseline data. The first two were taken directly from literature and consisted of the standard General Linear Model (GLM) without using a small SD separation measurement and the adaptive filtering (AF) method developed by Zhang et al [147]. The third one was a simultaneous static deconvolution and regression and will be called the static estimator (SE) here for simplicity. The last one was a dynamic Kalman filter estimator (KF).

2.2.3 Signal modeling

For all the methods used in this study, the discrete-time hemodynamic response h at sample time n was reconstructed with a set of temporal basis functions

$$h[n] = \sum_{i=1}^{N_w} w_i b_i [n] \tag{2.1}$$

where $b_i[n]$ are normalized Gaussian functions with a standard deviation of 0.5 s and their means separated by 0.5 s over the regression time as shown in Fig. 2-2 a). N_w is the number of Gaussian functions used to model the hemodynamic response and was set to 15 in our work. Using this set, the noise-free simulated HbO response was fit with a Pearson R^2 of 1.00 and a mean square error (MSE) of 9.2×10^{-5} and the noise-free simulated HbR response was fit with an R^2 of 1.00 and an MSE of 2.1×10^{-5}. The MSE was lower for HbR only because the amplitude of the simulated HbR response was lower. These fits are shown in Fig. 2-2 b). The weights for the temporal bases w_i were estimated using the four different methods described in the following sections.

For the standard block average estimator, we modeled the concentration signal in the 3 cm separation channel $y_3[n]$ by

$$y_3[n] = \sum_{k=-\infty}^{\infty} h[k]\, u[n-k]\,. \tag{2.2}$$

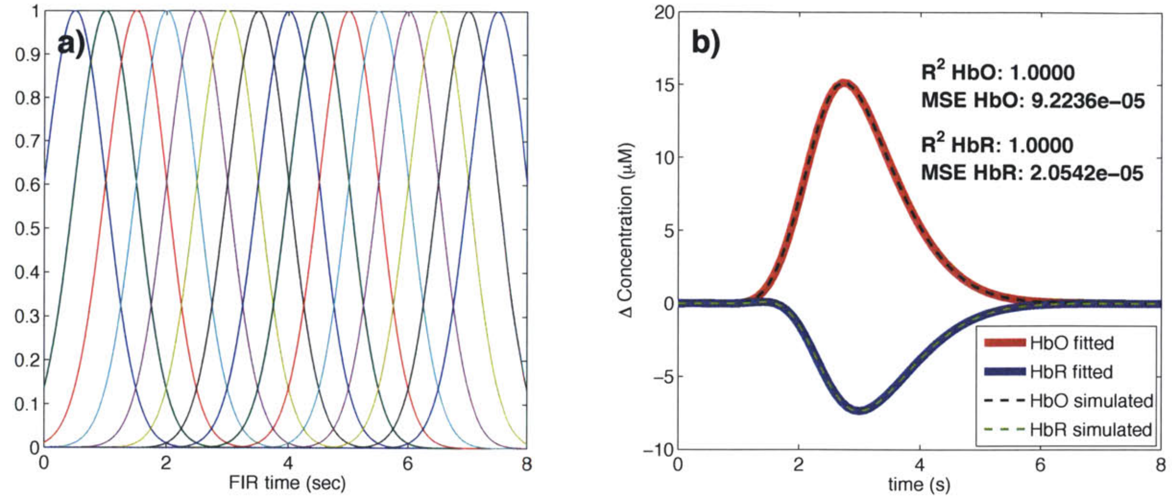

Figure 2-2: a) Temporal basis set used in the analysis. The finite impulse response (FIR) of the temporal basis functions ranged from 0 to 8 s after the onset of the simulated response. b) Noise-free simulated responses (dotted lines) overlapped with the responses recovered with a least-square fit (continuous lines) using the temporal basis set. The R^2 and the MSE of the fit are indicated for both HbO and HbR.

$u[n]$ is called the onset vector and is a binary vector taking the value 1 when n corresponds to a time where the stimulation starts and 0 otherwise.

For our static simultaneous estimator and our dynamic Kalman filter simultaneous estimator, we modeled the signal in the 3 cm separation channel $y_3[n]$ by a linear combination of the 1 cm separation signal $y_1[n]$ and the hemodynamic response $h[n]$ by

$$y_3[n] = \sum_{k=-\infty}^{\infty} h[k]\, u[n-k] + \sum_{i=1}^{N_a} a_i y_1[n+1-i]. \tag{2.3}$$

N_a is the number of time points taken from the 1 cm separation channel to model the superficial signal in the 3 cm separation channel. This value was set to 1 in our work for all three estimators using short SD separation measurements but could be any integer in principle. The a_i's are the weights used to model the superficial signal in the 3 cm separation channel from the linear combination of the 1 cm separation signal. The states to be estimated by the static and the Kalman filter estimators were the weights for the superficial contribution a_i and the weights for the temporal bases w_i. All those weights were assumed stationary in the case of the static estimator, and time-varying in the case of the Kalman filter estimator.

The motivation for Eq. 4.2 is that the residual between the 3 cm channel and the 1 cm channel corresponds to the hemodynamic response of the brain. This is well justified when the brain activation is detected only in the 3 cm separation channel and when the systemic physiology pollutes both the 1 cm and the 3 cm separation channels. It is a reasonable assumption for cognitive NIRS measurements performed on an adult head. In this case, the hemodynamic response is expected to occur only in the brain tissue and the 1 cm separation channel does not reach the cerebral cortex, making the 1 cm measurement sensitive to scalp and skull fluctuations only. This would also be justified for cognitive measurements on babies by reducing the separation of the 1 cm signal to ensure that this channel remains insensitive to brain hemodynamics. However, our assumption would be violated for specific stimuli (e.g. the Valsalva maneuver) for which the hemodynamic response occurs more globally across the head. Other scenarios that could be troublesome would be if the systemic physiology occurs only in the brain tissue (e.g. an activation-like oscillation a few seconds after the true stimulus response) or if the interference is phase-locked with the stimulus. In this case, the systemic physiology could potentially be modeled by our temporal basis set (overfitting).

2.2.4 Standard General Linear Model

For this first method, and only for this one, the 1 cm SD separation channels were not used. The pre-filtered concentrations from the 3 cm SD separation were further lowpass filtered at 0.5 Hz using a 3$^{\text{rd}}$ order Butterworth filter. Re-expressing Eq. 2.2 in matrix form, we get

$$\mathbf{y}_3 = \mathbf{U}\mathbf{w} \tag{2.4}$$

where $\mathbf{y}_3$ is simply the length N_t time course vector $y_3\left[n\right]$

$$\mathbf{y}_3 = \left[\begin{array}{ccc} y_3\left[1\right] & \ldots & y_3\left[N_t\right] \end{array} \right]^T . \tag{2.5}$$

The columns of $\mathbf{U}$ are the linear convolution of the onset vector $u[n]$ with each temporal basis function $b_i[n]$

$$\mathbf{U} = \begin{bmatrix} u * b_1[n] & \cdots & u * b_{N_w}[n] \end{bmatrix} \tag{2.6}$$

and $\mathbf{w}$ is the vector containing the weights for the temporal basis w_i

$$\mathbf{w} = \begin{bmatrix} w_1 & \ldots & w_{N_w} \end{bmatrix}^T. \tag{2.7}$$

The estimates of the weights $\hat{\mathbf{w}}$ are found by inverting Eq. 2.4 using the Moore-Penrose pseudoinverse

$$\hat{\mathbf{w}} = \left(\mathbf{U}^T\mathbf{U}\right)^{-1}\mathbf{U}^T\mathbf{y}_3 \tag{2.8}$$

and the hemodynamic response is finally reconstructed with the estimates of the temporal basis weights $\hat{w}_i$ obtained from $\hat{\mathbf{w}}$.

When the GLM was used without any other estimator (i.e. not as the last step of the adaptive filter or the Kalman filter), we included a 3rd order polynomial drift as a regressor. This procedure is used regularly in fMRI analysis. In this case, the matrix $\mathbf{U}$ is expanded

$$\mathbf{G} = \begin{bmatrix} \mathbf{U} & \mathbf{D} \end{bmatrix} \tag{2.9}$$

where $\mathbf{D}$ is an N_t by 4 drift matrix given in the 2.6. The estimates of the weights $\hat{\mathbf{w}}$ are found by inverting

$$\hat{\mathbf{w}} = \left(\mathbf{G}^T\mathbf{G}\right)^{-1}\mathbf{G}^T\mathbf{y}_3. \tag{2.10}$$

2.2.5 Adaptive filtering

The adaptive filtering technique was taken directly from [147]. Only the salient points are outlined here. The HbO and the HbR responses were recovered independently and the adaptive filter was used for both. The two pre-filtered concentration signals at 1 cm (y_1) and 3 cm (y_3) were first normalized with respect to their respective standard deviation. This was to ensure that the standard deviation of the two signals used in the computation were close 1 to accelerate the convergence of the algorithm [147]. The output of the filter, $e[n]$, is then given by

$$e[n] = y_3[n] - \sum_{k=0}^{N_a} w_{k,n}\, y_1[n-k] \tag{2.11}$$

where the coefficient of the filter, $w_{k,n}$, is updated via the Widrow-Hoff least mean square algorithm [55]:

$$w_{k,n} = w_{k,n-1} + 2\mu e[n-1]\, y_1[n-k]. \tag{2.12}$$

In our study, w was initialized at $w_{k,1} = [1\ 0\ 0\ ...]^T$ and μ was set to $1\text{x}10^{-4}$ as in [147]. After trying different values for N_a, we identified $N_a = 1$ as the value minimizing the MSE between our simulated and recovered hemodynamic responses. The output $e[n]$ was then multiplied by the original standard deviation of y_3 to rescale it back to its original scale. The output of the filter was then further lowpass filtered at 0.5 Hz and the hemodynamic response was finally estimated using the standard GLM method (with no drift) by substituting $\mathbf{y_3}$ by $\mathbf{e}$ in Eq. 2.8

$$\hat{\mathbf{w}} = \left(\mathbf{U}^T\mathbf{U}\right)^{-1}\mathbf{U}^T\mathbf{e} \tag{2.13}$$

where $\mathbf{e}$ is simply the length N_t time course vector $e[n]$

$$\mathbf{e} = \begin{bmatrix} e[1] & \dots & e[N_t] \end{bmatrix}^T \tag{2.14}$$

and again the hemodynamic response is finally reconstructed with the estimates of the temporal basis weights $\hat{w}_i$ obtained from $\hat{\mathbf{w}}$.

2.2.6 Static estimator

Our static estimator is an improved version of the linear minimum mean square estimator (LMMSE) developed by Saager et al [116, 117]. In their work, they used the small separation signal and an LMMSE to estimate the contribution of the superficial signal in the large separation signal. This superficial contamination was then removed from the large separation signal and the hemodynamic response was then estimated from the residual (large separation signal without the superficial contamination). In our study, we simultaneously removed the contribution of the superficial signal in the 3 cm separation signal and estimated the hemodynamic response.

Eqs. 4.2 and 4.1 can be re-expressed in matrix form

$$\mathbf{y}_3 = \mathbf{A}\mathbf{x} \tag{2.15}$$

where $\mathbf{y}_3$ is the vector representing the signal in the 3 cm channel and is given by Eq. 2.5, $\mathbf{x}$ is the concatenation of the w_i's and a_i's

$$\mathbf{x} = \begin{bmatrix} w_1 & \dots & w_{N_w} & a_1 & \dots & a_{N_a} \end{bmatrix}^T \tag{2.16}$$

and $\mathbf{A}$ is the concatenation of the N_t by N_w matrix $\mathbf{U}$ given by Eq. 4.7 and the N_t by N_a matrix $\mathbf{Y}$

$$\mathbf{A} = \begin{bmatrix} \mathbf{U} & \mathbf{Y} \end{bmatrix} \tag{2.17}$$

where

$$\mathbf{Y} = \begin{bmatrix} y_1[1] & 0 & \dots \\ y_1[2] & y_1[1] & 0 \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{2.18}$$

The first N_w columns of $\mathbf{A}$ are the linear convolution of the onset vector $u[n]$ with each temporal basis function $b_i[n]$ and the last N_a columns of $\mathbf{A}$ are simply the signal from the 1 cm separation channel $y_1[n]$ delayed by one more sample in each column. In order to compare the different estimators on the same footing, N_a was set to 1 for all three estimators using short SD separations. A more explicit expression for $\mathbf{A}$ is given in 2.6. The estimates of the weights $\hat{\mathbf{x}}$ are found by inverting Eq. 2.15 using the Moore-Penrose pseudoinverse

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{y}_3 \tag{2.19}$$

and the hemodynamic response is finally reconstructed with the estimates of the temporal basis weights $\hat{w}_i$ obtained from $\hat{\mathbf{x}}$. This reconstructed response was further lowpass filtered at 0.5 Hz.

2.2.7 Kalman filter estimator

For our dynamic Kalman filter estimator, Eqs. 4.2 and 4.1 need to be re-express in state-space form:

$$\mathbf{x}[n+1] = \mathbf{I}\mathbf{x}[n] + \mathbf{w}[n] \tag{2.20}$$

$$y_3[n] = \mathbf{C}[n]\,\mathbf{x}[n] + \mathbf{v}[n] \tag{2.21}$$

where $\mathbf{w}[n]$ and $\mathbf{v}[n]$ are the process and the measurement noise respectively. $\mathbf{x}[n]$ is the sample n of $\mathbf{x}$ given by Eq. 4.6, $\mathbf{I}$ is an $N_w + N_a$ by $N_w + N_a$ identity matrix and $\mathbf{C}[n]$ is an $N_w + N_a$ by 1 vector whose entries correspond to the n^{th} row of $\mathbf{A}$ in Eq. 2.17. The estimate $\hat{\mathbf{x}}[n]$ at each sample n is then computed using the Kalman filter [77] followed by the Rauch–Tung–Striebel smoother [113]. The Kalman filter recursions require initialization of the state vector estimate $\hat{\mathbf{x}}[0]$ and estimated state covariance $\mathbf{P}[0]$. In our study, the initial state vector estimate $\hat{\mathbf{x}}[0]$ was set to the values obtained using our static estimator and the initial state covariance estimate $\mathbf{P}[0]$ was set to an identity matrix with diagonal entries of $1\text{x}10^{-1}$ for the temporal basis states and $5\text{x}10^{-4}$ for the superficial contribution state. The Kalman filter algorithm was run a first time to estimate the initial state covariance and then run a second time. The initial covariance estimate for the second run was set to the final covariance estimate of the first run. Running the filter twice makes the method less sensitive to the initial guess $\mathbf{P}[0]$. Statistical covariance priors must also be specified for the state process noise $\operatorname{cov}(\mathbf{w}) = \mathbf{Q}$ and the measurement noise $\operatorname{cov}(\mathbf{v}) = \mathbf{R}$. The process noise determines how big the states are allowed to vary at each time step. If this value is small, the estimator will approach the static estimator. If it is large, the state will be allowed to vary significantly over time. In this work, the process noise covariance only contained nonzero terms on the diagonal elements. Those diagonal terms were set to $2.5\text{x}10^{-6}$ for the temporal basis state and $5\text{x}10^{-6}$ for the superficial contribution states. This imbalance in state update noise was also used by Diamond et al [28] and caused the functional response model to evolve more slowly than the superficial contribution model. Practically, the measurement noise determine how well we trust the measurements during the recovery procedure. In our study, the measurement noise covariance was set to an identity matrix scaled by $5\text{x}10^{-2}$. Different values have been tried for the process noise and the measurement noise covariances. Changing the value of $\mathbf{Q}$ and $\mathbf{R}$ over two orders of magnitude did not result in notable performance changes and we could have drawn all the same conclusions presented in this paper using these alternative $\mathbf{Q}$ and $\mathbf{R}$ values. The values for $\mathbf{Q}$ and $\mathbf{R}$ presented above were empirically determined to minimize the MSE between the recovered and the simulated hemodynamic response. The algorithm was then processed with the following prediction-correction recursion [46].

Since the state update matrix is the identity matrix in Eq. 4.4, the state vector $\mathbf{x}$ and state covariance $\mathbf{P}$ are predicted with

$$\hat{\mathbf{x}}[n|n-1] = \hat{\mathbf{x}}[n-1|n-1] \tag{2.22}$$

$$\hat{\mathbf{P}}[n|n-1] = \hat{\mathbf{P}}[n-1|n-1] + \mathbf{Q}. \tag{2.23}$$

The Kalman gain $\mathbf{K}$ is then computed

$$\mathbf{K}[n] = \hat{\mathbf{P}}[n|n-1]\,\mathbf{C}[n]^T \left(\mathbf{C}[n]\,\hat{\mathbf{P}}[n|n-1]\,\mathbf{C}[n]^T + \mathbf{R}\right)^{-1} \tag{2.24}$$

and the state vector $\mathbf{x}$ and state covariance $\mathbf{P}$ predictions are corrected with the most recent measurements $y_3[n]$

$$\hat{\mathbf{x}}[n|n] = \hat{\mathbf{x}}[n|n-1] + \mathbf{K}_n\left(y_3[n] - \mathbf{C}[n]\,\hat{\mathbf{x}}[n|n-1]\right) \tag{2.25}$$

$$\hat{\mathbf{P}}[n|n] = \left(\mathbf{I} - \mathbf{K}[n]\,\mathbf{C}[n]\right)\hat{\mathbf{P}}[n|n-1]\,. \tag{2.26}$$

After the Kalman algorithm was applied twice, the Rauch–Tung–Striebel smoother was applied in the backward direction. With the identity matrix as the state-update matrix in Eq. 4.4, the algorithm is given by [56]:

$$\hat{\mathbf{x}}[n|N_t] = \hat{\mathbf{x}}[n|n] + \hat{\mathbf{P}}[n|n]\,\hat{\mathbf{P}}[n+1|n]^{-1}\left(\hat{\mathbf{x}}[n+1|N_t] - \hat{\mathbf{x}}[n+1|n]\right). \tag{2.27}$$

The complete time course of the estimated hemodynamic response $\hat{h}[n]$ was then reconstructed for each sample time n using the final state estimates $\hat{\mathbf{x}}[n|N_t]$ and the temporal basis set contained in $\mathbf{C}[n]$

$$\hat{h}[n] = \mathbf{C}[n]\,\hat{\mathbf{x}}[n|N_t]\,. \tag{2.28}$$

This reconstructed hemodynamic response time course $\hat{h}[n]$ was further lowpass filtered at 0.5 Hz and the standard GLM estimator (with no polynomial drift) was then applied

$$\hat{\mathbf{w}} = \left(\mathbf{U}^T\mathbf{U}\right)^{-1}\mathbf{U}^T\hat{\mathbf{h}} \tag{2.29}$$

where $\mathbf{U}$ is the matrix defined in Eq. 4.7 and

$$\hat{\mathbf{h}} = \begin{bmatrix} \hat{h}[1] & \ldots & \hat{h}[N_t] \end{bmatrix}^T \tag{2.30}$$

to obtain the final weights $\hat{w}_i$ used to reconstructed the final estimate of the hemodynamic response. We observed that these last filtering and averaging steps further improved the estimate of the hemodynamic response compared to reconstructing the hemodynamic response from the final state estimates of the smoother.

2.2.8 Statistical analysis

Only specific channels based on the following criteria were kept in the analysis. The raw hemoglobin concentrations were bandpass filtered with a 3^{rd} order Butterworth-type filter between 0.01 Hz and 1.25 Hz [148]. The Pearson correlation coefficient R^2 between each 1 cm HbO channel and its 4 closest neighbor 3 cm HbO channels (before adding the synthetic hemodynamic response) were then computed and the SD pairs for which $R^2 < 0.1$ were discarded for the analysis. The mean R^2 across the selected channels was 0.47 for HbO and 0.22 for HbR. We also computed the Pearson correlation coefficient after adding the synthetic hemodynamic response and similar results were obtained. The mean differences between the R^2's computed before and after adding the synthetic response was 0.01 for HbO and 0.003 for HbR, with the highest value obtained before adding the synthetic response to the real data. Those small differences emphasize the fact that the signals were dominated by systemic physiology in our simulations. This result also suggests that no resting state measurement is required to select the channels which would benefit from the small separation measurement since the correlation can be estimated from the time course containing brain activation. Zhang et al [146] showed that the adaptive filter method was working well when the correlation between the short and the long separation channel for HbO was greater than 0.6. We used 0.1 in this work to include more channels in the analysis and to show that our state-space method was working well when the initial correlation was lower than 0.5. Using this criterion, 94 out of the 144 possible channels (6 subjects × 3 runs × 8 channels) were kept for further analysis. This represented 65 % of the original data set. The numbers of channels kept for each of the subjects were 16, 14, 13, 17, 19 and 15 respectively. The signal to noise ratio (SNR) for each channel was computed as the amplitude of the simulated hemodynamic response divided by the standard deviation of the time course of the signal. The mean SNR across the selected channels was 0.45 for HbO and 0.38 for HbR.

We used two different metrics to compare the performance of the different algorithms. The first one was the Pearson correlation coefficient R^2 between the true synthetic hemodynamic response and the recovered response given by each algorithm. This metric was used to access the level of oscillation in the recovered hemodynamic response created by the global interference not removed by the algorithms and still contaminating the signal. Since the R^2 coefficient is scale invariant, it could not give any information about the accuracy of the amplitude of the recovered hemodynamic response. To overcome this problem, we also used the mean square error (MSE) as a metric to compare the performance of the different algorithms.

Since the random position of the trials across the same time course can greatly affect the accuracy of the recovered hemodynamic response, we repeated the procedure 30 times with 30 different random onset time instances for each of the 94 selected channels. The mean and the standard deviation of the 2820 R^2 coefficients (94 channels × 30 instances) for each algorithm were then computed after applying the Fisher transformation

$$z = \tanh^{-1}\left(R^2\right) \tag{2.31}$$

and the results were then inverse transformed. The mean and the standard deviation of the 2820 MSEs were also computed. This procedure was repeated independently for 10, 30 and 60 trials in each six-minute data series. The different algorithms were compared together by computing two-tailed paired t-tests on their MSEs and Fisher transformed R^2 coefficients.

2.3 Results

Typical time courses of the recovered hemodynamic response overlapped with the true simulated response are shown in Fig. 2-3 a) to d) for the four algorithms tested. The SNR for this particular simulation was 0.33 for HbO and 0.81 for HbR. The R^2's and the MSEs for HbO and HbR are shown in the legend of each individual panel.

Those individual results were obtained from a single simulation with 10 trials. The time courses for this specific simulation are shown in panel e) for HbO and f) for HbR. Both the initial 1 cm channel and the 3 cm channel containing the added synthetic hemodynamic responses are shown as well as the position of the 10 individual onset times. The R^2 between the initial 1 cm channel and the initial 3 cm channel (no response added) is also shown in the legend of panel e) and f) for HbO and HbR respectively. All concentrations are expressed in micromolar (μM) units.

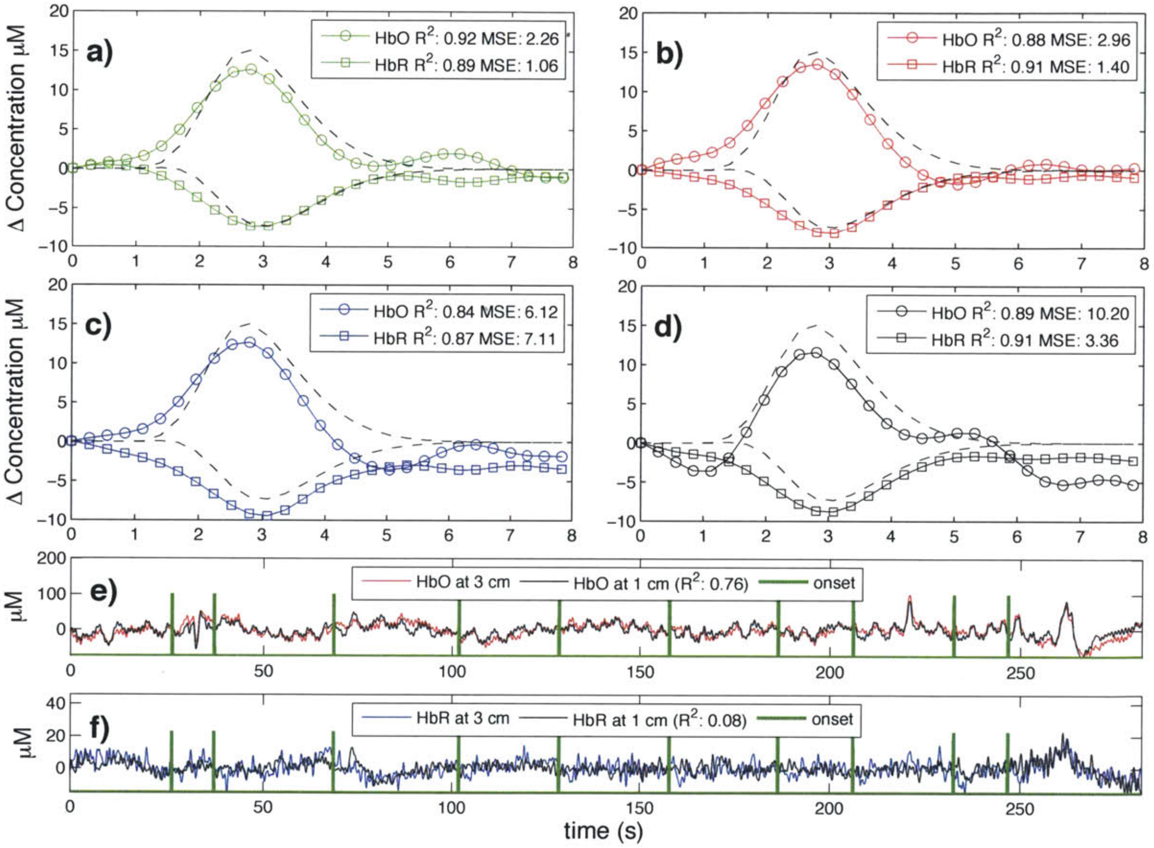

Figure 2-3: a) to d) Typical time courses of the recovered hemodynamic responses overlapped with the simulated hemodynamic response. For these specific traces, the SNR was 0.33 for HbO and 0.81 for HbR. R^2 coefficients and MSEs between the recovered (circles) and the simulated (dashed) response are shown in the legends. a) Kalman filter estimator b) Static estimator c) Adaptive filter d) Standard GLM with 3^{rd} order drift. e) HbO and f) HbR time courses of the 3 cm channel (with synthetic responses added) overlapped with the 1 cm channel. The positions of the onset time are also shown and the correlation coefficients between the 1 cm and the 3 cm channels (before adding synthetic responses) are indicated in parenthesis.

The summary R^2 statistics over all subjects, all channels and all instances are shown in a bar graph in Fig. 2-4 for both HbO and HbR. These values represent the Pearson R^2 coefficients computed between the recovered and the simulated hemodynamic responses. The bars represent the mean and the error bars represent the standard deviation. Both the mean and the standard deviation were computed on the Fisher transformed values and then inverse transformed. Two-tailed paired t-tests on the Fisher transformed values were performed between all the different estimators and statistical significance at the level $p < 0.05$ is illustrated by a black line over the bars for which a significant difference was observed. In our three simulations using 10, 30 and 60 trials respectively, the R^2's for HbO and HbR obtained using our Kalman filter dynamic estimator were significantly higher ($p < 0.05$) than the ones obtained using the adaptive filter. Moreover, the R^2's obtained were higher with the Kalman filter than with the static estimator. These differences were significant ($p < 0.05$) except in our 10 trial simulation for HbO.

Similarly, the summary MSE statistics over all subjects, all channels and all instances are shown in Fig. 2-5. These values represent the mean square error computed between the recovered and the simulated hemodynamic responses. The bars represent the mean while the error bars represent the standard deviation. Two-tailed paired t-tests were performed between all the different estimators and statistical significance at the level $p < 0.05$ is illustrated by a black line over the bars for which a significant difference was observed. The MSEs obtained for HbO and HbR in our three simulations (10, 30 and 60 trials) were significantly lower ($p < 0.05$) with our Kalman filter estimator than with the adaptive filter. Futhermore, the MSEs obtained with the Kalman filter were also lower ($p < 0.05$) than the ones obtained with the static estimator for both HbO and HbR in our three simulations.

Table 2.1 summarizes the statistical analysis over all the subjects, all the channels and all the instances for both HbO and HbR and for the simulations with 10, 30 and 60 trials. Each algorithm was compared to every other. The values shown are the p-values obtained from a two-tailed paired t-test. Statistical differences at the level $p < 0.05$ are indicated with bold script. These p-values were computed from the data

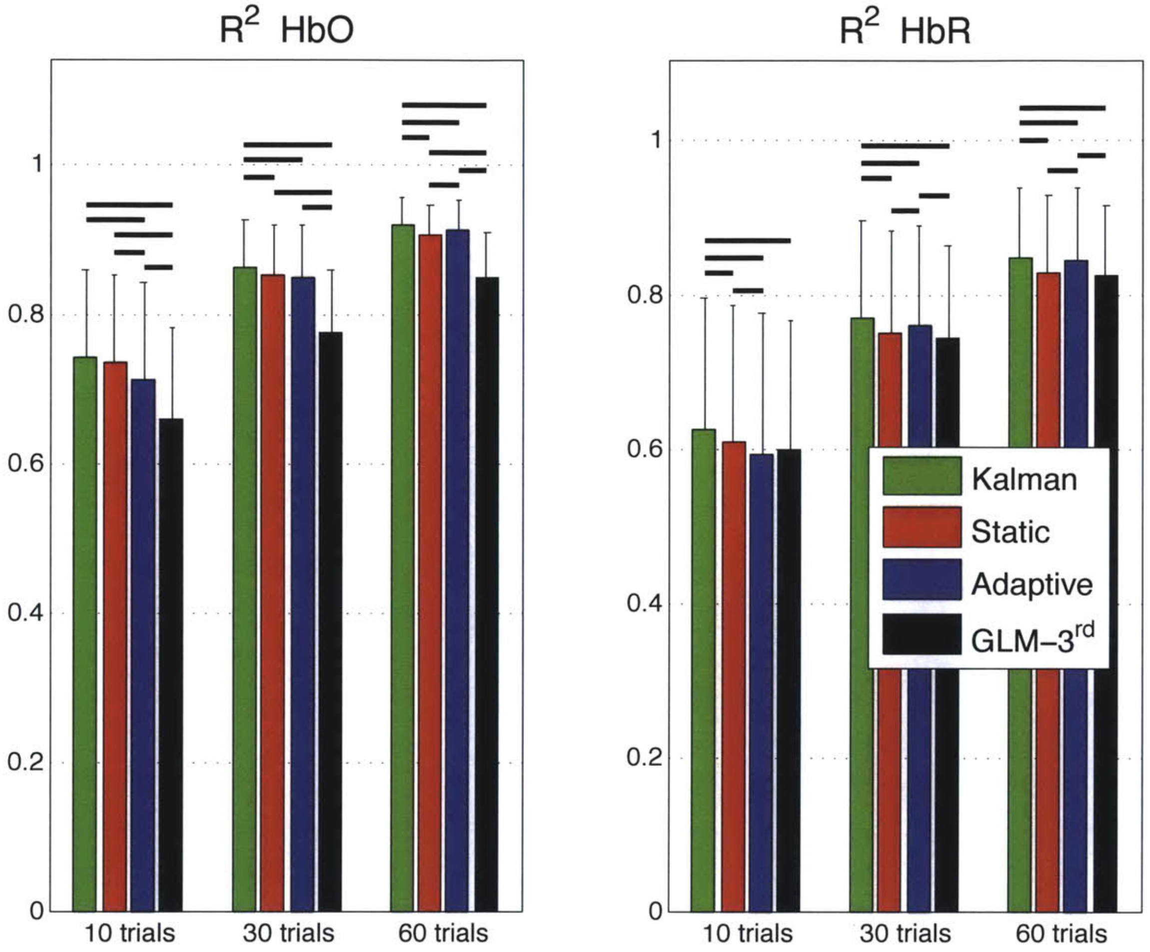

Figure 2-4: Pearson R^2 coefficients between simulated and recovered hemodynamic responses. The bars represent the means and the error bars represent standard deviations computed accross all subjects, all channels and all intances. The means and the standard deviation were computed in the Fisher space and then inverse transformed. Two-tailed paired t-tests were performed on the Fisher transformed R^2's. Statistical differences ($p < 0.05$) between the four algorithms are indicated by black horizontal lines over the corresponding bars.

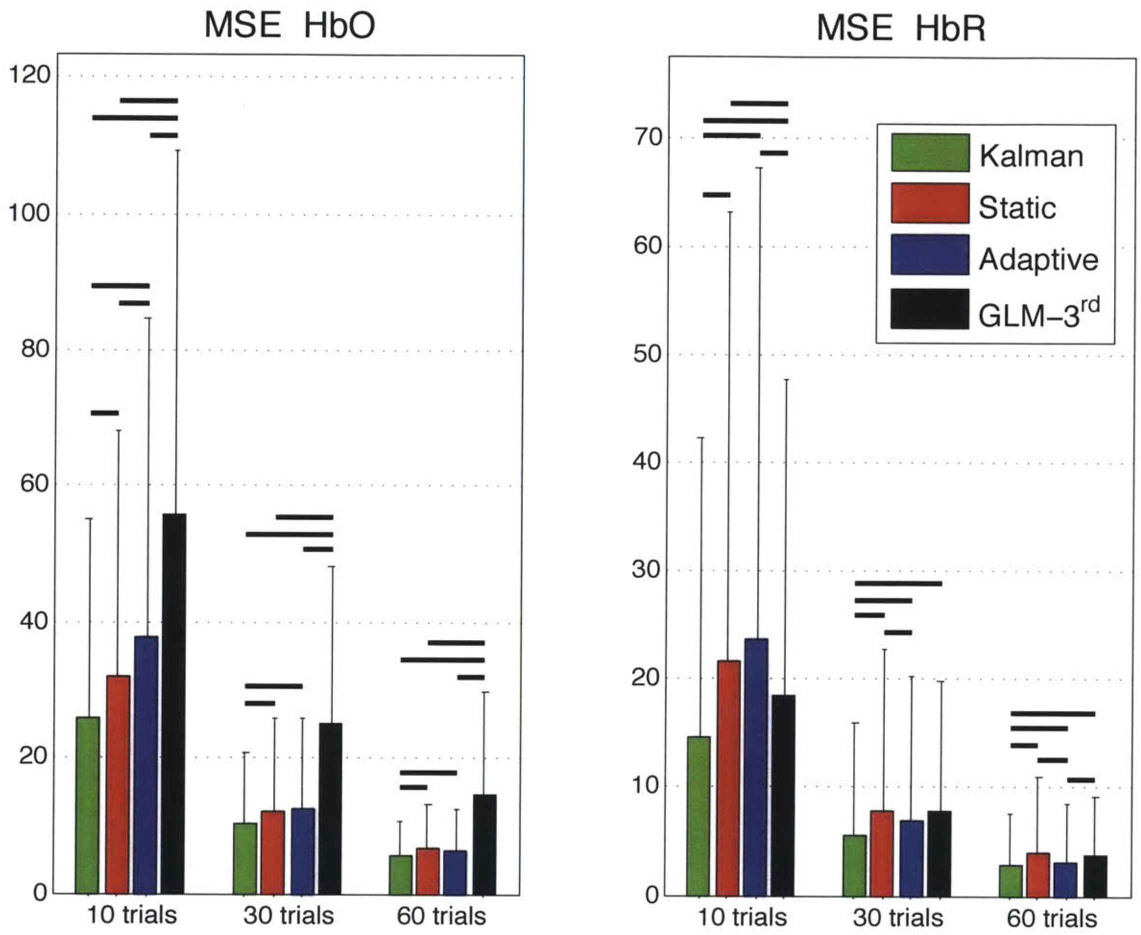

Figure 2-5: Mean squared errors (MSE) between simulated and recovered hemodynamic responses. The bars represent the means and the error bars represent the standard deviations computed accross all subjects, all channels and all instances. Two-tailed paired t-tests were performed between the four estimators and statistical differences at the level $p < 0.05$ are indicated by black horizontal lines over the corresponding bars.

summarized in the bar graphs shown in Figs. 2-4 and 2-5.

Table 2.1: Cross-comparison of the different algorithms. P-values for the two-tailed paired t-tests accross all subjects, all channels and all intances are shown. For the R^2 coefficients, the tests were performed on the Fisher transformed values. **Bold face** indicates significant difference at the $p < 0.05$ level. KF: Kalman filter estimator, SE: Static estimator, AF: Adaptive filter, GLM: Standard GLM with 3^{rd} order drift.

	10 trials			30 trials			60 trials		
	KF	SE	AF	KF	SE	AF	KF	SE	AF
R^2 HbO									
SE	6e-02	-	-	**3e-03**	-	-	**2e-07**	-	-
AF	**2e-07**	**6e-06**	-	**1e-03**	5e-01	-	**4e-04**	**2e-03**	-
GLM	**3e-15**	**2e-13**	**4e-07**	**8e-16**	**2e-12**	**4e-13**	**6e-15**	**1e-10**	**5e-14**
R^2 HbR									
SE	**5e-05**	-	-	**2e-09**	-	-	**1e-08**	-	-
AF	**2e-07**	**2e-03**	-	**3e-05**	**5e-03**	-	**4e-02**	**6e-10**	-
GLM	**5e-04**	2e-01	**4e-01**	**7e-04**	3e-01	**4e-02**	**1e-04**	6e-01	**7e-04**
MSE HbO									
SE	**4e-06**	-	-	**3e-05**	-	-	**4e-06**	-	-
AF	**2e-05**	**7e-03**	-	**1e-06**	7e-01	-	**2e-05**	6e-02	-
GLM	**2e-12**	**2e-09**	**5e-05**	**1e-13**	**7e-11**	**1e-11**	**4e-10**	**3e-08**	**2e-09**
MSE HbR									
SE	**2e-05**	-	-	**4e-05**	-	-	**1e-04**	-	-
AF	**3e-05**	2e-01	-	**2e-04**	**4e-03**	-	**7e-03**	**7e-05**	-
GLM	**6e-08**	**3e-02**	**1e-02**	**3e-07**	9e-01	9e-02	**4e-06**	1e+00	**5e-04**

2.4 Discussion

2.4.1 Simultaneous filtering and estimation

One of the salient features of our Kalman filter estimator is that it filters the global interference and simultaneously estimates the hemodynamic response. This feature resulted in a more accurate recovery of the hemodynamic response with our Kalman filter estimator compared to the adaptive filter, for which the filtering and the estimation were performed in two distinct steps. Independent regression of the small separation channel potentially removes contributions of the hemodynamic response in the signal which lead to an underestimation of the hemodynamic response thereafter.

Our Kalman filter estimator avoids this pitfall. Compared to the adaptive filter, our Kalman filter estimator showed significant improvements at the $p < 0.05$ level in both HbO and HbR recoveries for our 10, 30 and 60 trial simulations. Those improvements were observed in both Pearson R^2 and MSE metrics.

2.4.2 Dynamic versus static estimation

The systemic interference present in NIRS data is non-stationary. This has been nicely shown by Lina et al [84] who performed a detailed wavelet analysis of resting NIRS data with blood pressure, respiratory and heart rate data acquired simultaneously on awake human subjects. The amplitude of the systemic physiology measured by the 1 cm and the 3 cm channel depends on the respective pathlength of the light for each channel. Systemic physiology could alter the optical properties of the tissue over time. As a result, a sustained change in absorption could modify the pathlength of the light independently in the 1 cm and the 3 cm channel, modifying at the same time the relative amplitude of the systemic physiology detected in each channel. This feature of the systemic interference explains why our Kalman filter, which is a dynamic estimator, performed better than the static estimator. Using our Kalman filter estimator, improvements in the HbO and HbR recovery were observed in both the Pearson R^2 and the MSE metrics compared to the static estimator. All these improvements were significant at the $p < 0.05$ level except for the HbO Pearson R^2 improvement which was not significant in our 10 trial simulation.

2.4.3 HbO versus HbR

In their wavelet analysis, Lina et al [84] also showed that the HbO time courses were more contaminated by global interference than the HbR time courses. As such, the correlation between the 1 cm and 3 cm channel should be higher for HbO than HbR, and filtering methods using 1 cm SD separations should work better for HbO than for HbR. In our data, the mean initial Pearson R^2 correlation between the 1 cm and 3 cm signals were higher for HbO than HbR (0.47 vs 0.22). Comparing our Kalman filter estimator with the standard block average estimator, the p-values obtained in the t-tests performed on the Fisher transformed Pearson R^2's and the MSEs were at least five orders of magnitude lower for HbO than HbR. This indicates that the improvements observed with our Kalman filter were more prominent for HbO than HbR. This better performance in the recovery of HbO over HbR using a small separation method was also reported by Zhang et al [146] using their adaptive filter.

2.4.4 Impact of initial correlation

In the case where the systemic physiology present in the 3 cm separation did not correlate with the systemic physiology present in the 1 cm channel, the performance of the Kalman filter was similar to the standard GLM. In this case, the model cannot reproduce the data and the a_i coefficients in Eq. 4.2 converge to zero. As such, the w_i's estimated by the Kalman filter are very close to the ones obtained using the GLM. An important point is that in the case of low initial R^2 coefficients ($0.1 < R^2 < 0.2$), taking into account the 1 cm channel with the Kalman filter did not decrease the performance of the recovery compared to the GLM. On the other hand, the performance of the adaptive filter for ($0.1 <$ initial $R^2 < 0.2$) was worst than the GLM. This counter-performance of the adaptive filter for poor initial correlation between the short and the long channel was also reported by Zhang et al [146]. These findings suggest that the Kalman filter can be used even if the correlation between the 1 cm and the 3 cm channel is low as opposed to the adaptive filter. In the worst case, the Kalman filter will be as good as the standard GLM. However, the higher the initial correlation between the 1 cm and the 3 cm channel is, the more significant is the improvement using a small separation measurement. This is illustrated by the larger improvement obtained for HbO than HbR when using a small separation measurement together with our Kalman filter.

2.4.5 Technical notes

The MSEs obtained in our simulations and presented in Fig. 2-5 were lower for HbR than HbO. This occurred because the amplitude of the simulated HbR response was lower than the simulated HbO response which resulted in lower MSEs for HbR. This is illustrated for noise-free data in Fig. 2-2b.

For all the results presented in this paper, a single time point was taken from the 1 cm channel to regress the 3 cm channel. In practice, this value could be any integer. A simple phase shift (delay) between the 3 cm and 1 cm channel would be taken into account by using multiple time points from the 1 cm. In this case, all the a's in Eq. 4.2 would converge to zero except for one a at the value of i corresponding to the shift between the two signals in terms of number of sample points. Different values for N_a were tested during our simulations. With the adaptive filter, we obtained better results using a single point than using 100 points as in Zhang et al [147]. Using 100 points results in overfitting the signal which removes more of the hemodynamic response contribution than using a single point. This is another pitfall of the non-simultaneous recovery and filtering feature of the adaptive filter which is avoided with our Kalman filter. Finally, we did not observe any improvement when using multiple points with our Kalman filter, suggesting that no delays were present in our data between the 1 cm and the 3 cm channel.

The Gaussian temporal basis functions used in this work allow us to model different hemodynamic responses with different shapes and components. This includes a potential initial dip and post-stimulus undershoot, responses with a double bump and negative responses. It is also easy to use additional Gaussian functions to extend this method for longer stimuli, making the temporal basis set used in the present work very general and less restrictive. However, as stated in section 2.2.3, the drawback for using a more general set is the potential overfitting of phase-locked systemic physiology. This could be avoided using a more restrictive temporal basis set such as a gamma-variant function and its derivatives [67, 1, 64, 48], and at the same time could potentially reduce the number of parameters to estimate.

We tested different values for the separation between the basis and also different values for the width of the Gaussians. The values of 0.5 second for both the separation and the width presented in this paper resulted in the lowest MSEs between the recovered and the simulated responses and highest R^2's. The separation between our temporal basis Gaussians and their widths was three times lower than the values used by Diamond et al [28].

In order to compare the four methods used in this work on the same footing, we used temporal basis functions for each estimator. For the standard GLM estimator, the adaptive filter and the Kalman filter, we have also tried to replace the final step of using the GLM with a temporal basis set by a simple block average without using any temporal prior. For all these three estimators, using temporal basis functions in the final step further improved the recovery of both HbO and HbR. The MSEs between the recovered and the simulated hemodynamic response were lower when temporal basis were used than when a simple block average without temporal basis was applied. Similarly, the R^2's computed between the recovered and the simulated responses were higher when temporal basis were used in the final block average step. This result raises the importance of using temporal priors to reduce the dimensionality of the estimation problem.

As stated in section 2.2.7, changing the state process noise and the measurement noise priors over two orders of magnitude did not affect the performance of our Kalman estimator. For HbO, no differences could be observed (two-tailed paired t-test, $p < 0.05$) between the MSEs recovered using values for the process noise or the measurement noise ten times lower or higher than the ones presented in section 2.2.7. For HbR, small differences in the MSEs were observed but these results did not change any conclusions drawn in this paper. The MSEs recovered with our Kalman filter in this case were still the lowest of the four estimators.

2.4.6 Future directions

As mentioned in Zhang et al [146], an important question is whether an additional short separation optode is required for each longer separation optode or whether a single one is sufficient. Although the systemic interference is thought to be global in the brain, it might be reflected differently in the NIRS data collected over different regions of the head. Sources of variation include blood vessel size which might affect the amplitude of the recovered response but also blood vessel length and geometry which might give rise to phase mismatches between different NIRS channels. Studies using multiple small SD separation optodes at different locations over the head should be performed in the future to address this question.

2.5 Summary

In summary, we filtered the global interference present in NIRS data by using additional small separation optodes and we simultaneously estimated the hemodynamic response using a dynamic algorithm. Our dynamic Kalman filter performed better than the traditional adaptive filter, the static estimator and the standard block average estimator for both HbO and HbR recovery. These results were consistent with the fact that dynamic estimation better captures the non-stationary behavior of the systemic interferences in NIRS and that the simultaneous filtering and estimation prevents underestimation of the hemodynamic response. The algorithm is easily implementable and suitable for a wide range of NIRS studies.

2.6 Appendix: Design matrix

The explicit expression for $\mathbf{D}$ in Eq. 2.9 is given by

$$\mathbf{D} = \begin{bmatrix} 1 & 1/N_t & 1^2/N_t^2 & 1^3/N_t^3 \\ 1 & 2/N_t & 2^2/N_t^2 & 2^3/N_t^3 \\ 1 & 3/N_t & 3^2/N_t^2 & 3^3/N_t^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & N_t/N_t & N_t^2/N_t^2 & N_t^3/N_t^3 \end{bmatrix}.$$

The dimension of the matrix $\mathbf{D}$ is N_r by 4. Each column is normalized by its highest value to keep the matrix $\mathbf{G}$ well conditioned and to avoid numerical errors during the inversion in Eq.2.10.

The explicit expression for $\mathbf{A}$ in Eq. 2.17 is given by

$$\mathbf{A} = \begin{bmatrix} b_1[1] & b_2[1] & \dots & b_{N_w}[1] & y_1[1] & 0 & \dots & 0 \\ b_1[2] & b_2[2] & \dots & b_{N_w}[2] & y_1[2] & y_1[1] & & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_1[N_b] & b_2[N_b] & \dots & b_{N_w}[N_b] & \vdots & \vdots & & y_1[1] \\ 0 & 0 & & 0 & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 & \vdots & \vdots & & \vdots \\ b_1[1] & b_2[1] & \dots & b_{N_w}[1] & \vdots & \vdots & & \vdots \\ b_1[2] & b_2[2] & \dots & b_{N_w}[2] & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_1[N_b] & b_2[N_b] & \dots & b_{N_w}[N_b] & \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & y_1[N_t] & y_1[N_t-1] & \dots & y_1[N_t-N_a+1] \end{bmatrix}.$$

N_b is the length of each temporal basis function and was 80 in our work due to the 10 Hz temporal resolution and 8 s FIR for our temporal basis functions. The vertical dimension of matrix **A** corresponds to N_t, the total number of time points in the entire time course. The number of copies of the temporal basis functions corresponds to the number of trials (or stimuli) in the specific time course (i.e. if the run contained 10 trials, then 10 copies of the temporal basis set will appear in the corresponding **A** matrix).

Chapter 3

Impact of the short channel location

This section was published in:

Gagnon, L., Cooper, R. J., Yucel, M. A., Perdue, K., Greve, D. N., and Boas, D. A. (2011). "Short separation channel location impacts the performance of short channel regression in NIRS." *NeuroImage* 59: 2518-2528

The main contribution of this chapter is to quantify the performance of the short separation method as a function of the relative distance between 3 cm NIRS channels containing the brain signal and 1 cm channels used as a regressors. We investigated this relationship with both simulations and real functional data. NIRS measurements including several short separation channels spread across the probe were acquired on 6 human subjects. The simulations were performed by adding a synthetic hemodynamic response to the resting-state NIRS data. NIRS signals were also collected during a series of finger tapping blocks for each of the 6 subjects. In both cases, the performance of the short separation regression was characterized for different short SD regressors located at different distances from the standard 3 cm channel.

3.1 Introduction

Over the past 15 years, Near-Infrared Spectroscopy (NIRS) [138, 99, 47, 59, 63] has emerged as a complement to functional Magnetic Resonance Imaging (fMRI) for mapping the hemodynamic response associated with cerebral activity. NIRS non-invasively measures the temporal variations of the two dominant chromophores in the near-infrared window: oxygenated hemoglobin (HbO) and deoxygenated or reduced hemoglobin (HbR).

The advantages of NIRS for the investigation of brain activity include the measurement of both HbO and HbR concentrations, its low cost, and its portability. The portability of NIRS enables long-term monitoring of the hemodynamic reponse associated with, for instance, epilectic activity at the bedside [115]. Disadvantages of NIRS include modest spatial resolution of the order of one to three centimeters and limited penetration depth [7].

A common problem with NIRS recordings is the presence of strong physiology-based systemic interference in the signal which reduces the accuracy of NIRS for detecting brain activation. This interference arises from cardiac activity, respiration and other homeostatic processes [98, 133, 29, 105]. The contribution of this interference in the NIRS signal is amplified because the light is both introduced and collected at the surface of the scalp. This back-reflection geometry makes NIRS very sensitive to the superficial layers of the head which contain no brain signal but exhibit strong systemic fluctuations. As such, the NIRS signal is often dominated by systemic interference occurring in the superficial layers of the head including the scalp and the skull.

Several methods have been described in the literature to remove the systemic interference from NIRS measurements. Some post-processing algorithms include low pass filtering [40, 74], principal component analysis [145, 41] and wavelet filtering [84, 94, 73, 85]. Multi-distance NIRS measurements with layered models and path length weighted methods have also being investigated [136, 137, 141]. Other methods include the subtraction of a NIRS channel acquired in a non-activated region of the brain from the signal of interest to reduce the systemic interference [40].

A more refined version of this method is to simultaneously collect additional NIRS measurements using short source-detector (SD) separation channels (generally shorter than 1 cm), which are sensitive to superficial layers only [116]. Assuming that the signal collected with these additional short separation measurements is dominated by the same systemic interference present in the longer SD channels, the small separation signals can be used as regressors to filter the systemic interference from the longer SD measurements. Several algorithms have been developed to perform the regression of the small separation measurements. These include linear minimum mean square estimation (LMMSE) [116, 117, 51, 118], adaptive filtering [147, 148, 146] and state-space modeling with Kalman filter estimation [44].

An important question which was not addressed in these previous papers is the impact that the relative location of the short and long SD channels has on the performance of the short separation method. If good performance is obtained using a short separation channel located far away from the standard long SD channel, then a single short separation channel can be used as a regressor for all longer SD channels on the probe. On the other hand, the performance of the short separation method potentially worsen as the relative distance between the short and the long SD channel increases. In this case several short separation channels would be required and only those closest to the long SD channels would be suitable for regression.

The main contribution of this paper is to quantify the performance of the short separation method as a function of the relative distance between long SD NIRS channels (3 cm) containing the brain signal and short separation (1 cm) channels used as regressors. We investigated this relationship with both simulations and real functional data. NIRS measurements including several short separation channels spread across the probe were acquired on 6 human subjects. The simulations were performed by adding a synthetic hemodynamic response to the resting-state NIRS data. NIRS signals were also collected during a series of finger tapping blocks for each of the 6 subjects. In both cases, the performance of the short separation regression was characterized for different short SD regressors located at different distances from the standard 3 cm channel.

3.2 Methods

3.2.1 Experimental data

For this study, 6 healthy adult subjects were recruited. The Massachusetts General Hospital Institutional Review Board approved the study and all subjects gave written informed consent. Data were collected using a TechEn CW6 system operating at 690 and 830 nm. The NIRS probe contained 5 sources and 12 detectors as shown in Fig. 4-1a. This source-detector geometry resulted in 14 long SD measurements (3 cm) and 7 short SD measurements (1 cm). A set of 200 μm-core fibers was used for the short separation detector optodes to avoid saturation of the photodiode. These fibers are illustrated in orange in Fig. 4-1. An alternative to avoid photodiode saturation could be the use of standard NIRS fibers with optical filters at the tip of the probe to attenuate light intensity. The probe was secured over the left motor region of each subject as illustrated in Fig. 4-1b. One of the short separation measurements was acquired over the forehead. In this probe, the relative distances center-to-center between the short and the long channels take the values 1.4, 1.7, 2.4, 3.3, 4.2, 5.2 or 6.2 cm. Examples are given for each case in Fig. 4-1c. The forehead short separation channel was located more than 10 cm away from any 3 cm channel.

During the experiment, subjects were sitting in a comfortable chair in front of a computer screen with a black background. The functional runs were divided as shown in Fig. 3-2. Each run lasted 390 seconds and contained six blocks of 30 s finger tapping interleaved with 30 s resting blocks. Three functional runs were acquired for

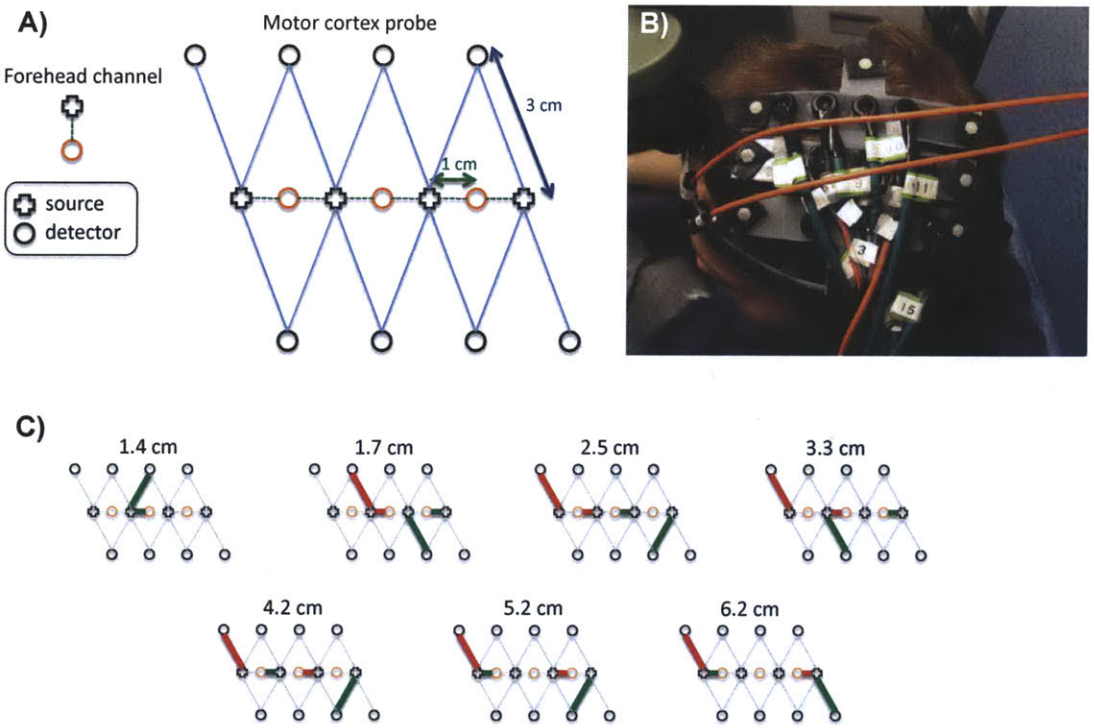

Figure 3-1: (A) Geometry of the optical probe. Two different SD separations were used: 1 cm and 3 cm. (B) Location of the probe on the subjects. The probe was secured over the motor region. (C) Examples of short and long channel pairs. With this probe arrangement, the possible relative distances between the short and the long SD channels were 1.4, 1.7, 2.5, 3.3, 4.2, 5.2 and 6.2 cm.

each subject. During the resting blocks, a small 0.5-by-0.5 cm white square located at the middle of the screen appeared and the subjects were asked to fixate on this square. During the finger tapping blocks, the instruction "tap your fingers" was displayed in white characters on the computer screen using the Psychophysics toolbox in Matlab [11]. At that time, the subjects were asked to touch their right thumb with each of the fingers of their right hand alternately at a rate of 3 Hz. Following the three functional runs, three baseline runs of 5 minutes each were acquired. During the baseline runs, the subjects were asked to simply close their eyes and remain still.

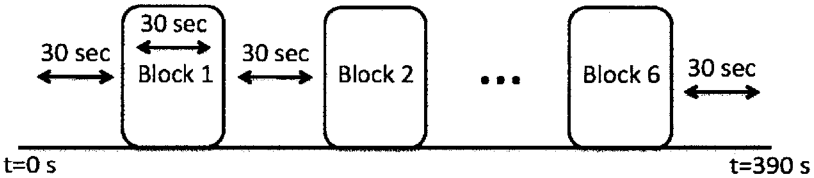

Figure 3-2: Overview of the finger tapping protocol. A run consisted of 6 blocks of 30 seconds of finger tapping interleaved with 30 seconds of rest. Each runs started and ended with a 30 second resting period. 3 functional runs were acquired for each of the 6 subjects.

3.2.2 Data processing

An overview of the procedure is shown in Fig. 3-3. Both the short and long SD measurements were bandpass filtered at 0.01-1.25 Hz. Even though the data will be further low pass filtered at 0.5 Hz in the processing stream, it is important to keep the 0.5-1.25 Hz frequency band here, since most of the cardiac oscillations are contained in this frequency band. These cardiac oscillations are strongly present in both the short and the long SD measurements and this increases the baseline correlation between the short and the long separation channel. These cardiac oscillations guide the dynamic estimation of the superficial contamination to more accurately estimate the HRF. As weve shown recently [44], prefiltering the cardiac oscillations reduces the performance of the dynamic estimation and results in a poorer estimate of the HRF. For both the simulations and the real functional data analysis, the Kalman filter algorithm was used to regress the short separation measurement and recover the hemodynamic response simultaneously. This algorithm was described in detail previously [44] and only the salient points are reviewed here.

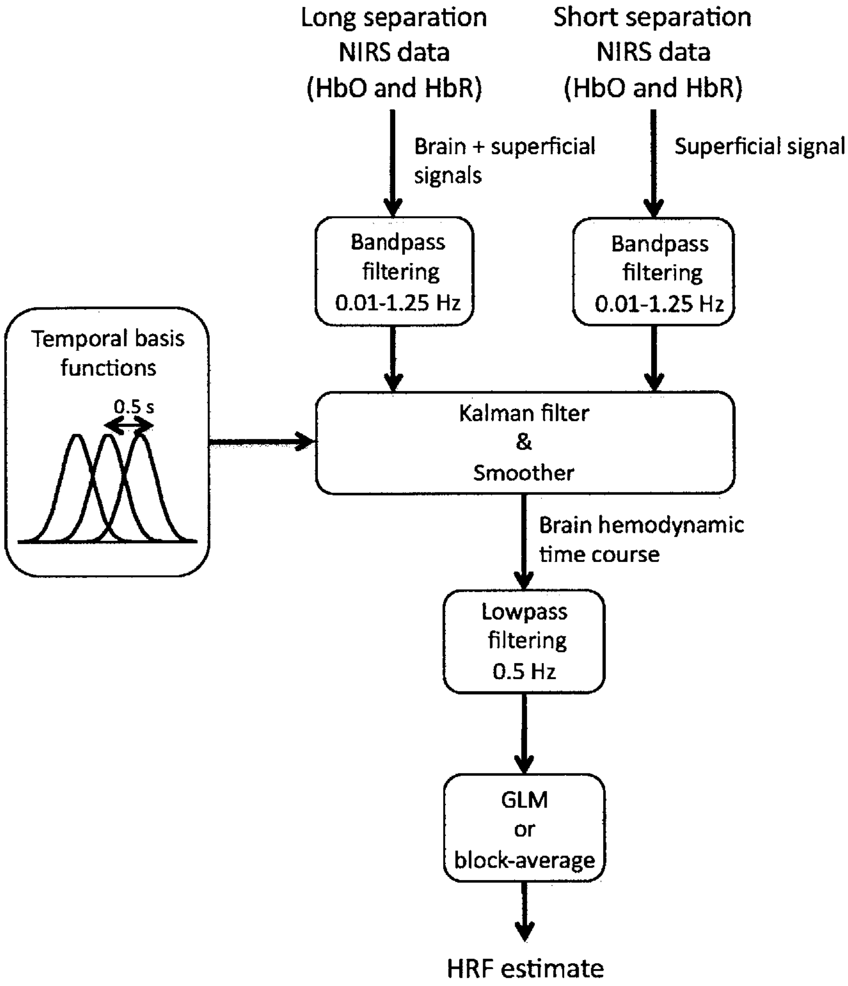

Figure 3-3: Schematic of the NIRS data analysis. The NIRS data from both the 1 cm and the 3 cm separation channels were first converted to HbO and HbR time courses and bandpass filtered. HbO and HbR were analyzed separately. The 1 cm and 3 cm bandpass time courses were passed to the Kalman algorithm [44] and then further lowpass filtered. The HRF was finally estimated using the GLM or a standard block-average.

The hemodynamic response was modeled by

$$h[n] = \sum_{i=1}^{N_w} w_i b_i [n] . \tag{3.1}$$

where $b_i[n]$ are normalized Gaussian functions with a standard deviation of 0.5 s and their means separated by 0.5 s. N_w is the number of Gaussian functions used to model the hemodynamic response and was set to 15 for our simulations (section 3.2.3) and 79 for our finger tapping data to recover the HRF over 0-8 sec and 0-40 sec respectively. The signal in the 3 cm separation channel $y_3[n]$ was modelled by a linear combination of the 1 cm separation signal $y_1[n]$ and the brain response $y_b[n]$. The expression for the 3 cm signal is given by

$$y_3[n] = y_b[n] + a\, y_1[n] \tag{3.2}$$

with

$$y_b[n] = \sum_{k=-\infty}^{\infty} h[k]\, u[n-k]. \tag{3.3}$$

and where $u[n]$ is the onset vector which is a binary vector taking the value 1 when n corresponds to a time when the stimulus was presented and 0 otherwise. It is to note that $u[n]$ is equal to 1 only at the onset of the stimulus and not throughout the duration of the stimulus.

The variable a is the dynamic weight used to model the superficial signal in the 3 cm separation channel as a linear combination of the short separation signal. Only a single time delay was taken from the short separation channel to model the superficial signal in the 3 cm channel since this has been shown to result in a better performance in our previous paper [44]. The states to be estimated by the Kalman filter were the weight of the superficial contribution a and the weights of the temporal bases w_i. All these weights were assumed to be time-varying. Eqs. (4.1), (4.2) and (4.3) can be re-written in state-space form:

$$\mathbf{x}[n+1] = \mathbf{I}\mathbf{x}[n] + \mathbf{w}[n] \tag{3.4}$$

$$y_3[n] = \mathbf{C}[n]\,\mathbf{x}[n] + \mathbf{v}[n] \tag{3.5}$$

where $\mathbf{w}[n]$ and $\mathbf{v}[n]$ are the process and the measurement noise respectively. $\mathbf{x}[n]$

is the n^{th} instance of $\mathbf{x}$ given by

$$\mathbf{x} = \left[\begin{array}{cccc} w_1 & \ldots & w_{N_w} & a \end{array} \right]^T . \tag{3.6}$$

The quantity $\mathbf{I}$ is an $N_w + 1$ by $N_w + 1$ identity matrix and $\mathbf{C}[n]$ is a 1 by $N_w + 1$ vector given by

$$\mathbf{C}[n] = \left[\begin{array}{cccc} u * b_1[n] & \cdots & u * b_{N_w}[n] & y_1[n] \end{array} \right] . \tag{3.7}$$

where "$*$" denotes the convolution operator. The estimate $\hat{\mathbf{x}}[n]$ at each sample n is then computed using the Kalman filter [77] followed by the Rauch–Tung–Striebel smoother [113].

The convergence of the Kalman filter depends on the initial estimate of the state vector $\hat{\mathbf{x}}[0]$. To overcome this problem, $\hat{\mathbf{x}}[0]$ was set to the values obtained using a static least-squares estimator as in [44] to ensure a fast convergence. Moreover, to overcome the problem of selecting a good initial guess for the state covariance estimate $\hat{\mathbf{P}}[0]$, the Kalman filter algorithm was run twice and the initial covariance estimate for the second run was set to the final covariance estimate of the first run. This process makes the performance of the filter almost insensitive to the initial covariance estimate. For the first pass of the Kalman filter, we set $\hat{\mathbf{P}}[0]$ to an identity matrix with diagonal entries of $1\text{x}10^{-1}$ for the temporal basis states and $5\text{x}10^{-4}$ for the superficial contribution state. The process noise covariance $\mathbf{Q}$ only contained nonzero terms on the diagonal elements. Those diagonal terms were set to $2.5\text{x}10^{-6}$ for the temporal basis states and $5\text{x}10^{-6}$ for the superficial contribution state. The measurement noise covariance $\mathbf{R}$ was set to an identity matrix scaled by $5\text{x}10^{-2}$. These values were extensively studied in our previous paper [44] and multiplying or dividing these values by factor of 10 did not significantly affect the performance of our method. The Kalman filter algorithm was then processed with the following prediction-correction recursion [46]:

$$\hat{\mathbf{x}}\left[n|n-1\right] = \hat{\mathbf{x}}\left[n-1|n-1\right] \tag{3.8}$$

$$\hat{\mathbf{P}}\left[n|n-1\right] = \hat{\mathbf{P}}\left[n-1|n-1\right] + \mathbf{Q}. \tag{3.9}$$

$$\mathbf{K}\left[n\right] = \hat{\mathbf{P}}\left[n|n-1\right]\mathbf{C}\left[n\right]^T\left(\mathbf{C}\left[n\right]\hat{\mathbf{P}}\left[n|n-1\right]\mathbf{C}\left[n\right]^T + \mathbf{R}\right)^{-1} \tag{3.10}$$

$$\hat{\mathbf{x}}\left[n|n\right] = \hat{\mathbf{x}}\left[n|n-1\right] + \mathbf{K}_n\left(y_3\left[n\right] - \mathbf{C}\left[n\right]\hat{\mathbf{x}}\left[n|n-1\right]\right) \tag{3.11}$$

$$\hat{\mathbf{P}}\left[n|n\right] = \left(\mathbf{I} - \mathbf{K}\left[n\right]\mathbf{C}\left[n\right]\right)\hat{\mathbf{P}}\left[n|n-1\right]. \tag{3.12}$$

After the Kalman algorithm was applied twice, the Rauch–Tung–Striebel smoother was applied in the backward direction [56]:

$$\hat{\mathbf{x}}\left[n|N_t\right] = \hat{\mathbf{x}}\left[n|n\right] + \hat{\mathbf{P}}\left[n|n\right]\hat{\mathbf{P}}\left[n+1|n\right]^{-1}\left(\hat{\mathbf{x}}\left[n+1|N_t\right] - \hat{\mathbf{x}}\left[n+1|n\right]\right) \tag{3.13}$$

with N_t the number of time points in the data. The complete time course of the filtered brain signal $\hat{y}_b\left[n\right]$ containing the estimated hemodynamic response $\hat{h}\left[n\right]$ was then reconstructed for each sample time n using the first N_w final state estimates $\hat{\mathbf{x}}_b = \left[w_1 \cdots w_{N_w}\right]^T$ and the temporal basis set contained in $\mathbf{C}\left[n\right]$

$$\hat{y}_b\left[n\right] = \mathbf{C}\left[n\right]\hat{\mathbf{x}}_b\left[n|N_t\right]. \tag{3.14}$$

This reconstructed filtered brain signal time course $\hat{y}_b\left[n\right]$ was further low pass filtered at 0.5 Hz to remove any cardiac fluctuations potentially present in the time course and the final estimate of the hemodynamic response $\hat{h}\left[n\right]$ was obtained either by applying a standard General Linear Model (GLM) procedure (without any cosine bases or short separation regressor) containing the same temporal basis function as in Eq. (4.1) or by block-averaging $\hat{y}_b\left[n\right]$. More details can be found in our previous paper [44].

3.2.3 Simulations

For each baseline measurement, the changes in optical density were converted to changes in hemoglobin concentrations using the modified Beer-Lambert relationship [19, 24, 7]. A pathlength correction factor of 6 and a partial volume correction factor of 50 were applied [69, 70]. The variance in all 252 (6 subjects x 3 runs x 14 pairs) baseline HbO and HbR time courses from the 3 cm measurements were then computed. To ensure a uniform distribution of the noise in our simulations, only the time courses showing a variance below 25 μM^2 were kept in the analysis, corresponding to 28.2 % of the data (71 of the 252 baseline time courses). We have also tested and confirmed that our method was working for higher levels of noise. Due to the non uniform level of noise across the probe, this threshold of 25 μM^2 was required in order to compare all distances on equal footing.

Ten individual evoked responses were added over all 71 selected 3 cm baseline measurements at random onset times with an inter-stimulus interval taken randomly from a uniform distribution (10-30 sec). This procedure was repeated 30 times for each baseline measurement to create 30 simulated time courses with 30 different onset times and ensure reproducible averaged results. The duration of the synthetic response was 8 seconds. The HbO time course increased by 15 μM at the peak while the HbR time course decreased by 7 μM. The synthetic hemodynamic response was the same used in our previous paper [44]. The resulting 2130 time courses (71 time courses x 30 simulated runs) were then bandpass filtered (0.01-1.25 Hz) and passed to the Kalman filter algorithm (Fig. 3-3) using each of the seven short separation (1 cm) measurements available as a regressor. The HRF was also recovered using a standard GLM with a set of cosine basis with 64 s period cutoff [42] for comparison (no short separation used). This resulted in 17,040 estimated HRFs (2130 time courses x 8 regressors (7 short separations + 1 standard GLM with cosine basis set)). The HbO and HbR responses were recovered independently. For each 1 cm-3 cm combination, the baseline R^2 coefficient before adding the synthetic HRF to the 3 cm channel was computed.

For each short separation used, the relative center-to-center distance between the 3 cm and the short-separation channel was computed. With the probe shown in Fig. 4-1a, the possible relative distances were 1.4, 1.7, 2.4, 3.3, 4.2, 5.2 or 6.2 cm as well as $>$ 10 cm for the forehead channel and are illustrated in Fig. 4-1c.

The quality of each recovered HRF was quantified by three different metrics: (1) the Pearson correlation coefficient R^2 between the true synthetic HRF (tHRF) and the recovered HRF (rHRF), (2) the mean square error (MSE) between tHRF and rHRF and (3) the Contrast-to-noise ratio (CNR) defined as the amplitude of rHRF divided by the root-mean-square (RMS) of the residual of tHRF and rHRF

$$\mathrm{CNR} = \frac{\max(\mathrm{rHRF})}{\mathrm{RMS}(\mathrm{rHRF} - \mathrm{tHRF})}. \tag{3.15}$$

The average for each of these three metrics across all the recovered HRFs for each specific relative distance was computed and the results were compared to the corresponding averaged metrics obtained from the HRFs recovered with the standard GLM (no short separation) using a two-tailed paired t-test. As in our previous paper [44], we used a paired t-test to resolve for small systematic differences. For the Pearson R^2 metric, the average was taken after applying a Fisher transformation and the resulting average was then back transformed. This comparison was performed for all 8 relative distances (1.4, 1.7, 2.4, 3.3, 4.2, 5.2, 6.2 cm and forehead).

This entire procedure was then repeated after introducing a time-lag in the short separation channel. For each 1 cm-3 cm combination, the cross-correlation function between the two channels before adding the synthetic HRF was also computed and a time-lag corresponding to the maximum of the cross-correlation function was applied to the short separation measurement. This time-lag could be any number in the interval $\{-N_t, N_t\}$ with N_t the number of time point in the NIRS time course but typical values obtained from our data ranged from -4 to 4 seconds for both HbO and HbR. The values for R^2, MSE and CNR obtained by introducing a time-lag were also compared with the zero-lag values with a two-tailed paired t-test.

The cross-correlation function used to identify the optimal time-lag was normalized such that the zero-lag value corresponded to the Pearson R^2 coefficient

$$\bar{R}_{y_1 y_3}[m] = \frac{R_{y_1 y_3}[m]}{\sigma_{y_1} \sigma_{y_3}} \tag{3.16}$$

with σ_{y_3} the standard deviation of $y_3[n]$. The maximum of this normalized cross-correlation function is the equivalent to shifting one of the channels by the optimal time-lag before computing the standard correlation and thus we will refer to this value as the optimal time-lag correlation for the rest of the text. To avoid any confusion, we will refer to the standard R^2 correlation as the zero-lag correlation. The zero-lag and the optimal time-lag correlations were also compared using a two-tailed paired t-test of their Fisher transformed values.

3.2.4 Functional data

The functional data were analysed in the same way as above with the Kalman filter, but the HRFs were recovered from 0 to 40 seconds after the stimulus onsets. Each 3 cm channel was analyzed using each of the seven short separation channels available and also with a standard block-average for comparison.

3.3 Results

3.3.1 Baseline correlation

The correlation (Pearson R^2) between the baseline NIRS time courses are shown in Fig. 3-4. In panels (A) and (B), the correlation between the 3 cm separation and the short separation channels are plotted as a function of their relative distance on the probe. These values are identified by the label "no time-lag" in the legends. The optimal time-lag correlation values are also plotted and identified by the label "optimum time-lag" in the legends. We observed a fast decay of both the zero-lag and the optimal time-lag correlations as the distance between the two channels went from 1.4 to 2.4 cm and the correlation then plateaued from 2.4 to 6.2 cm. This trend was observed for both HbO and HbR. The optimal time-lag correlation values obtained were slightly higher (but significant at the $p < 0.05$ level, two-tailed paired t-test) than the zero-lag correlation for all relative distances on the probe. It is good to re-emphasize that the statistical test is pairwise such that a small but constant difference across the sample is marked as significant. The optimal time-lags obtained increased with increasing relative distances indicating that this slight improvement in correlation was real and not due to potential artifact in the processing. Finally, the increases in correlation obtained by introducing time-lag were slightly more prominent for HbR than HbO.

In panels (C) and (D) of Fig. 3-4, both the zero-lag and the optimal time-lag correlations between two short separation channels are plotted as a function of their relative distance on the probe. A similar fast decay was observed as the relative distance between the two channels increased from 1 to 2 cm on the probe and then the correlation plateaued for longer distances. The values for the optimal-time lag correlations in this case were significantly higher ($p < 0.05$, two-tailed paired t-test) compared to the zero-lag correlations.

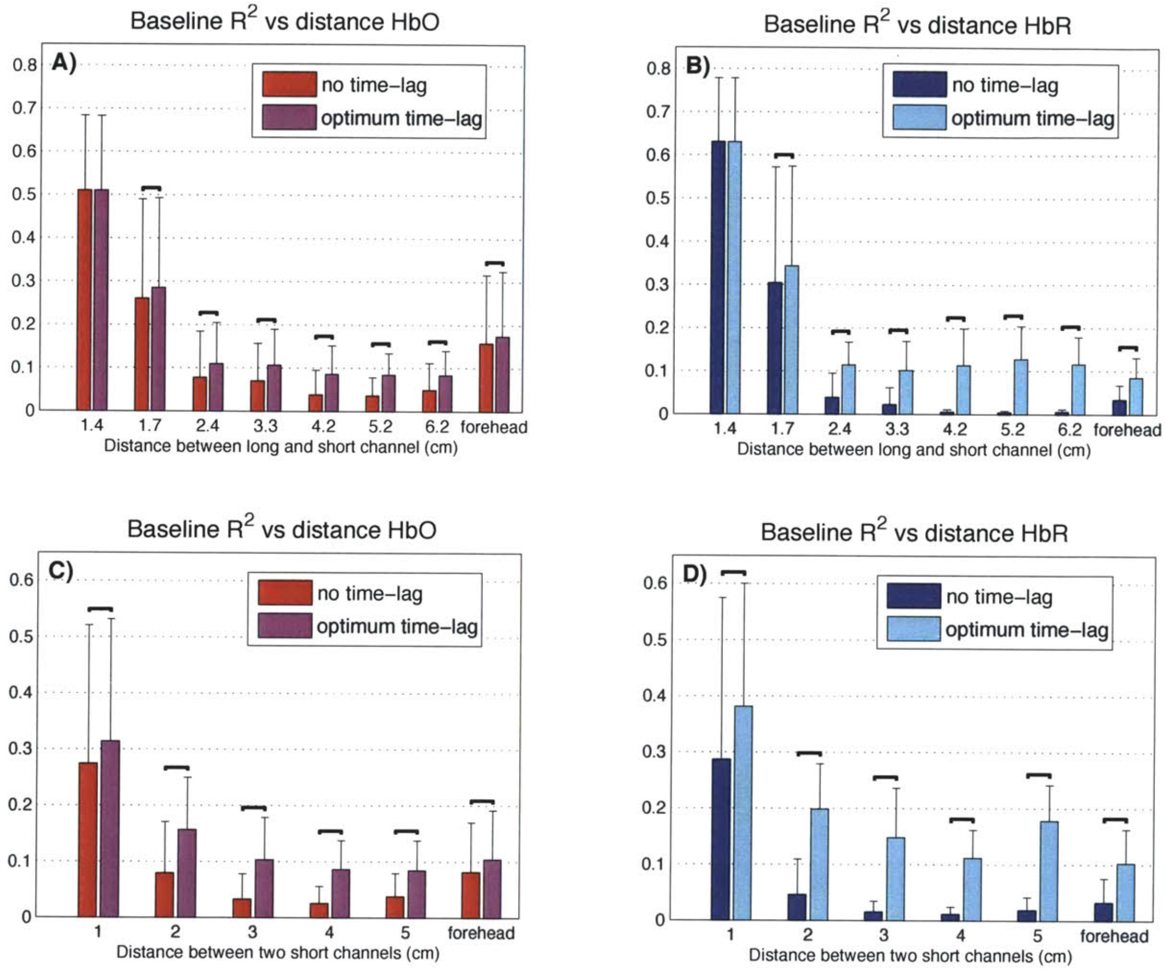

Figure 3-4: Effect of the relative distance on the initial baseline correlation between the channels. The baseline data were bandpass filtered between 0.01 and 1.25 Hz before the R^2 correlation was computed. The values labeled "optimal time-lag" were computed by taking the maximum of the normalized cross-correlation function (Eq. 3.16) while the "no time-lag" values are the standard Pearson R^2 coefficient. Statistical differences at the $p < 0.05$ level are indicated with horizontal black lines (two-tail paired t-test). The statistical test is pairwise such that a small but constant difference across the sample is marked as significant. (A)-(B) Initial baseline R^2 between the long and the short channels as a function of the relative distance between them. (C)-(D) Initial baseline R^2 between two short separation channels as a function of the relative distance between them.

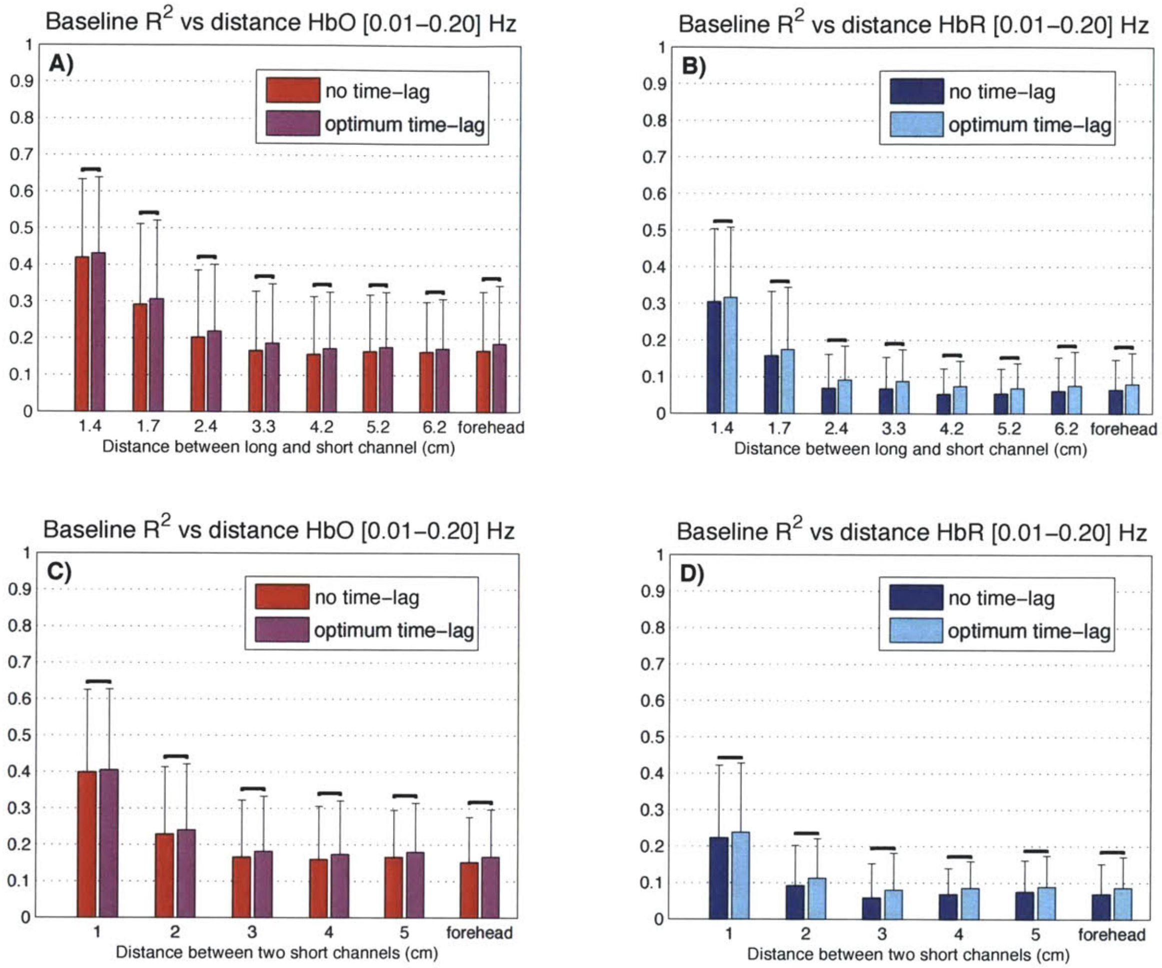

Figure 3-5: Low frequency band. Effect of the relative distance on the initial baseline correlation between the channels. The baseline data were bandpass filtered between 0.01 and 0.2 Hz (low frequency oscillations) before the R^2 correlation was computed. The values labeled "optimal time-lag" were computed by taking the maximum of the normalized cross-correlation function (Eq. 3.16) while the "no time-lag" values are the standard Pearson R^2 coefficient. Statistical differences at the $p < 0.05$ level are indicated with horizontal black lines (two-tail paired t-test). The statistical test is pairwise such that a small but constant difference across the sample is marked as significant. (A)-(B) Initial baseline R^2 between the long and the short channels as a function of the relative distance between them. (C)-(D) Initial baseline R^2 between two short separation channels as a function of the relative distance between them.

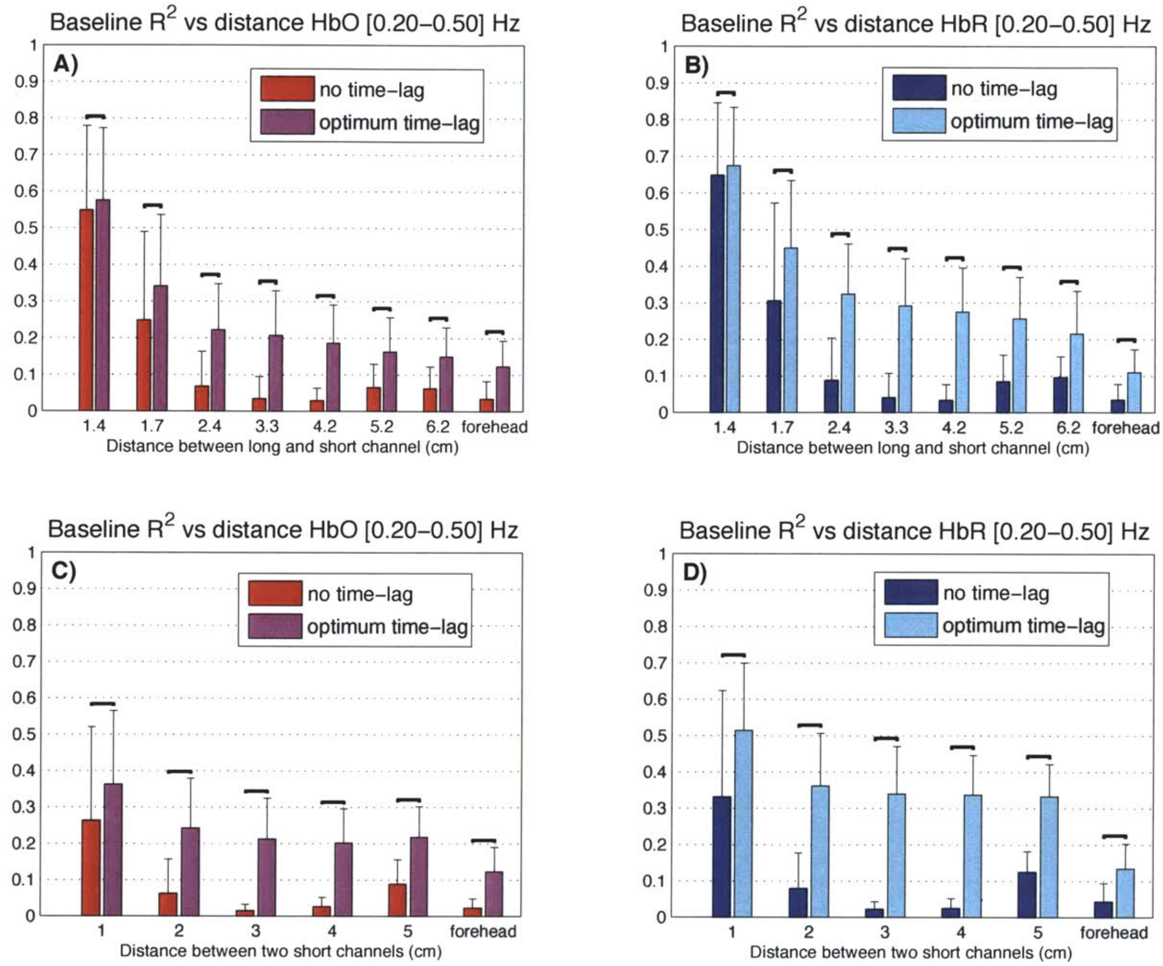

Figure 3-6: Respiratory band. Effect of the relative distance on the initial baseline correlation between the channels. The baseline data were bandpass filtered between 0.2 and 0.5 Hz (respiratory oscillations) before the R^2 correlation was computed. The values labeled "optimal time-lag" were computed by taking the maximum of the normalized cross-correlation function (Eq. 3.16) while the "no time-lag" values are the standard Pearson R^2 coefficient. Statistical differences at the $p < 0.05$ level are indicated with horizontal black lines (two-tail paired t-test). The statistical test is pairwise such that a small but constant difference across the sample is marked as significant. (A)-(B) Initial baseline R^2 between the long and the short channels as a function of the relative distance between them. (C)-(D) Initial baseline R^2 between two short separation channels as a function of the relative distance between them.

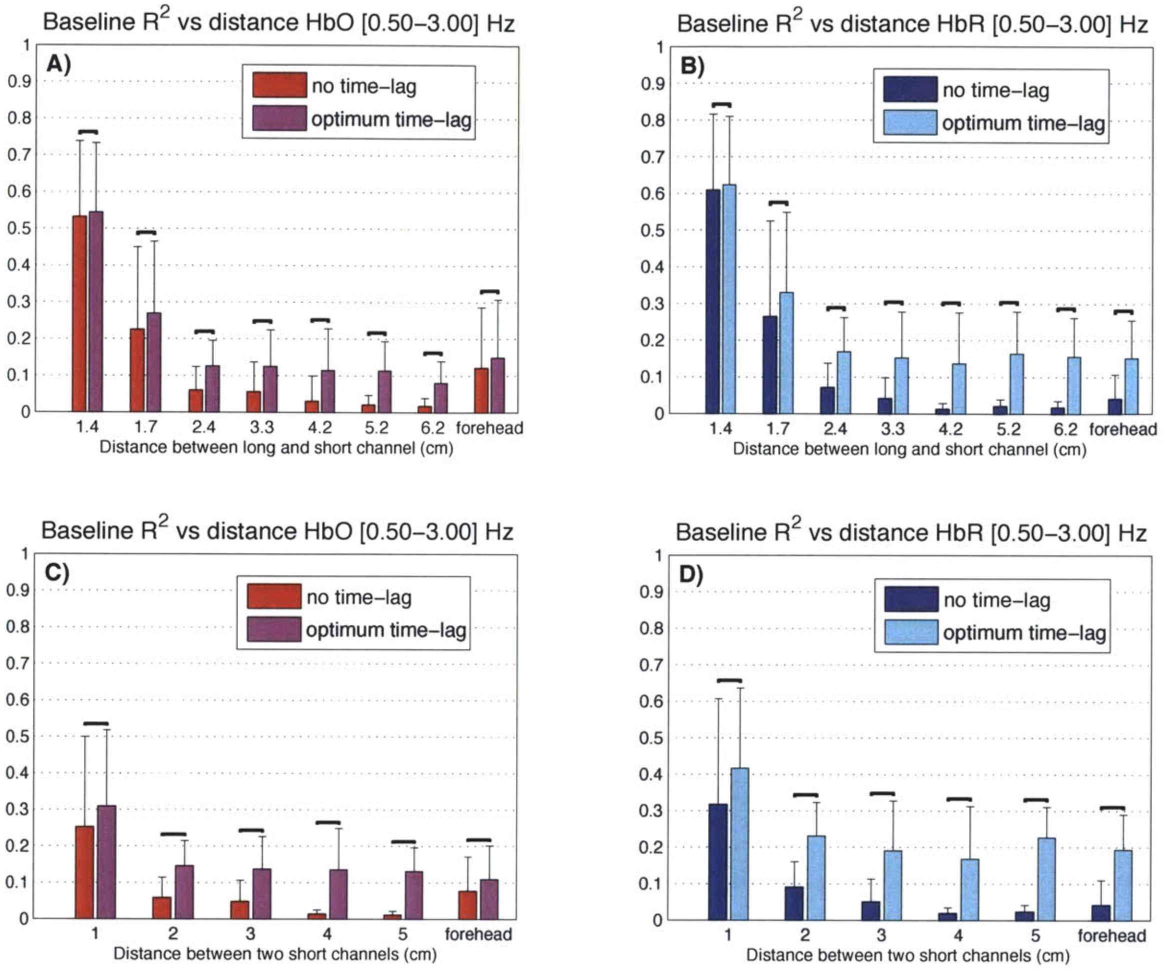

Figure 3-7: Cardiac band. Effect of the relative distance on the initial baseline correlation between the channels. The baseline data were bandpass filtered between 0.5 and 3 Hz (cardiac oscillations) before the R^2 correlation was computed. The values labeled "optimal time-lag" were computed by taking the maximum of the normalized cross-correlation function (Eq. 3.16) while the "no time-lag" values are the standard Pearson R^2 coefficient. Statistical differences at the $p < 0.05$ level are indicated with horizontal black lines (two-tail paired t-test). The statistical test is pairwise such that a small but constant difference across the sample is marked as significant.(A)-(B) Initial baseline R^2 between the long and the short channels as a function of the relative distance between them. (C)-(D) Initial baseline R^2 between two short separation channels as a function of the relative distance between them.

3.3.2 Simulation results

The results for the synthetic HRF simulations are shown in Figs. 3-8, 3-9 and 3-10 for the R^2, MSE and CNR metric respectively. On panels (A) and (B) of all three figures, the three metrics are plotted as a function of the relative distance between the 3 cm separation and the short separation channel used as a regressor. Values obtained by introducing a time-lag in the short separation channel are also shown as well as the corresponding values obtained using a standard GLM. We observed a fast decrease of the improvement obtained by the Kalman filter as the relative distance between the 3 cm and the short separation channel was increased from 1.4 to 2.4 cm. The performance then plateaued for longer relative distances. Both the R^2 (Fig. 3-8) and the CNR (Fig. 3-10) decreased as the relative distance between the long- and short separation channels was increased, while the MSE (Fig. 3-9) increased. Using a short separation channel located 1.4 cm away from the channel containing the synthetic HRF resulted in a mean increase in CNR of 50 % for HbO and 100 % for HbR relative to the GLM method. Using a short separation channel located farther than 2 cm away from the channel containing the synthetic HRF resulted in significant ($p < 0.05$, two-tailed paired t-test) but negligible improvements of the order of a few percent compared to the standard GLM procedure. Again, we re-emphasize that the statistical test is pairwise such that a small but constant difference across the sample is marked as significant. However the magnitude of the difference is small.

On panels (C) and (D) of Figs. 3-8, 3-9 and 3-10, the same R^2, MSE and CNR results are plotted as a function of the baseline zero-lag correlation (Pearson R^2) between the 3 cm and the short separation channels. Results obtained by introducing a time-lag in the short separation channel are also plotted as a function of the baseline optimal time-lag correlation. We observed a linear relationship between the improvement obtained with the Kalman filter and both the baseline zero-lag correlations and optimal time-lag correlations between the two channels. A baseline correlation greater than 0.8 resulted in a mean improvement in CNR of 50 % and 100 % compared to the standard

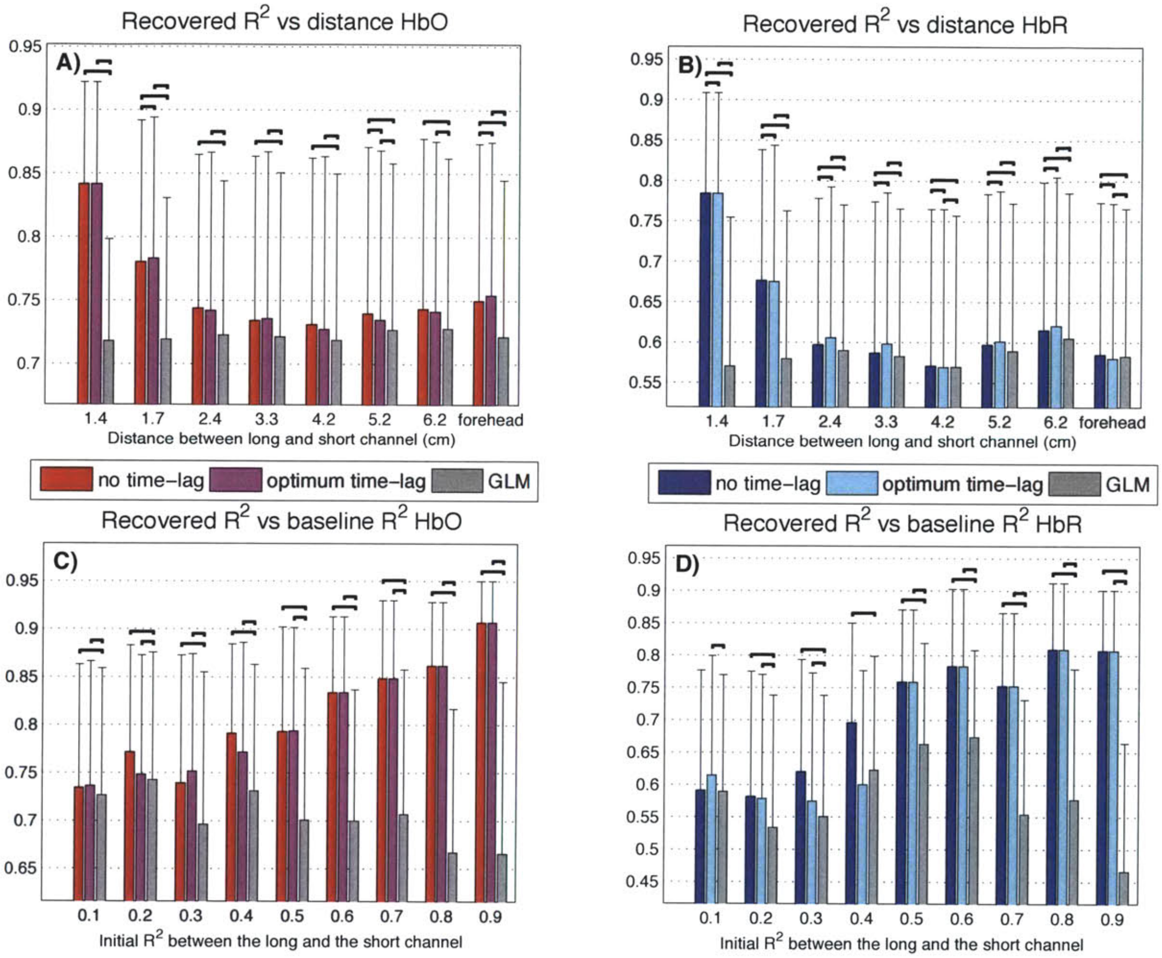

Figure 3-8: Effect of the relative distance on the correlation between the recovered HRF and the true HRF. Panels (A)-(B) show the recovered R^2 as a function of the distance between the short and the long NIRS channel. Panels (C)-(D) show the recovered R^2 as a function of the baseline R^2 between the short and the long NIRS channel. The value recovered using a standard GLM with no small separation is also shown for comparison. Statistical differences at the $p < 0.05$ level are indicated with horizontal black lines (two-tail paired t-test). The statistical test is pairwise such that a small but constant difference across the sample is marked as significant.

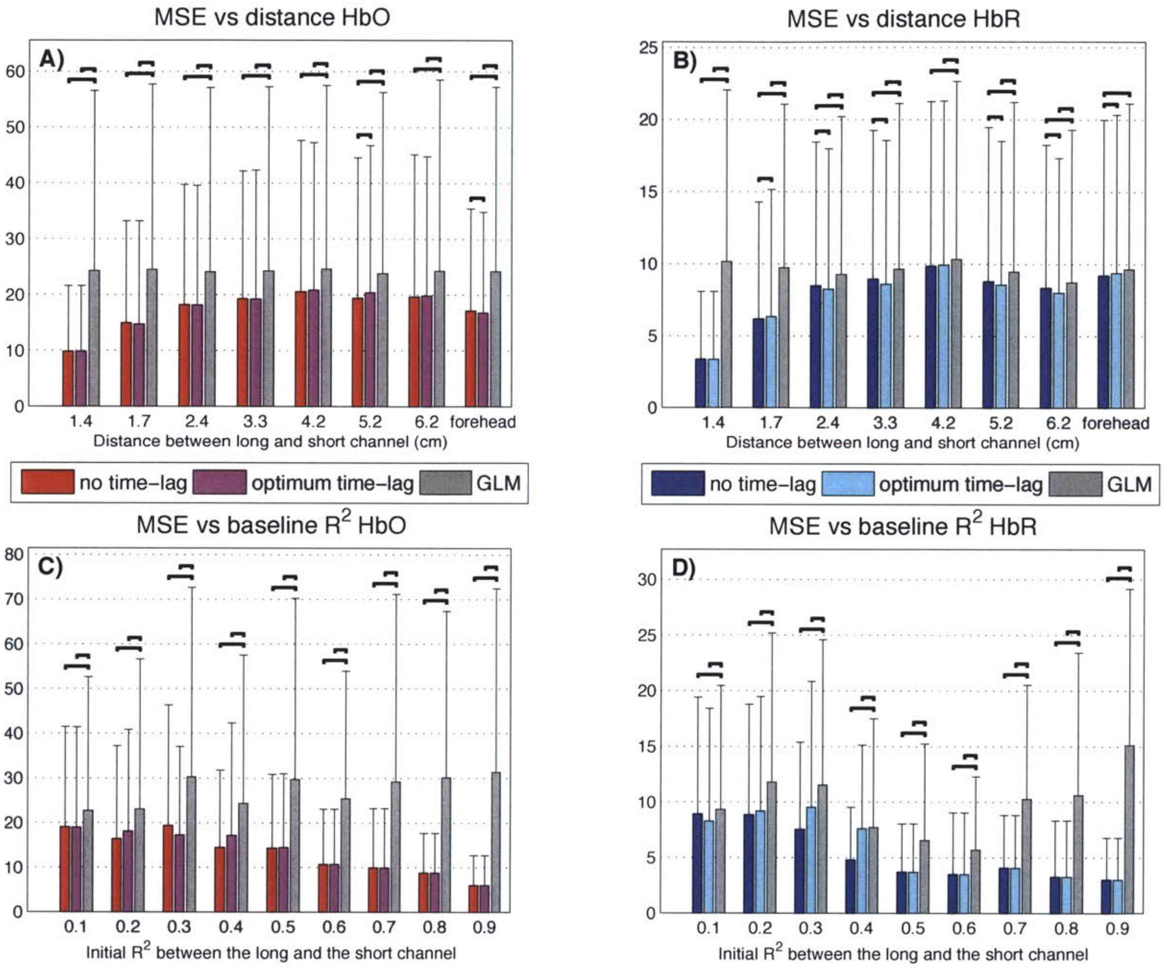

Figure 3-9: Effect of the relative distance on the MSE between the recovered HRF and the true HRF. Panels (A)-(B) show the MSE as a function of the distance between the short and the long NIRS channel. Panels (C)-(D) show the MSE as a function of the baseline R^2 between the short and the long NIRS channel. The value recovered using a standard GLM with no small separation is also shown for comparison. Statistical differences at the $p < 0.05$ level are indicated with horizontal black lines (two-tail paired t-test). The statistical test is pairwise such that a small but constant difference across the sample is marked as significant.

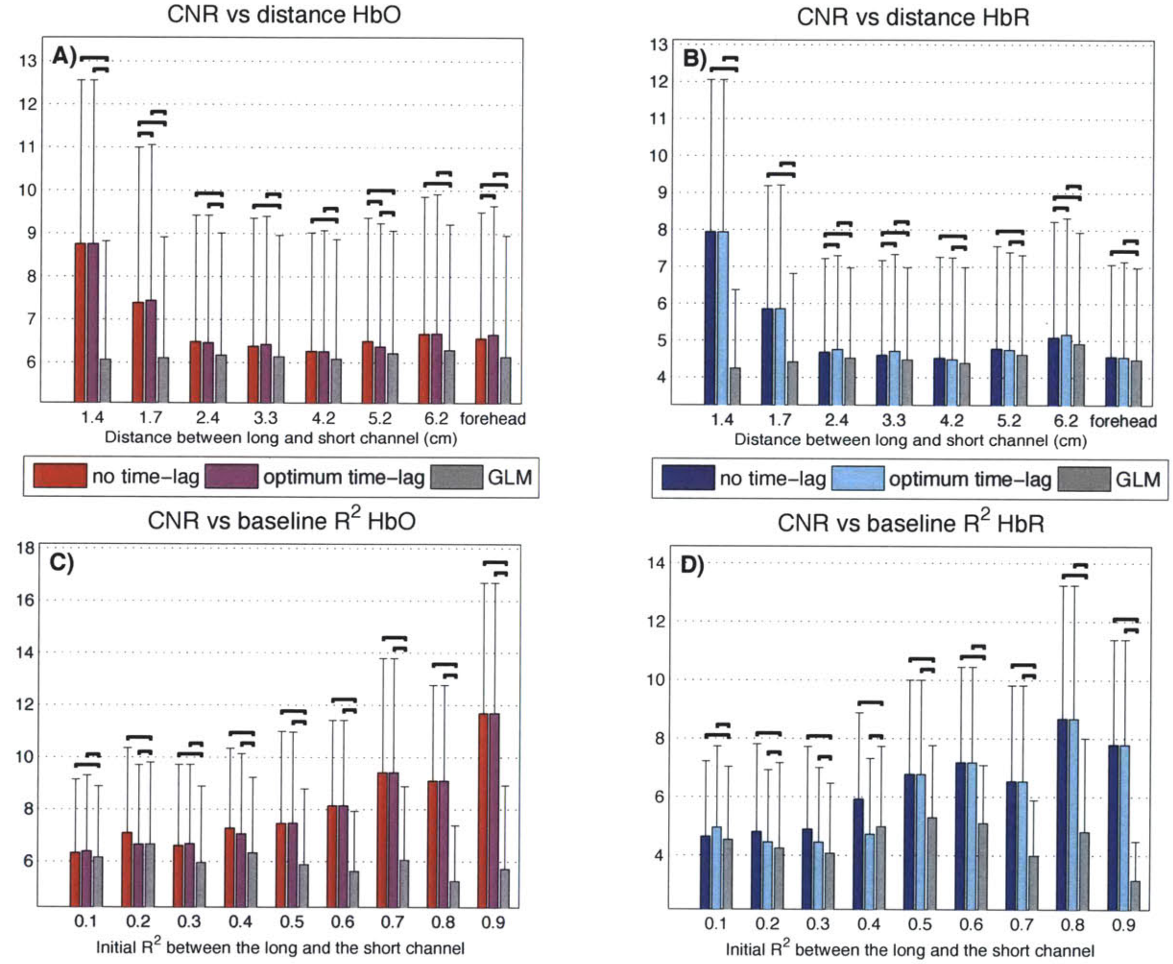

Figure 3-10: Effect of the relative distance on the Contrast-to-noise ratio (CNR) defined in Eq. 3.15. Panels (A)-(B) show the CNR as a function of the distance between the short and the long NIRS channel. Panels (C)-(D) show the CNR as a function of the baseline R^2 between the short and the long NIRS channel. The value recovered using a standard GLM with no small separation is also shown for comparison. Statistical differences at the $p < 0.05$ level are indicated with horizontal black lines (two-tail paired t-test). The statistical test is pairwise such that a small but constant difference across the sample is marked as significant.

GLM for HbO and HbR respectively.

3.3.3 Functional data results

Each run of finger tapping was analysed independently for each subject. The SD pair showing the strongest functional response was selected manually for each subject. To avoid any bias towards the Kalman filter method, both the Kalman filter and the standard block-average results were taken into account independently. The criteria for selecting the responses were a sustained increase in HbO and a sustained decrease in HbR [20], as well as a sustained increase in HbT to avoid pial vein washout contamination. Based on these criteria, the selected channel from the block-average matched the one selected from the Kalman filter result for each subject, although the HRF recovered with the block-average showed weak activation for three of the six subjects. For each individual subject, the HRFs recovered from the selected channel are shown for the first run in Fig. 3-11. Results from a single run are presented to illustrate the power of our method and the high CNR achieved with only 6 individual finger tapping blocks. Results from the second and the third run were very similar. Columns 1-4 illustrate the corresponding HRFs (same 3 cm channels) recovered using: (1) a block-average with no small-separation channel, (2) the Kalman filter with the small-separation channel located in the forehead as a regressor, (3) the Kalman filter with the short separation channel located 2.4 cm away as a regressor and (4) the Kalman filter with the closest short separation channel (located 1.4 cm away) as a regressor.

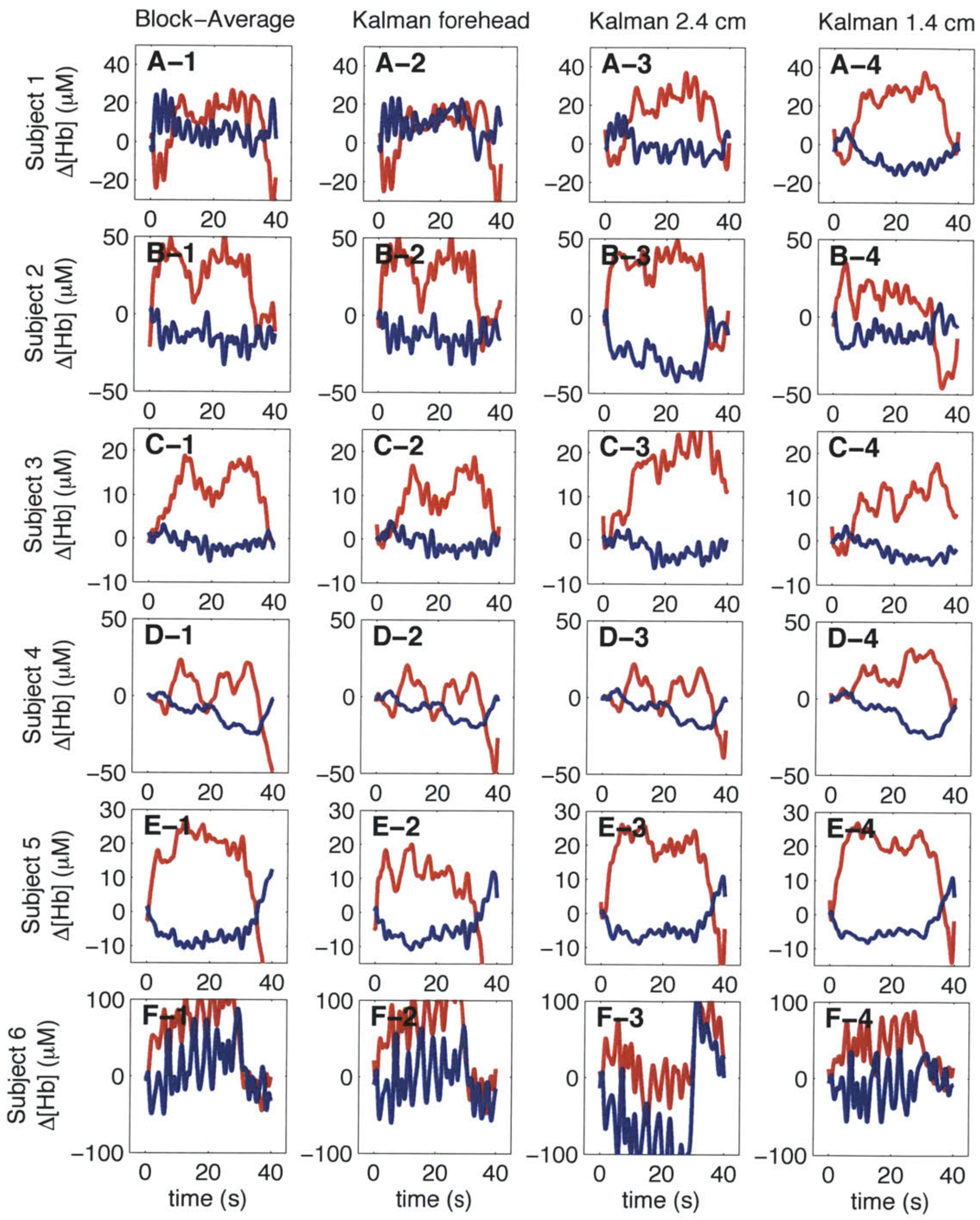

Figure 3-11: Finger tapping results from a single run containing 6 individual blocks. For each subject, the trace from the channel containing the strongest functional response is shown. (1) Standard block-average (2) Kalman filter with the forehead short separation channel. (3) Kalman filter with the short separation channel located at 2.4 cm. (4) Kalman filter with the short separation channel located at 1.4 cm.

For subjects 1 ,4 and 6, the HRFs recovered without using the closest short separation channel (columns 1-3) showed weak activation patterns only. However, the activation became clear on the same channels using the Kalman filter together with the closest short separation channel (column 4). For subjects 2, 3 and 5, a clear activation was present whether the Kalman filter was used or not. For subjects 2, 3 and 4, a strong artifact present between 15 and 30 s. was removed using the Kalman filter with the closest short separation channel. Removing this artifact made the HRFs more constant during the duration of the stimulus (0-30 s).

3.4 Discussion

3.4.1 Systemic interference measured by NIRS is inhomogeneous across the scalp

Systemic interference measured in NIRS has been termed "global" interference previously in the literature [116, 147, 148, 136, 118]. In contradiction, our present results indicate that systemic interference is actually inhomogeneous across the surface of the scalp, that is, the correlation between systemic interference measured at two different locations decreases with the increasing relative distance between the two measurements. Although the short separation channel measurements might contain some cortical signal, Monte Carlo simulations have shown that this contribution is negligible for a SD distance of 1 cm [147]. As such, panels (C) and (D) of Fig. 3-4 indicate that the origin of this decorrelation is located in the superficial layers of the head and therefore, is not due to autoregulation mechanisms occurring in the brain tissue [105].

Although introducing a time-lag re-established one third of the correlation between the two short separation measurements, the other two thirds of the R^2 correlation was still lost after introducing time delays. This finding indicates that only part of the correlation decay can be explained by simple transit time effects across different locations on the scalp. The systemic interferences measured in NIRS are oscillatory processes containing three dominant components [84]. These are the cardiac pulsations around 1 Hz, respiratory oscillations around 0.4 Hz and other low frequency oscillations (including Mayer waves [76]) around 0.1 Hz. Analyses similar to the one presented in Fig. 3-4 were performed with the NIRS data bandpass filtered at 0.01-0.2 Hz, 0.2-0.5 Hz and at 0.5-3 Hz (see supplementary figures 3-5, 3-6 and 3-7). These frequency bands correspond to the low frequency, respiratory and cardiac oscillations respectively. These analyses revealed a decay in correlation with increasing relative distances in all these three frequency bands. Although the correlation decayed, it never reached zero even for low frequency oscillations, which is in agreement with recent findings by Tong and Frederick ([132]). Up to 3/4 and 1/2 of the correlation lost in the 0.2-0.5 Hz band and the 0.5-3 Hz band respectiveley could be re-established by introducing a time-lag. However, introducing a time-lag in the 0.01-0.2 Hz frequency band resulted in only negligible improvements in correlation. These findings are in agreement with a recent paper from Tian et al ([130]) indicating that cardiac fluctuations (~1 Hz) are more global while low frequency oscillations (~0.1 Hz) are less spatially coherent.

From our results, one can conclude that (1) slow oscillations are inhomogeneous across the surface of the scalp and (2) a significant proportion of the correlation decay in the higher frequency bands (cardiac and respiration) is attributed to transit time effects across different spatial regions. These phase mismatches of the cardiac and respiratory pulsation over different locations arise potentially from spatial heterogeneity of the vasculature such as blood vessel length, orientation, size, depth and dilation [147, 146]. However, 1/4 and 1/2 of the correlation lost in the respiration and cardiac frequency band respectively could not be re-established by introducing a time-lag and future studies will be required to investigate the origin of this correlation decay.

3.4.2 Impact on the short separation method

This fast decrease in correlation with increasing relative distance has an important impact on the performance of the short separation method. Panels (C) and (D) of Figs. 3-8, 3-9 and 3-10 illustrate that all three metrics used (R^2, MSE and CNR) varied linearly with the baseline correlation, whether a time-lag was used or not. Since the baseline correlation decreased with the relative distance as shown in Fig. 3-4, we expected the performance of the short separation method assessed with these three metrics to decrease with the relative distance. This expectation was confirmed by our simulations and are shown in panels (A) and (B) of Figs. 3-8, 3-9 and 3-10. For relative distances larger than 2 cm, only mild improvements were obtained using our short separation approach compared to the standard GLM method. No decrease in performance was observed using our Kalman filter with any of the available short separation. This is consistent with our previous findings [44] that the Kalman filter improves or doesn't change recovery of the hemodynamic response. At worst, the recovered response will be the same as the one recovered with a standard GLM with no small separation used. To obtain larger improvements that are useful in practice, the short separation channel must be located no more than 1.5 cm from the 3 cm channel from which the HRF is to be recovered.

For large NIRS probes containing several long SD measurements spread over several centimeters, our results indicate that multiple short separation channels are required in order to combine each long-separation measurement with a short separation channel within a 1.5 cm radius. In this case, our simulations indicate that the improvement in CNR using the short separation method is of the order of 50 % for HbO and 100 % for HbR, as shown respectively in panels (A) and (B) of Fig. 3-10. As in our previous paper [44], we observed an improvement for both HbO and HbR, in contradiction with Zhang et al. [146] where no improvement was observed for HbR using an adaptive filter method. In the simulations of our previous paper [44], we also observed a decrease in performance for HbR using an adaptive filter. As we showed, our Kalman filter approach overcomes this pitfall by regressing the short separation measurement and simultaneously estimating the hemodynamic response.

The necessity of using several short separation channels on the probe was also confirmed by our finger tapping experiment. Using the short separation channel located at 2.4 cm from the 3 cm SD measurement, the HRF obtained from subjects 1, 4 and 6 showed a weak activation pattern only, as shown in Fig. 3-11. On the other hand, the activation became very clear on these same channels using the short separation channel located 1.4 cm away. The baseline correlations between the short separation and the long-separation channels were around 0.3 for the short separation channel located 2.4 cm away and around 0.5 for the short separation channel located 1.4 cm away. This re-emphasizes the fact that the initial baseline correlation between the short separation and the 3 cm channel is an important factor in determining the performance of the Kalman filter algorithm. In practice, this baseline correlation can be computed to predict the impact of using short optode separations. In our previous paper [44], we showed that the presence of a hemodynamic response in the 3 cm channel does not impact the baseline initial correlation between the short separation and the 3 cm channel. This occurs because the contribution of systemic interference in NIRS largely dominates the contribution of the hemodynamic response. This was confirmed here with our real functional data.

3.4.3 How many regressors should be used?

In all the simulations presented in this paper, we used a single regressor at a time. We also ran simulations using multiple regressors simultaneously but all of them resulted in a significantly higher MSE and lower CNR compared to using the single short separation channel. This is explained by a potential overfitting of the data. When more than one regressor is used at the same time, we start fitting noise in the data which introduces errors in the estimation of the hemodynamic response. The same thing occurred in our previous paper [44] when a single regressor was used but multiple time delays from this regressor were used in the regression. To avoid overfitting the data and to obtain the maximum power from the short separation method, we have found that on average a single regressor performs better than two or more.

3.4.4 Alternatives for modelling the physiological interference

Alternative methods for modelling the physiological interference have been proposed in the literature. [111] used a set of sine and cosine functions to model the oscillatory behaviour of the systemic physiology. The linear coefficients of these temporal bases were included as additional states in the state-space model. Alternatively, [1] used a set of sine functions only but included the phase as an additional state. These methods were implemented and compared with the short separation method. In both cases, the short separation approach performed better compared to these modelling techniques. These models contain a higher number of degrees of freedom (i.e. larger number of state) which potentially introduces crosstalk between the state corresponding to the hemodynamic response and the state corresponding to the systemic physiology. This phenomenon degrades the estimation of the hemodynamic response. The short separation method reduces the number of degrees of freedom and reduces crosstalk by measuring directly the systemic interference in the superficial layers of the head and therefore, results in a more accurate estimation of the hemodynamic response.

3.4.5 Future studies

In this work, a small detector fiber was placed in proximity of each source fiber, resulting in a single small separation channel for every long SD channel. An interesting question is whether or not an additional short separation channel located in proximity of the detector fiber would further improve the recovery of the hemodynamic response. Doing so would maximize the overlap between the pathlength of the small separation channels and the longer SD channel, and might result in a further improvement of the small separation method. However, care must be taken to ensure that the additional source fibers do not inadvertently saturate the detectors and to ensure that the fiber optic probe remains flexible for efficient coupling with the scalp. We have begun to investigate the improvement of additional short separation measurements on the long separation source and long separation detector simultaneously. Our preliminary evidence indicates that as expected the results get better. For sure, we can still use a single short separation regressor but choose the one which has a higher correlation with the long separation measurement; either the short separation coincident with the long separation source, or the long separation detector. When we use both short separation regressors we have to ensure that we are not over fitting the data. Sometimes using both short separation regressors will actually degrade our estimate of the HRF, and thus statistical model testing needs to be implemented to determine if one or two regressors should be used.

3.5 Summary

In this study, we have determined that the position of the short separation NIRS channel relative to the long-separation channel impacts the performance of the short separation regression method to improve the recovery of the hemodynamic response in NIRS. We showed that the relative distance between the channel of interest and the regressor must be less than 1.5 cm to have a meaningful impact on the recovery of the hemodynamic response. In this case, improvements in CNR were of the order of 50 % for HbO and 100 % for HbR compared to the standard GLM approach. When a short separation channel located farther than 2 cm was used as regressor, only minor improvements were obtained compared to the standard GLM method, which are of little practical use. This decrease in performance for longer relative distances results from a decrease in the baseline correlation between the channel of interest and the regressor. Our results indicate that this correlation decay is due to spatially inhomogeneous hemodynamics in the superficial layers of the head.

Chapter 4

Double short separation measurements

This section was published in:
Gagnon, L., Yucel, M. A., Boas, D. A. and Cooper, R. J. (2013). "Further improvement in reducing superficial contamination in NIRS using double short separation measurements." *NeuroImage* in press

The main contribution of this chapter is to demonstrate that using two short separation measurements, one at the source optode and one at the detector optode, further increases the performance of the short separation regression method compared to using a single short separation measurement. Our work emphasizes the importance of integrating short separation measurements both at the source and at the detector optode of the standard channels from which the hemodynamic response is to be recovered. While the implementation of short separation sources presents some difficulties experimentally, the improvement in noise reduction is significant enough to justify the practical challenges.

4.1 Introduction

Near-Infrared Spectroscopy (NIRS) is a non-invasive neuro-investigatory technique used to measure the hemodynamic changes associated with evoked brain activity [99, 63, 88, 36]. In adult populations, the majority of functional NIRS studies are performed using a back-reflection geometry, with near-infrared light carried to and from the head via optical fibers. This back-reflection geometry ensures that the NIRS signal is extremely vulnerable to contamination by the hemodynamics within the superficial layers of the head.

It has recently been shown that this interference can be significantly reduced with the use of short separation (SS) recordings which are sensitive to superficial layers only, including the scalp and the skull [116, 147, 136]. These SS measurements are used as regressors during the post-processing of the NIRS signal and have been shown to improve the detection of evoked brain activity using NIRS [146, 118, 141, 44].

In our previous work [43], we showed that the superfical NIRS signal obtained by SS channels is spatially inhomogeneous across the scalp. In order to successfully regress the superficial signal from that of the brain, the SS measurement must therefore be located as close as possible to the long separation (LS) NIRS channel. For NIRS measurements performed on adult humans, the source and detector optodes are generally separated by around 3 cm, which is large enough to observe the hemodynamics of the cortex but short enough that enough light returns to the detector [7]. Any change measured by a standard NIRS channel is the result of an integration of all optical changes which occur throughout the volume of tissue traversed by the NIR light. The NIRS signal will therefore invariably contain superficial interference from two different locations: beneath the source optode and beneath the detector optode and these interference signals may not correlate with one another. As a result, it is reasonable to hypothesize that the performance of the SS regression method will improve if the superficial NIRS signal is independently recorded from both the tissue beneath the source optode and from the tissue beneath the detector optode.

In this work, we investigated the performance of using two such SS measurements to regress systemic physiological signals from NIRS data.

4.2 Methods

4.2.1 Experimental measurements

Five healthy adult subjects were recruited for this study. The Massachusetts General Hospital Institutional Review Board approved the study and all subjects gave written informed consent. Data were collected using a TechEn CW6 system operating at 690 and 830 nm. The NIRS probe con- tained 8 sources and 8 detectors as shown in Fig. 4-1. The probe was designed to contain a SS measurement at each source and at each detector optode of the LS measurement to provide maximal overlap between the LS and the SS measurements. In order to avoid detector saturation, 200 μm-core fibers were used for the SS detectors (shown in red in Fig. 4-1) and an optical filter (Kodak WRATTEN ND 2.00) was glued to the tip of standard NIRS fiber bundles (shown in green in Fig. 4-1) for the additional sources of the SS measurements. The probe was secured over the left motor region of each subject.

During the experiment, subjects were sitting in a comfortable chair in front of a computer screen with a black background. The 6-minute long functional runs consisted of stimulus trials of 5-seconds of a finger tapping task, with an inter-trial interval varying randomly between 12 and 18 sec. Each functional run contained between 19 and 21 trials and three functional runs were acquired for each subject. Between stimulus periods, a small 0.5-by-0.5 cm white square appeared at the center of the computer screen and the subjects were asked to fixate on this square. During the finger tapping periods, the instruction tap your fingers was displayed in white characters on the computer screen using the Psychophysics toolbox in Matlab [11]. The finger tapping task required the subject to touch their right thumb to each finger of their

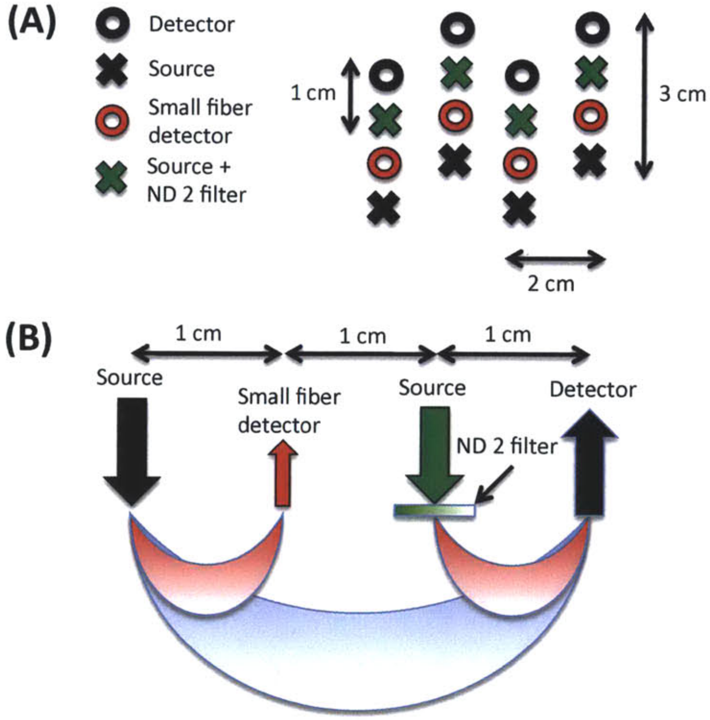

Figure 4-1: NIRS probe. (A) Geometry (B) Sensitivity In order to avoid detector saturation, 200 μm-core fibers were used for the SS detectors (shown in red) and a piece of optical filter (Kodak WRATTEN ND 2.00) was glued at the tips of standard NIRS fibers (shown in green) for the additional sources of the SS measurements.

right hand in order as quickly as possible and repeat the process until the instruction left the computer screen. Following the three functional runs, three baseline runs of 5 minutes each were acquired. During the baseline runs, the subjects were asked to simply close their eyes and remain still.

4.2.2 Data analysis

An overview of the data analysis procedure is shown in Fig. 4-2. Both the SS and LS measurements were bandpass filtered at 0.01-1.25 Hz and then used simultaneously in a Kalman filter. The Kalman filter was identical to that described in our previous papers [44, 43] with the addition of a second short separation regressor.

The hemodynamic response was modeled by

$$h\left[n\right]=\sum_{i=1}^{N_w} w_i b_i\left[n\right] \tag{4.1}$$

where $b_i\left[n\right]$ are normalized Gaussian functions with a standard deviation of 0.5 s and their means separated by 0.5 s. N_w is the number of Gaussian functions used to model the hemodynamic response and was set to 15 for our simulations to recover the HRF over 0-8 sec. The signal in the LS channel $y_{LS}\left[n\right]$ was modelled by a linear combination of the two SS signals $y_{SS}^{Src}\left[n\right]$ and $y_{SS}^{Det}\left[n\right]$ and the brain response $y_b\left[n\right]$. The expression for the LS signal is given by

$$y_{LS}\left[n\right]=y_b\left[n\right]+a_{Src}\, y_{SS}^{Src}\left[n\right]+a_{Det}\, y_{SS}^{Det}\left[n\right] \tag{4.2}$$

with

$$y_b\left[n\right]=\sum_{k=-\infty}^{\infty} h\left[k\right]u\left[n-k\right] \tag{4.3}$$

and where $u\left[n\right]$ is the onset vector which is a binary vector taking the value 1 when n corresponds to a time when the stimulus was presented and 0 otherwise. It is to note that $u\left[n\right]$ is equal to 1 only at the onset of the stimulus and not throughout the duration of the stimulus.

The variables a_{Src} and a_{Det} are the dynamic weights used to model the superficial signals in the LS separation channel as a linear combination of the two SS signals. Only a single time delay was taken from the short separation channels to model the superficial signals in the LS channel since this has been shown to result in a better performance in our previous paper [44]. The states to be estimated by the Kalman filter were the weights of the superficial contribution a_{Src} and a_{Det} and the weights of the temporal bases w_i. All these weights were assumed to be time-varying. Eqs. (4.1), (4.2) and (4.3) can be re-written in state-space form:

$$\mathbf{x}\left[n+1\right]=\mathbf{I}\mathbf{x}\left[n\right]+\mathbf{w}\left[n\right] \tag{4.4}$$

$$y_{LS}[n] = \mathbf{C}[n]\,\mathbf{x}[n] + \mathbf{v}[n] \tag{4.5}$$

where $\mathbf{w}[n]$ and $\mathbf{v}[n]$ are the process and the measurement noise respectively. $\mathbf{x}[n]$ is the n^{th} instance of $\mathbf{x}$ given by

$$\mathbf{x} = \begin{bmatrix} w_1 & \ldots & w_{N_w} & a_{Src} & a_{Det} \end{bmatrix}^T. \tag{4.6}$$

The quantity $\mathbf{I}$ is an $N_w + 2$ by $N_w + 2$ identity matrix and $\mathbf{C}[n]$ is a 1 by $N_w + 2$ vector given by

$$\mathbf{C}[n] = \begin{bmatrix} u * b_1[n] & \cdots & u * b_{N_w}[n] & y_{SS}^{Src}[n] & y_{SS}^{Det}[n] \end{bmatrix}. \tag{4.7}$$

where "$*$" denotes the convolution operator. The estimate $\hat{\mathbf{x}}[n]$ at each sample n is then computed using the Kalman filter [77] followed by the Rauch–Tung–Striebel smoother [113].

The convergence of the Kalman filter depends on the initial estimate of the state vector $\hat{\mathbf{x}}[0]$. To overcome this problem, $\hat{\mathbf{x}}[0]$ was set to the values obtained using a static least-squares estimator as in [44] to ensure a fast convergence. Moreover, to overcome the problem of selecting a good initial guess for the state covariance estimate $\hat{\mathbf{P}}[0]$, the Kalman filter algorithm was run twice and the initial covariance estimate for the second run was set to the final covariance estimate of the first run. This process makes the performance of the filter almost insensitive to the initial covariance estimate. For the first pass of the Kalman filter, we set $\hat{\mathbf{P}}[0]$ to an identity matrix with diagonal entries of $1\text{x}10^{-1}$ for the temporal basis states and $5\text{x}10^{-4}$ for the superficial contribution state. The process noise covariance $\mathbf{Q}$ only contained nonzero terms on the diagonal elements. Those diagonal terms were set to $2.5\text{x}10^{-6}$ for the temporal basis states and $5\text{x}10^{-6}$ for the superficial contribution state. The measurement noise covariance $\mathbf{R}$ was set to an identity matrix scaled by $5\text{x}10^{-2}$. These values were extensively studied in our previous paper [44] and multiplying or dividing these values by factor of 10 did not significantly affect the performance of our method. The Kalman filter algorithm was then processed with the following prediction-correction recursion [46]:

$$\hat{\mathbf{x}}\left[n|n-1\right] = \hat{\mathbf{x}}\left[n-1|n-1\right] \tag{4.8}$$

$$\hat{\mathbf{P}}\left[n|n-1\right] = \hat{\mathbf{P}}\left[n-1|n-1\right] + \mathbf{Q}. \tag{4.9}$$

$$\mathbf{K}\left[n\right] = \hat{\mathbf{P}}\left[n|n-1\right]\mathbf{C}\left[n\right]^{T}\left(\mathbf{C}\left[n\right]\hat{\mathbf{P}}\left[n|n-1\right]\mathbf{C}\left[n\right]^{T} + \mathbf{R}\right)^{-1} \tag{4.10}$$

$$\hat{\mathbf{x}}\left[n|n\right] = \hat{\mathbf{x}}\left[n|n-1\right] + \mathbf{K}\left[n\right]\left(y_{LS}\left[n\right] - \mathbf{C}\left[n\right]\hat{\mathbf{x}}\left[n|n-1\right]\right) \tag{4.11}$$

$$\hat{\mathbf{P}}\left[n|n\right] = \left(\mathbf{I} - \mathbf{K}\left[n\right]\mathbf{C}\left[n\right]\right)\hat{\mathbf{P}}\left[n|n-1\right]. \tag{4.12}$$

After the Kalman algorithm was applied twice, the Rauch–Tung–Striebel smoother was applied in the backward direction [56]:

$$\hat{\mathbf{x}}\left[n|N_t\right] = \hat{\mathbf{x}}\left[n|n\right] + \hat{\mathbf{P}}\left[n|n\right]\hat{\mathbf{P}}\left[n+1|n\right]^{-1}\left(\hat{\mathbf{x}}\left[n+1|N_t\right] - \hat{\mathbf{x}}\left[n+1|n\right]\right) \tag{4.13}$$

with N_t the number of time points in the data. The complete time course of the filtered brain signal $\hat{y}_b\left[n\right]$ containing the estimated hemodynamic response $\hat{h}\left[n\right]$ was then reconstructed for each sample time n using the first N_w final state estimates $\hat{\mathbf{x}}_b = \left[w_1 \cdots w_{N_w}\right]^T$ and the temporal basis set contained in $\mathbf{C}\left[n\right]$

$$\hat{y}_b\left[n\right] = \mathbf{C}\left[n\right]\hat{\mathbf{x}}_b\left[n|N_t\right]. \tag{4.14}$$

This reconstructed filtered brain signal time course $\hat{y}_b\left[n\right]$ was further low pass filtered at 0.5 Hz to remove any cardiac fluctuations potentially present in the time course and the final estimate of the hemodynamic response $\hat{h}\left[n\right]$ was obtained by applying a standard General Linear Model (GLM) procedure (without any cosine bases or short separation regressor) containing the same temporal basis function as in Eq. (4.1). More details of the algorithm and the parameters used can be found in our previous papers [44, 43].

4.2.3 Simulations

Simulations were performed as described previously [44, 43] by adding a synthetic HRF to the three baseline NIRS recordings acquired in each subject. For each baseline measurement, the changes in optical density were converted to changes in hemoglobin concentrations using the modified Beer-Lambert relationship [24]. A pathlength correction factor of 6 and a partial volume correction factor of 50 were applied [126, 69, 70].

An overview of the simulation procedure is shown in Fig. 4-3. Thirty individual simulated evoked responses were added over each of the 60 LS channel baseline measurements (5 subjects x 3 runs x 4 LS channels), with randomized onset times such that the inter-stimulus intervals varied between 12 and 30 seconds.

There is increasing evidence that many functional activation tasks give rise to systemic changes in oxyhemoglobin concentration which are time locked to the stimulus [79, 129]. These evoked systemic signals are likely due to an increase in heart rate and blood pressure associated with the performance of a task. We therefore added an evoked systemic artifact that was phase-locked with the stimulus to better represent real functional experiments. The artifact was added to both the LS and SS HbO signals. We used a sinusoidal function to model the artifact and the phase was chosen such that the peak of the artifact matches the peak of the hemodynamic response. No artifacts were added to the HbR traces since no evoked systemic artifacts have been reported for HbR signals [79, 129]. Baseline signals for SS and LS optodes are also shown at the bottom panels of Fig. 4-3. The green box emphasizes a portion where the Source SS signal resembles the LS signal while the yellow box emphasizes a portion where the Detector SS signal resembles the LS signal. These boxes illustrate the main hypothesis behind the double short separation method that the superficial interferences beneath the source and the detector may be different and that using two short separation measurements allows one to better capture the interference.

This procedure was repeated 30 times for each baseline measurement to create 30 simulated time courses with 30 different onset times to ensure robust averaged results. The duration of the synthetic response was 8 seconds as shown in Fig. 4-3. The resulting 1800 time courses (60 time courses x 30 simulated runs) were then band-pass filtered (0.01-1.25 Hz) and passed to the Kalman filter algorithm as depicted in Fig. 4-2. This algorithm was employed 3 times, using three different regressor cases: the SS measurement located close to the detector (Det SS), the SS measurement located close to the source (Src SS) and both SS measurements simultaneously (Src&Det). To provide a comparison where no SS measurements are used, the HRF was also recovered using a standard GLM approach, using a set of cosine bases with a 64 s period cutoff [42] to model the baseline physiology and the same temporal basis set used in the Kalman filter to model the HRF. Although simultaneous HbO and HbR estimation is possible [28], the HbO and HbR responses were recovered independently in this work to prevent potential crosstalk introduced by noise in the regression [126]. For each SS-LS combination, the baseline Pearson R^2 correlation coefficient after adding the synthetic HRF to the LS channel was computed.

The quality of each recovered HRF was quantified using four different metrics: (1) the Pearson correlation coefficient R^2 between the true synthetic HRF and the recovered HRF, (2) the mean square error (MSE) between the true HRF and the recovered HRF, (3) the baseline noise defined by taking the standard deviation over the intervals -5 to 0 sec prior to the stimulus onset and (4) the inter-trial variability obtained by computing the standard deviation across individual trials and taking the mean of the obtained variability over the interval 0 to 12 sec following the stimulus onset.

The average for each of these four metrics across all the recovered HRFs for each specific regressor was computed and the results obtained were compared using two-tailed paired t-tests. As in our previous papers [44, 43], we used a paired t-test to resolve for small systematic differences. For the Pearson R^2 metric, the average was taken after applying a Fisher transformation and the resulting average was then back transformed.

4.2.4 Experimental finger tapping

The functional data acquired during the finger tapping task were analyzed in the same way as above using the Kalman filter algorithm, but the HRFs were recovered from -5 to 18 sec after the stimulus onsets. Each of the four LS channels was analyzed using the three regressor cases derived from its closest SS measurements (Det SS, Src SS and Src&Det). The noise and the inter-trial variability metrics described above were also computed. The R^2 and the MSE metrics were not computed for the experimental finger tapping data since the true HRF is unknown.

4.3 Results

4.3.1 Baseline R^2 correlation results

The baseline R^2 coefficients between the SS and LS channel were com- puted and are shown in Table 4.1. The correlations computed for the simulations as well as for the experimental finger tapping are shown. For the simulations, the correlation was computed after the synthetic HRFs were added to the baseline LS trace and the evoked systemic artifacts were added to both the LS and SS traces. These correlations were calculated after all NIRS signals were bandpass filtered between 0.01 and 1.25 Hz to remove instrumental noise and drifts. As shown in Table 4.1, the baseline correlation was higher between the LS signal and the source SS signal than between the LS signal and the detector SS signal.

Table 4.1: Baseline R^2 correlation between LS and SS signals.

Optode	Sim HbO	Exp HbO	Sim HbR	Exp HbR
Src SS	0.39	0.38	0.24	0.13
Det SS	0.28	0.29	0.08	0.04

4.3.2 Simulation results

The metrics described in section 4.2.3 were computed over the 1800 recovered HRFs for each of the three regressor cases (Det SS, Src SS and Src&Det). To address how often the SS method improves the recovery of the HRF, we computed a likelihood ratio (as a percentage) which describes how often the SS method improved the recovery of the HRF in the 1800 traces of our simulations. The comparison was based on the performance obtained with the standard GLM and the likelihood was computed individually for each regressor case and using each of the four metrics. These likelihoods are shown in Table 4.2. An improvement was identified as an increase for R^2, and as a decrease for MSE, Noise and inter-trial variability. Likelihoods were computed by taking the ratio of the number of recovered HRFs which meet these criteria and dividing by 1800, the total number of recovered HRFs, and converting to a percentage.

Table 4.2 illustrates that the likelihood of improvement is higher for HbO compared to HbR. This pattern is consistent regardless of the metric used to assess the performance of the regression. Using two SS measurements resulted in an improved HRF 96% of the time for HbO and 53% of the time for HbR (based on the MSE metric).

The recovered HRF averaged across all subjects, all channels, all trials and all repetitions are shown in Fig. 4-4 for each of the four recovery methods (3 regressors cases + GLM). The width of the traces indicates the standard deviation taken across all simulations (5 subjects, 3 runs, 4 channels, 30 trials, 30 repetitions). The true simulated HRF is illustrated by dotted lines.

Table 4.2: Likelihood of improving the recovery of the HRF with the SS method compared to the standard GLM method. The likelihoods were computed by comparing the HRFs recovered over the 1800 simulated time courses of our simulations.

	Det SS (%)	Src SS (%)	Src&Det (%)
HbO			
R^2	77	83	85
MSE	89	94	96
Noise	88	92	95
Variability	95	88	95
HbR			
R^2	47	50	51
MSE	49	52	53
Noise	52	54	55
Variability	47	50	48

The evoked systemic artifact is still present in the HbO response recovered with the GLM and this resulted in an overestimation of the HbO response when no SS regressor was applied. The removal of the evoked systemic artifact was improved progressively as the detector SS, the source SS and both the source and the detector SSs were applied

A quantitative comparison of the different regressor cases is shown in Fig. 4-5. The values of the four performance metrics described in section 2.3 were compared for the three different regressor cases as well as for the standard GLM without SS regression. A significant improvement in performance was observed for all metrics ($p<0.05$, two-tailed paired t-test) when two SSs were used compared to using a single SS measurement.

4.3.3 Experimental finger tapping results

Fig. 4-6 shows the HRFs recovered during the finger tapping task for each of the four recovery methods. Results are shown for a single run and for a single subject to illustrate the high quality of the recovery and the traces corresponding to the 18 individual trials are shown to illustrate the variability. The mean HRFs are also illustrated by a black dotted line in each case.

Quantitative comparisons of the noise and inter-trial variability were computed over the 15 runs (5 subjects, 3 runs each). Results are shown in the bar graphs of Fig. 4-7. A decrease in both noise and inter-trial variability ($p<0.05$, two-tailed paired t-test) was observed when two SS signals were used compared to a single SS. This trend was observed for both HbO and HbR.

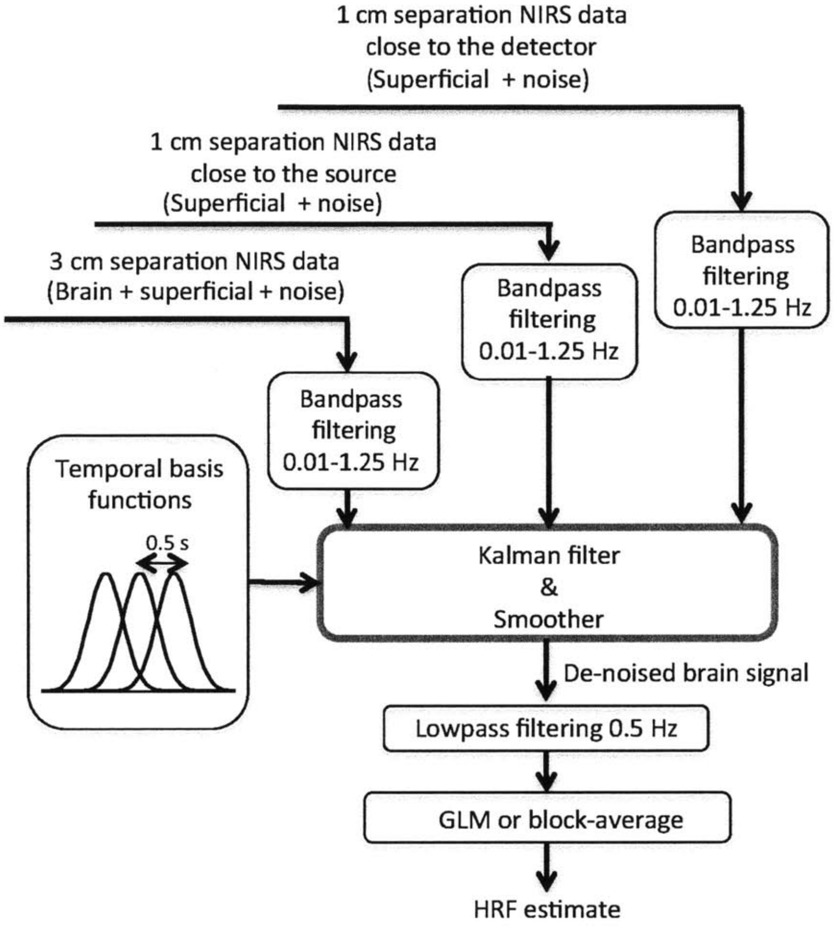

Figure 4-2: Overview of the algorithm for data analysis. Both the SS and LS measurements were bandpass filtered at 0.01-1.25 Hz and then used simultaneoulsy as regressors in the Kalman filter. A set of temporal basis functions was used to lower the dimensionality of the problem. The output of the Kalman filter was further low pass filtered at 0.5 Hz to remove any cardiac fluctuations potentially present in the time course and the final estimate of the hemodynamic response was obtained by applying a standard General Linear Model (GLM) procedure.

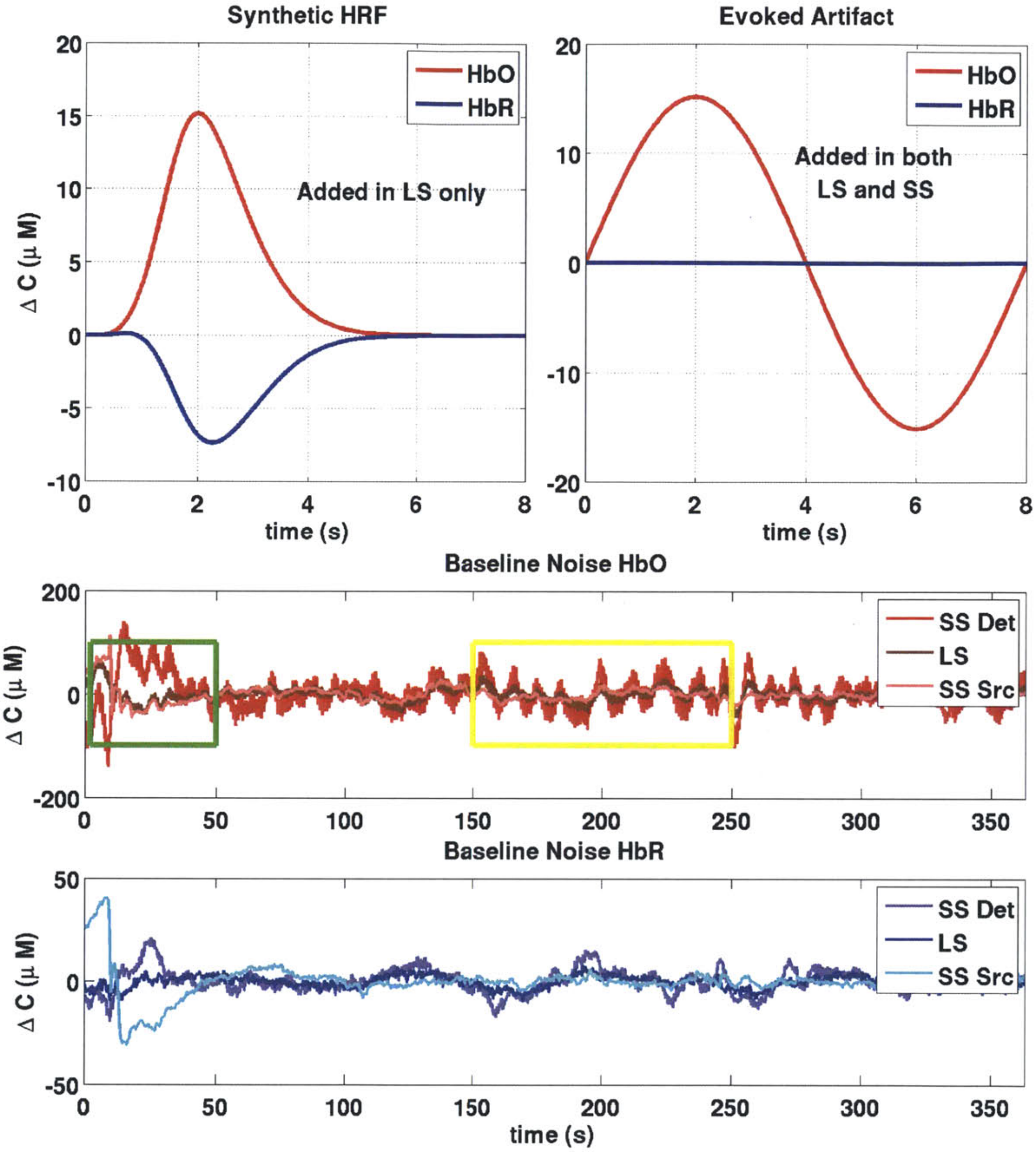

Figure 4-3: Construction of synthetic data. Thirty individual evoked responses were added over all 60 LS baseline measurements (5 subjects x 3 runs x 4 LS channels) at random onset times with an inter-stimulus interval taken randomly from a uniform distribution (12-30 sec). We also added an evoked systemic artifact to the HbO signals (LS and SS) that was phase-locked with the stimulus onset. The green box emphasizes a portion of the baseline signal where the Source SS signal resembles the LS signal while the yellow box emphasizes a portion where the Detector SS signal resembles the LS signal.

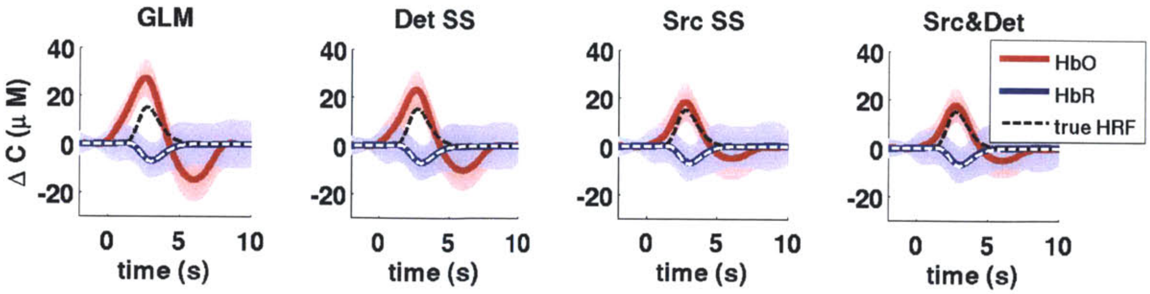

Figure 4-4: Averaged recovered HRF across all simulations. The width of the traces represent uncertainty given by one standard deviation taken across all simulations (5 subjects, 3 runs, 4 channels, 30 trials, 30 noise instances). The true simulated HRF is shown with dotted lines in each case.

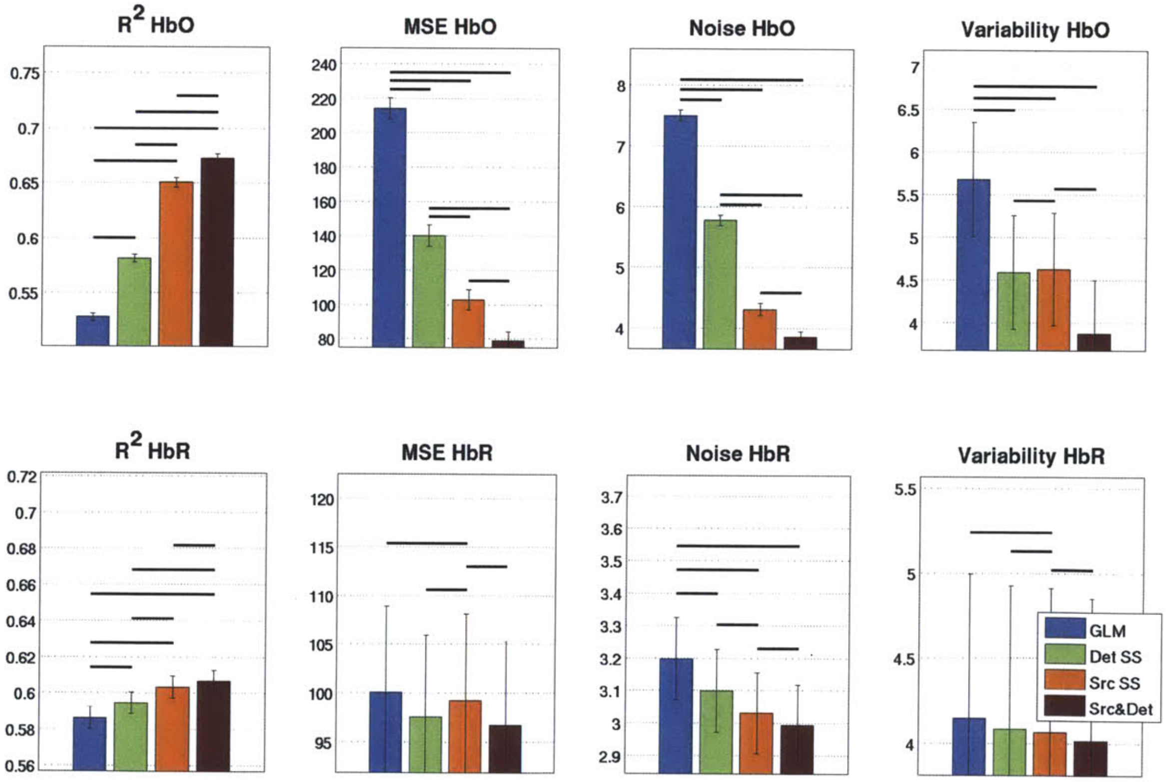

Figure 4-5: Quantitative analysis of the 1800 simulations. Bars represent the mean taken across all simulated HRFs and error bars represent the standard error on the mean. Statistical significance at the $p<0.05$ level (two-tailed paired t-test) is illustrated by black lines over the bars.

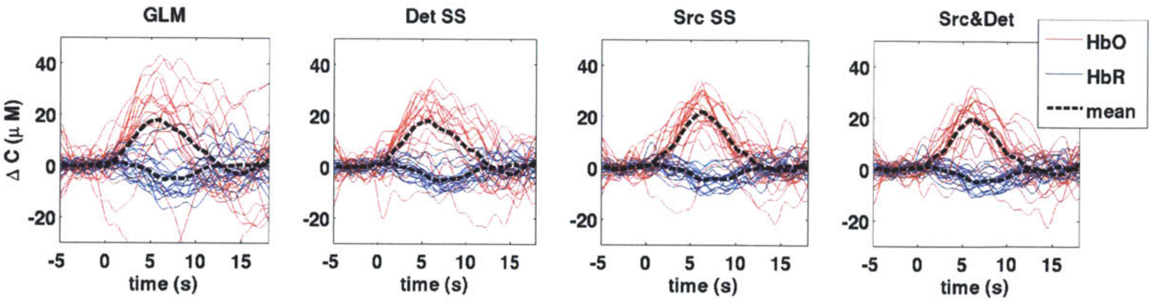

Figure 4-6: HRFs recovered during the experimental finger tapping for a representative subject for a single run. The HRFs for each of the 18 individual trials are shown in red for HbO and blue for HbR. The mean HRF taken across the 18 trials is shown by a black dotted line.

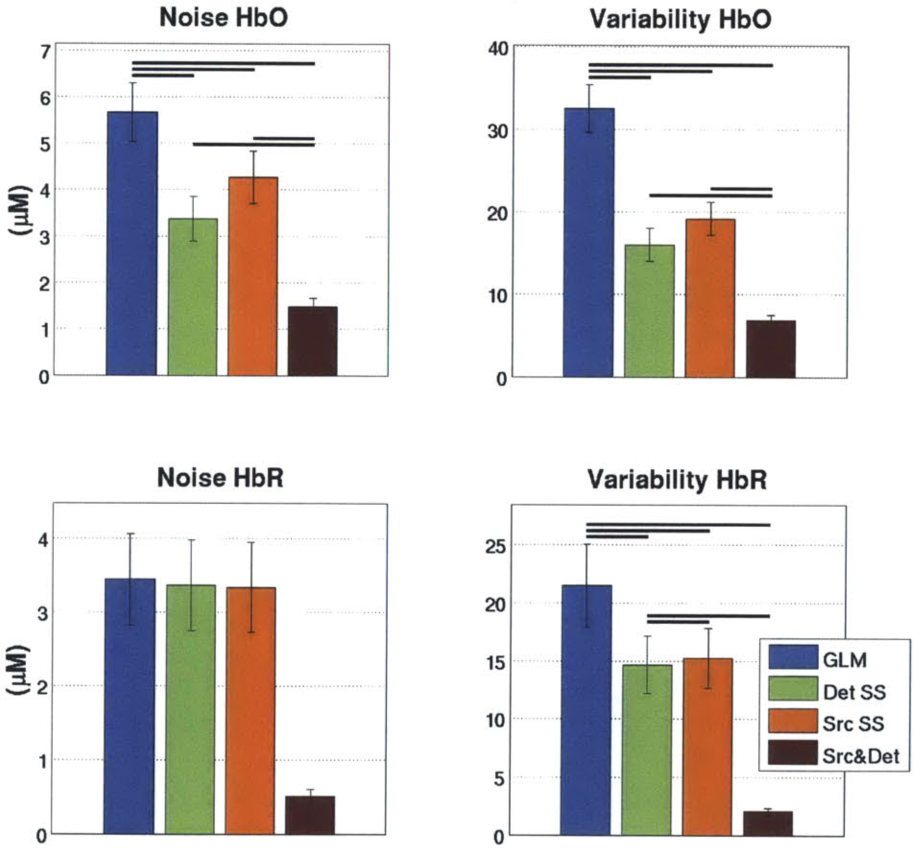

Figure 4-7: Quantitative analysis of the experimental finger tapping. Bars represent the mean taken across all runs and error bars represent the standard error on the mean. Statistical significance at the $p<0.05$ level (two-tailed paired t-test) is illustrated by black lines over the bars.

4.3.4 Combined results

To illustrate quantitatively the gain obtained by using two SS measurements, the noise levels and the inter-trial variability in the recovered HRFs were compared explicitly for the four different recovery methods across both the simulations and the experimental finger tapping results. Values are shown in Table 4.3.

Across all simulation and experimental results, the noise level for HbO decreased on average from 6.6 μM when using the GLM to 4.6 and 4.3 μM when using the detector SS and the source SS respectively. This corresponds to a decrease of 31 and 35% respectively. When both SS signals are used, the noise further drops to 2.7 μM, a decrease of 59% when compared to the standard GLM. For HbR, the noise level dropped on average from 3.3 to 3.2 μM using the detector SS or the source SS. This corresponds to a decrease of 2.6%. When the two SSs are used, the noise further drops to 1.8 μM, a decease of 47% when compared to the standard GLM procedure.

The inter-trial variability for HbO went from 19.6 μM when using the GLM to 10.3 and 11.9 μM when using the detector SS and the source SS respectively. This corresponds to a decrease of 46 and 38% respectively. When both SS signals are used, the variability further drops to 5.4 μM, a decrease of 72% when compared to the GLM. For HbR, the variability went from 12.8 μM to 9.4 and 9.7 μM using the detector SS and the source SS respectively. This corresponds to a decrease of 27 and 25% respectively. When the two SSs are used, the inter-trial variability further drops to 3.0 μM, a decrease of 76% when compared to the GLM.

Table 4.3: Comparison of the level of noise in the recovered HRF for different regressors. Combined (simulations + experimental finger tapping) results are shown.

Method	Noise (μM)	Comparison with GLM (%)	Comparison with Det SS (%)	Comparison with Src SS (%)
Noise HbO				
GLM	6.6			
Det SS	4.6	-30.5		
Src SS	4.3	-34.7		
Src&Det	2.7	-59.3	-41.5	-37.7
Noise HbR				
GLM	3.3			
Det SS	3.2	-2.6		
Src SS	3.2	-2.6		
Src&Det	1.8	-47.1	-45.7	-44.9
Variability HbO				
GLM	19.1			
Det SS	10.3	-46.0		
Src SS	11.9	-37.6		
Src&Det	5.4	-71.9	-47.9	-54.9
Variability HbR				
GLM	12.8			
Det SS	9.4	-26.9		
Src SS	9.7	-24.7		
Src&Det	3.0	-76.4	-67.6	-68.6

4.4 Discussion

4.4.1 Using two SS measurements is better than using only one

Our results show that using two SS measurements, one located close the source optode of the LS channel and one located close to the detector optode of the LS channel results in better performance of the SS method compared to when a single SS regressor is used. From our results in Table 4.3 (simulation and experimental), further improvements from 33% (Det: 30.5%, Src: 34.7%) to 59% in noise reduction and from 42% (Det: 46.0%, Src: 37.6%) to 72% in inter-trial variability were achieved for HbO. These values improved from 3% (Det: 2.6%, Src: 2.6%) to 47% for noise reduction and from 26% (Det: 26.9%, Src: 24.7%) to 76% for inter-trial variability for HbR. These improvements strongly motivate the incorporation of short separation measurements both at the source and at the detector optode of the LS channel.

The fact that two SSs perform better than a single SS is a consequence of the inhomogeneity of systemic physiology across the surface of the scalp. This inhomogeneity of the systemic physiology was described in our previous work [43]. Because the two SS measurements are located 3 cm away from each other, they measure different superficial hemodynamic signals, which do not correlate exactly with one other. This is probably due to the inhomogeneous nature of the vascular network in the scalp.

Since the standard LS measurement constitutes an integratation of concentration changes throughout the volume traversed by the NIR light, the LS signal will contain systemic physiology that cannot entirely be described by a single SS signal, but will be better modelled by a linear combination of our two SS signals, as illustrated in Fig. 4-3. For this reason, using two SS measurements as regressor inputs to our Kalman filtering algorithm allows for improved recovery of the functional hemodynamic response signal of the brain

4.4.2 Source SS versus Detector SS

Because of the inherent symmetry of the measurement sensitivity between source and detector, a single SS measurement should perform identically whether it is placed close to the source optode or the detector optode of the LS measurement. In practice, small differences in performance can occur. As shown in Fig. 4-1, it was necessary to place an optical filter at the tip of the additional source to avoid saturation of the standard LS detector. In theory, the attenuation of the filter should be chosen such that the LS detector detects an equal amount of light coming from both sources. Because of the limited number of optical filter attenuations available, this condition is not always met. In our study, the attenuation of the filter selected (ND 2.00) was a little too high, resulting in a noisier signal and therefore a lower correlation between the SS signal close to the detector optode and the LS signal, as shown in Table 4.1. This resulted in a better performance of the Src SS compared to the Det SS. Using a NIRS system allowing the intensity of the sources to be adjusted easily would solve this problem.

4.4.3 HbO vs HbR

The results in Table 4.2 show that the SS method is more likely to improve the recovery of the HRF for HbO than for HbR. Futhermore, the improvements obtained using the SS method were larger for HbO than for HbR. These results are in agreement with previous studies [146, 44].

While the SS method did not improve the HbR recovery in 47% of cases in our simulations (based on the MSE metric), the amplitude of the decreases in performance was very low compared to the amplitude of the increases in performance resulting in an overall improvement when results were averaged across all simulations ($p<0.05$, two-tailed paired t-test) as shown in Fig. 4-5. This result is consistent with our previous work [44] which reported no decrease in performance for HbR, though only averaged results were provided. The fact that the likelihood of improvement is lower for HbR should not prevent the use of the SS method for HbR since potential decreases in performance will not be meaningful compared to the potential benefits. We showed in our previous work [43] that a high improvement could be obtained for HbR as long as the baseline correlation between the LS and the SS was high. However, in practice, lower baseline correlations are obtained for HbR resulting in lower overall averaged performance. The reason for this lower baseline correlation is unclear. A potential explanation is that the interference mainly comes from the arteries, which contain mostly oxygenated blood (HbO). The HbR component of the arterial blood is comparatively much weaker (and therefore more noisy) which results in lower baseline correlation. Future investigation will be required to identify the underling physiological mechanism. Nevertheless, researchers can always predict the performance of the Kalman regression algorithm by computing the baseline correlation between the LS and the SS channels and thus decide whether this processing approach is the most appropriate.

4.4.4 Practical difficulties

Using SS measurements both at the source and at the detector presents two challenges. The first is the challenge of ensuring that the light levels match well the dynamic range of the photo-detector. As discussed in section 4.4.2, an optical filter must be placed at the tip of the additional source to overcome the low dynamical range and avoid detector saturation. The second difficulty is the high density of fibers required which makes the optical probe heavier and less flexible.

The results presented in this paper suggest that real benefits are obtained using two SS measurements compared to using a single one. These benefits are significant in cases where only a few trials can be performed such as in clinical studies [54, 12, 139, 93]. However, in cases where several trials can be averaged, using a single SS detector located close to the source might be sufficient and less troublesome. Another example where two SS might be beneficial is when the NIRS signals are used to compute the cerebral metabolic rate of oxygen (CMRO2) [66, 127, 128, 144] since higher levels of physiological noise could affect the accuracy of the CMRO2 estimation.

4.4.5 Future directions

It has been recently shown that tomographic reconstruction can accurately remove superficial artifacts from NIRS data [51]. The researchers also compared the performance of tomographic reconstruction with and without short separation signal regression where the short separation was regressed using a static linear minimum mean square estimator (LMMSE) as in [116, 118]. They concluded that each method improves the contrast-to-noise ratio when used alone, and that the two methods act synergistically, with greater improvements when used together. Since the Kalman filter used in the present study has been shown to perform better than the static estimator [44] used by Gregg et al ([51]), it would be interesting in future work to investigate (1) how the Kalman filter compared with tomographic reconstruction alone, and (2) how much is gained when used together with tomographic reconstruction.

4.5 Summary

This work has demonstrated that using two short separation measurements, one located close to the source optode and one located close to the detector optode further improves the performance of the Kalman short separation regression method compared to using a single short separation measurement. Using both simulated and experimental functional activation, reductions of 59% in noise level and 72% in inter-trial variability were obtained for HbO and reductions of 47% in baseline noise level and 76% in inter-trial variability were obtained for HbR compared to a standard GLM approach. Our work emphasizes the importance of designing multi-distance NIRS probes that incorporate short separation measurements at both the source and the detector optode of the standard 3 cm NIRS channels.

Chapter 5

Quantification of pial vein signal

This section was published in:

Gagnon, L., Yucel, M. A., Dehaes M., Cooper, R. J., Perdue K.L., Selb J., Huppert T. J., Hoge R. D. and Boas, D. A. and (2013). "Quantification of the cortical contribution to the NIRS signal over the motor cortex using concurrent NIRS-fMRI measurements." *NeuroImage* 59: 3933-3940

The main contribution of this chapter was to investigate the cortical contribution to the NIRS signal using (1) Monte Carlo simulations over a realistic geometry constructed from anatomical and vascular MRI and (2) multimodal NIRS-BOLD recordings during motor stimulation. A good agreement was found between the simulations and the modeling analysis of *in vivo* measurements. While different stimuli will result in different pial vein contributions, our finger tapping results do reveal the importance of considering the pial contribution.

5.1 Introduction

Near-infrared spectroscopy (NIRS) [138, 63, 59] is a non-invasive technique for monitoring the hemodynamic changes occuring in superficial regions of the cortex. Using non-ionizing light, NIRS measures the fluctuations of the two dominant biological chromophores in the near-infrared spectrum: oxygenated (HbO) and deoxygenated or reduced hemoglobin (HbR).

Over the past 15 years, NIRS has become an attractive alternative to functional Magnetic Resonance Imaging (fMRI), with several clinical advantages [71]. NIRS is portable and less susceptible to movement artifacts enabling long term monitoring of the hemodynamic activity at the bedside. However, the spatial resolution of NIRS is 1-3 cm [7] which is less than the resolution of standard fMRI scanners. Another disadvantage of NIRS is its penetration depth which limits its sensitivity to the upper 1 cm of the cerebral cortex [7].

The biophysical origin of the functional NIRS signal is the variation of HbO and HbR concentration resulting from changes in Cerebral Blood Flow (CBF) and Cerebral Metabolic Rate of Oxygen ($CMRO_2$). For evoked brain activity, the resulting variations in HbO and HbR are described by the Balloon model [15, 42, 14]. According to this model, the arterial dilation driven increase in CBF following brain activation induces a passive volume increase of the capillary and venous vasculature (termed the windkessel compartment) as well as an increase in oxygen saturation. Compartmental microscopic hemodynamic measurements [58, 30] have revealed that this evoked oxygenation increase propagates through the pial veins located at the surface of the cortex but that this pial compartment exhibits very little volume variation following brain activation. This pial vein signal is generally negligible in fMRI since the high anatomical resolution allows the signal coming from the cortical region to be isolated. Conversely, the NIRS signal is integrated through the different superficial layers of the head. This potentially gives rise to a pial vein contamination of the signal if the pial vessels happen to coincide with the path of the light during its propagation through the tissue.

Preliminary analysis of the impact of pial vasculature in NIRS has been performed by [23]. However, very few studies of the effect of pial vein oxygenation changes (termed "the washout effect") on the NIRS signal have been performed [38, 66].

In this paper, the effect of pial vein contamination of the NIRS signal is investigated over the motor cortex where superficial pial veins are present [50]. We first quantify cortical and pial vein contributions to the NIRS signal by Monte Carlo simulation performed on a realistic anatomical volume containing pial veins acquired with MRI. We then use a biophysical model of the fMRI signal to analyze concurrent NIRS-fMRI data acquired over the motor cortex of human subjects during a finger tapping task. The cortical contribution to the HbR and HbO signals relative to the cortical contribution of the HbT signal are both estimated by fitting the biophysical model to the multimodal data.

5.2 Theory

We used biophysical modelling to investigate the contribution of cortical changes in HbO and HbR to the NIRS signal taken from *in vivo* NIRS-BOLD measurements. The Obata model [97], a refined version of the original Balloon model [15], describes the fluctuations in the BOLD signal as a function of the changes in deoxyhemoglobin (HbR) concentration and cerebral blood volume (CBV) in a given voxel:

$$\frac{\Delta\mathrm{BOLD}\,(t)}{\mathrm{BOLD}} = \mathrm{V}_0\left[(k_1+k_2)\left(1-\frac{\mathrm{HbR}\,(t)}{\mathrm{HbR}_0}\right)-(k_2+k_3)\left(1-\frac{\mathrm{CBV}\,(t)}{\mathrm{CBV}_0}\right)\right]. \tag{5.1}$$

All the parameters involved in the Obata model are summarized in Table 5.1.

The general idea behind our method is the following: NIRS can measure independently variations in cerebral blood volume and in the concentration of deoxyhemoglobin. Based on the Obata model, these are the two physiological phenomena giving rise to the BOLD signal. Therefore, one could potentially predict the BOLD

Table 5.1: Parameters of the Obata model

Symbol	Value	Description
V_0	-	Resting venous blood volume fraction
r_0	$100\ s^{-1}$	Taylor expansion of intravascular relaxation rate ΔR_{2I}
ν_0	$80.6\ s^{-1}$	Frequency offset at the surface of magnetized vessel
TE	20 or 30 ms	Echo-time of the MRI sequence
ϵ	0.59 for TE=30ms 0.70 for TE=20ms	Intrinsic/Extrinsic signal ratio
E_0	0.4	Baseline oxygen extraction fraction
k_1	$4.3 \cdot \nu_0 \cdot E_0 \cdot TE$	Lumped constant
k_2	$\epsilon \cdot r_0 \cdot E_0 \cdot TE$	Lumped constant
k_3	$\epsilon - 1$	Lumped constant

signal from the NIRS measurements. This idea has been investigated previously by [70, 66]. However, the NIRS signal is contaminated by pial vein washout [21, 23]. While the pial compartment cannot be extracted from the NIRS measurements, this component can be removed from the fMRI data because of the high spatial resolution of BOLD-fMRI. Because NIRS suffers from pial contamination, the NIRS-predicted BOLD signal will agree with the measured BOLD signal only if the NIRS data are corrected for pial vein washout.

To apply the above methodology, the Obata model must be modified to account for two discrepancies between BOLD-fMRI and NIRS. (1) Continuous-wave NIRS cannot measure relative changes in hemoglobin but rather a quantity proportional to absolute variations. (2) NIRS measures variations in total hemoglobin (HbT) rather than CBV. The Obata model as written in Eq. (5.1) contains only dimensionless variables i.e. all the variables are normalized by their value at rest. Making use of the relation

$$\mathrm{HbT}_0 = \frac{\mathrm{Hct}}{\mathrm{MW}_{Hb}} \cdot \mathrm{V}_0, \tag{5.2}$$

where Hct represents the hematocrit and MW_{Hb} the molecular weight of hemoglobin, the Obata model can be re-written in terms of absolute hemoglobin variations which can be measured by NIRS. The new model will be referred to the NIRS-adapted

Obata model:

$$\frac{\Delta \mathrm{BOLD}\,(t)}{\mathrm{BOLD}} = a_1 \cdot \Delta \mathrm{HbT} - a_2 \cdot \Delta \mathrm{HbR} \tag{5.3}$$

with

$$a_1 = \frac{\mathrm{MW}_{Hb}}{\mathrm{Hct}} \cdot (k_2 + k_3) \cdot \gamma^{HbT} \cdot \mathrm{PVC}^{HbT} \tag{5.4}$$

and

$$a_2 = \frac{\mathrm{MW}_{Hb}}{\mathrm{Hct}} \cdot \frac{(k_1 + k_2)}{1 + \mathrm{S_aO_2}\,(\mathrm{E_0} - 1)} \cdot \gamma^{HbR} \cdot \mathrm{PVC}^{HbR}. \tag{5.5}$$

The definition and value of the parameters involved in this new version are summarized in Table 5.2.

Table 5.2: Parameters of the NIRS-adapted Obata model

Symbol	Value	Description
MW_{Hb}	64 500 g/mol	Molecular weight of hemoglobin
Hct	160 g/L	Hematocrit
$^*\gamma^{HbT}$	~ 1	Cortical weighting factor for HbT
$^*\gamma^{HbR}$	to be estimated	Cortical weighting factor for HbR
$^{**}\mathrm{PVC}^{HbT}$	50	Partial volume correction factor for HbT
$^{**}\mathrm{PVC}^{HbR}$	50	Partial volume correction factor for HbR
$\mathrm{S_aO_2}$	0.98	Oxygen saturation in arteries

* γ handles cortical vs pial (which are both part of the activated brain tissue)

** PVC handles activated brain tissue vs scalp+skull+non-activated brain tissue

Generally, the pathlength of the light in the head can be separated into scalp and skull layers, non-activated brain tissue and activated brain tissue [126]. In this paper, we further separate the activated brain tissue into two regions: the cortical tissue and the pial vasculature. More specifically, the quantity PVC represents a partial volume correction extracting the signal coming from the activated brain tissue from the signal coming from the rest of the head (i.e. skin/skull and non-activated brain tissue crossed by the light). Finally, the variable γ represents the fraction of the activated brain tissue signal that is coming from cortical tissue while $1 - \gamma$ represents the fraction originating from pial vein oxygenation changes.

The coefficients a_1 and a_2 in Eq. (5.3) were estimated from the multimodal data set using a standard least-square method. ΔBOLD/BOLD was extracted from the fMRI data while ΔHbT and ΔHbR were computed from the NIRS data. Once a_1 and a_2 were recovered, the cortical contribution of HbR relative to HbT defined as $\gamma_r^{HbR} = \gamma^{HbR}/\gamma^{HbT}$ was computed using the following relation:

$$\gamma_r^{HbR} = \underbrace{\frac{a_2}{a_1}}_{\text{fit}} \cdot \underbrace{\frac{(k_2 + k_3)\left[1 + \mathrm{S_aO_2}\left(1 - \mathrm{E_0}\right)\right]}{k_1 + k_2} \cdot \frac{\mathrm{PVC}^{HbT}}{\mathrm{PVC}^{HbR}}}_{\text{assumed}}. \tag{5.6}$$

where a_1 and a_2 were obtained from the least-square fit and the value for the rest of the parameters taken from the literature.

In our model, we separately account for the volume fraction of the activated brain tissue (with PVC) and the cortical vs pial composition of the activated tissue (with γ). With the above definition, PVC depends on the wavelengths of the NIRS sources and γ does not. [126] showed that measurements performed at 690 nm and 830 nm minimize the cross-talk between HbT and HbR introduced by incorrect values of PVC. Since these wavelengths were used in our measurements, crosstalk between HbR and HbT was negligible and it was reasonable to assume $\mathrm{PVC}^{HbT} = \mathrm{PVC}^{HbR}$. As such, these two factors cancel out in the estimation of γ_r^{HbR} with Eq. (5.6), as do Hct and MW_{Hb}. Our estimation of γ_r^{HbR} was therefore independent of the values assumed for these parameters. Under these assumptions, Eq. (5.6) reduces to

$$\gamma_r^{HbR} = \underbrace{\frac{a_2}{a_1}}_{\text{fit}} \cdot \underbrace{\frac{(k_2 + k_3)\left[1 + \mathrm{S_aO_2}\left(1 - \mathrm{E_0}\right)\right]}{k_1 + k_2}}_{\text{assumed}}. \tag{5.7}$$

[21] showed that the spatial extend of the activation region measured by HbT is smaller than the one measured by HbO or HbR. The larger activation region measured by HbO and HbR was attributed to potential washout of the deoxyhemoglobin in the pial vasculature during the activation. The same phenomenon of extended pial vein washout is handled by the parameter γ in our model. Pial veins exhibit only very small volume changes following brain activation [21, 58, 30]. Therefore, the value of

γ^{HbT} is very close to 1 indicating that most of the HbT signal is coming from the cortical region. Under this assumption, a pure washout of HbR is observed and thus the amplitude of the pial increase in HbO matches the amplitude of the pial decrease in HbR:

$$\left(1-\gamma^{HbO}\right)\Delta\mathrm{HbO} = -\left(1-\gamma^{HbR}\right)\Delta\mathrm{HbR} \tag{5.8}$$

where $(1-\gamma)$ indicates the pial fraction of the signal. One can derive the expression for the cortical weighting factor for HbO (γ^{HbO}) from Eq. (5.8):

$$\gamma^{HbO} = \frac{\Delta\mathrm{HbT} - \gamma^{HbR}\cdot\Delta\mathrm{HbR}}{\Delta\mathrm{HbO}}. \tag{5.9}$$

The parameter γ^{HbR} alone cannot be extracted from the *in vivo* data but under the assumption that $\gamma^{HbT} \approx 1$, the approximation $\gamma_r^{HbR} \approx \gamma^{HbR}$ holds. We will refer to γ_r^{HbO} as the estimation of γ^{HbO} under this assumption by substituting γ^{HbR} by γ_r^{HbR} in Eq. (5.9):

$$\gamma_r^{HbO} = \frac{\Delta\mathrm{HbT} - \gamma_r^{HbR}\cdot\Delta\mathrm{HbR}}{\Delta\mathrm{HbO}} \tag{5.10}$$

where the subscript "r" emphasizes the assumption of no pial volume changes.

Conversely, in the cortical tissues the amplitude of the increase in HbO does not match the decrease in HbR because vascular dilation gives rise to a change in blood volume [15]. Therefore, the cortical fraction of the signal, γ, given by the ratio of the cortical signal over the total activated tissue signal $\left(\frac{\text{cortical}}{\text{cortical+pial}}\right)$, will be different for HbR and HbO. Since the amplitude of the cortical ΔHbR is lower than the amplitude of the cortical ΔHbO, the relative cortical weighting factor will be lower for HbR compared to HbO ($\gamma_r^{HbR} < \gamma_r^{HbO}$).

5.3 Methodology

5.3.1 Monte Carlo simulations

The effect of pial vein washout was first investigated using Monte Carlo simulation. The anatomical model used in the simulation was the same as in [23]. A high resolution anatomical T1 MRI image was acquired and tissues where then segmented in four different types: scalp, skull, cerebrospinal fluid (CSF) and brain tissue (containing both white and gray matter). The segementation was performed using the Matlab package SPM8. A phase contrast MR angiography was also used to image the pial vasculature located at the surface of the cortex [23, 95]. A saturation band was applied to suppress the arterial blood signal such that only pial veins were kept in the final anatomical model. The selected velocity encoding was 5 mm/s for the plane described by the right-left and anterior-posterior encoding. The bandwidth was 300 Hz/px and imaging times were TR/TE = 80.9/10.3 ms. The flip angle was set to 15°. More information about the sequence parameters can be found in [23]. An illustration of the anatomical model with the position of the optical sources and detectors is shown in Fig. 5-1.

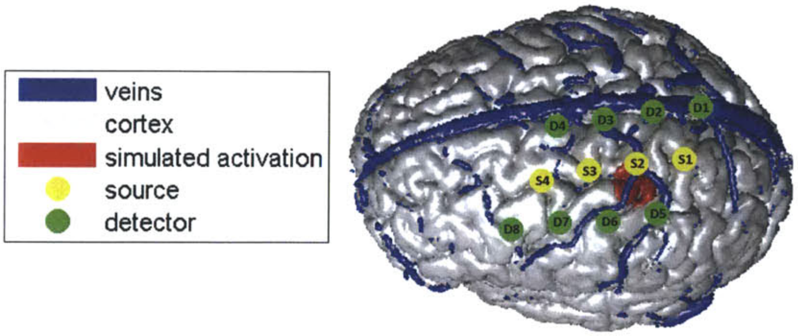

Figure 5-1: Anatomical model used in the Monte Carlo simulations. The position of the optical sources and detectors is also illustrated as well as the simulated region of activation.

Specific optical properties were assigned to each tissue type and are summarized in Table 5.3. These values were computed using the model from [126] which takes into

Table 5.3: Optical properties assigned to the different tissue types for the Monte Carlo simulations. Absorption coefficient μ_a [mm^{-1}], scattering coefficients μ_s [mm^{-1}], anisotropic factors g and refractive indexes n. The brain tissue includes gray and white matter. Values were computed using the method given in [126].

	$\lambda = 690$ nm				$\lambda = 830$ nm			
Tissu	μ_a	μ_s	g	n	μ_{a_0}	μ_s	g	n
Scalp	0.0159	8.000	0.900	1.4	0.0191	6.600	0.900	1.4
Skull	0.0101	10.00	0.900	1.4	0.0136	8.600	0.900	1.4
CSF	0.0004	0.100	0.900	1.4	0.0026	0.100	0.900	1.4
Brain	0.0178	12.50	0.900	1.4	0.0186	11.10	0.900	1.4
Pial veins	0.5745	74.50	0.985	1.4	0.4758	67.50	0.992	1.4

account the blood content and the oxygen saturation of each tissue type. A baseline oxygen saturation of 60 % was assumed in the pial veins [13]. In the specific cortical region shown in red in Fig. 5-1, we simulated an increase in HbO of +9 μM and a decrease in HbR of -3 μM. In our simulation, we also induced an increase in oxygen saturation in the pial veins (SpO$_2$) ranging from 5 to 10 %, as reported in the literature [13, 142].

Monte Carlo methods were used to simulate the photon fluence at each NIRS detector [6]. Fluences were computed for a baseline state i.e. using the optical properties given in Table 5.3 as well as for an activated state i.e. modifying slightly the optical properties of the brain tissue and the pial veins in Table 5.3 to take into account the cortical activation as well as the change in oxygen saturation in the pial veins. We then used the modified Beer-Lambert law [24, 19] to recover the hemoglobin concentration changes from the fluences detected during the baseline and the activated state. In our simulations, the relationship between the HbR change simulated in the cortical region ($\Delta\mathrm{HbR}_c^{sim}$) and the total HbR change recovered at the detector ($\Delta\mathrm{HbR}_{tot}^{det}$) containing both cortical and pial contributions is given by

$$\Delta\mathrm{HbR}_c^{sim} = \underbrace{\Delta\mathrm{HbR}_{tot}^{det} \cdot \mathrm{PVC}^{HbR}}_{\Delta\mathrm{HbR}_{tot}^{sim}} \cdot \gamma^{HbR} \tag{5.11}$$

with the Partial Volume Correction factor (PVC^{HbR}) relating the total detected HbR change ($\Delta\text{HbR}_{tot}^{det}$) to the total simulated HbR change ($\Delta\text{HbR}_{tot}^{sim}$) in the brain tissue. Similar equations hold for HbT

$$\Delta\text{HbT}_{c}^{sim} = \underbrace{\Delta\text{HbT}_{tot}^{det} \cdot \text{PVC}^{HbT}}_{\Delta\text{HbT}_{tot}^{sim}} \cdot \gamma^{HbT} \tag{5.12}$$

and for HbO

$$\Delta\text{HbO}_{c}^{sim} = \underbrace{\Delta\text{HbO}_{tot}^{det} \cdot \text{PVC}^{HbO}}_{\Delta\text{HbO}_{tot}^{sim}} \cdot \gamma^{HbO} \tag{5.13}$$

Without simulating volume changes in the pial veins, $\gamma^{HbT} = 1$ and from Eq. (5.12) we find that $\text{PVC}^{HbT} = \Delta\text{HbT}_{c}^{sim}/\Delta\text{HbT}_{tot}^{det}$. We first verified that using tissue optical properties for 690 nm and 830 nm, the partial volume correction factor (PVC) as defined in this paper was the same for HbR, HbT and HbO by running a simulation with an activation in the cortical region only i.e. $\gamma^{HbR} = 1$ and $\gamma^{HbO} = 1$ (see Fig. 7-2A). Once this assumption was justified, a second simulation was run with a washout effect in the pial vasculature on top of a cortical activation (see Fig. 7-2B). From this second simulation, the cortical weighting factor for HbR and HbO were computed using Eq. (5.11) with $\text{PVC}^{HbR} = \text{PVC}^{HbT}$ and Eq. (5.13) with $\text{PVC}^{HbO} = \text{PVC}^{HbT}$ respectively

$$\gamma^{HbR} = \frac{\Delta\text{HbR}_{c}^{sim}/\Delta\text{HbR}_{tot}^{det}}{\Delta\text{HbT}_{c}^{sim}/\Delta\text{HbT}_{tot}^{det}} \tag{5.14}$$

$$\gamma^{HbO} = \frac{\Delta\text{HbO}_{c}^{sim}/\Delta\text{HbO}_{tot}^{det}}{\Delta\text{HbT}_{c}^{sim}/\Delta\text{HbT}_{tot}^{det}}. \tag{5.15}$$

5.3.2 In vivo studies

Two different studies were performed. In each of them, we recorded concurrent NIRS-fMRI during a motor task which consisted of finger tapping blocks with a duration of either 2 s or 20 s.

In the first study, simultaneous NIRS-fMRI data were acquired during 2 s finger tapping blocks. These data have previously been described by [69, 70, 66]. Six subjects were scanned and completed six runs containing 30 individual events each. The NIRS data were bandpass filtered at 0.03-0.8 Hz before being block-averaged across all runs and all individual events. The source–detector pairs to be included in the region-of-interest average were chosen for each subject consistently with the fMRI data and from measurements showing statistically significant changes ($p < 0.05$). The grand average across the six subjects was then computed. The BOLD data were acquired using parameters TR/TE/θ=500 ms/30 ms/90°. The functional images were first motion corrected and spatially smoothed with a 6-mm Gaussian kernel. For each subject, a t-statistic map was generated and threshold was applied for $p < 0.05$. Significant voxels were then manually selected under the NIRS probe based on fiducial markers and the grand average was computed across all six subjects. The hemodynamic response functions were then calculated by an ordinary least-squares linear deconvolution with a third order polynomial drift. Each of the ΔHbO, ΔHbR, ΔHbT and ΔBOLD time courses were adjusted to cross y=0 at t=0 by subtracting their respective value at t=0.

In the second study, concurrent NIRS-ASL data were acquired during 20 s finger tapping blocks. These data have been previously published by [62]. The NIRS data were bandpass filtered at 0.0167-0.5 Hz. The source–detector pairs to be included in the region-of-interest average were chosen for each subject consistently with the fMRI data and from measurements showing better than $p < 0.05$ significance. The grand average across the five subjects was then computed. The BOLD signal was extracted from the control images of the ASL acquisition [62] with parameters TR/TE=2 s/20 ms. For each subject, we generated a t-statistic map to identify regions of significant response. Each t-map was thresholded at $p < 0.05$ to compute the hemodynamic response. This resulted in a focal ROI positioned on the precentral gyrus and located beneath the source-detector array of the optical probe. The hemodynamic responses were computed from the ROI using a linear deconvolution model with a third order polynomial drift. Each of the ΔHbO, ΔHbR, ΔHbT and ΔBOLD time courses were adjusted to cross y=0 at t=0 by subtracting their respective value at t=0.

For each study, we fitted Eq. (5.3) to the averaged ΔBOLD, ΔHbT and ΔHbR reponses with a least-square method to recover a_1 and a_2 before computing γ_r^{HbR} using Eq. (5.7). The parameter γ_r^{HbO} was then estimated using Eq. (5.10).

5.4 Results

5.4.1 Simulation results

Results of the Monte Carlo simulations arc summarized in Fig. 7-2. Panel (A) shows the simulated concentration changes when no increase in oxygen saturation was simulated in the pial veins. Using Eqs. (5.11-5.13) with $\gamma^{HbR} = \gamma^{HbT} = \gamma^{HbO} = 1$, the PVC^{HbR}/PVC^{HbT} and PVC^{HbO}/PVC^{HbT} ratios computed were very close to 1 confirming that the Partial Volume Correction factor (PVC) was the same for HbR, HbO and HbT. The region of interest (ROI) was defined by the three source-detector pairs showing the strongest activation. Panel (B) shows the simulated concentration changes when we added a 6% increase (60 to 66 %) in oxygen saturation in the pial veins on top of the cortical activation (+9 μM ΔHbO and -3 μM ΔHbR). The average cortical weighting factor γ computed with Eqs. (5.14-5.15) and taken over the three source-detector pairs of the ROI was 0.17 (or 17 %) for HbR and 72 % for HbO. We also ran the simulation inducing different oxygenation increases in the pial veins ranging from 5 to 10 %. The γ^{HbR} values recovered in each case are shown in Fig. 7-2C and ranged from 11 to 19%, while the γ^{HbO} values were less sensitive and decreased from 73 % to 70 % as shown in Fig. 7-2D.

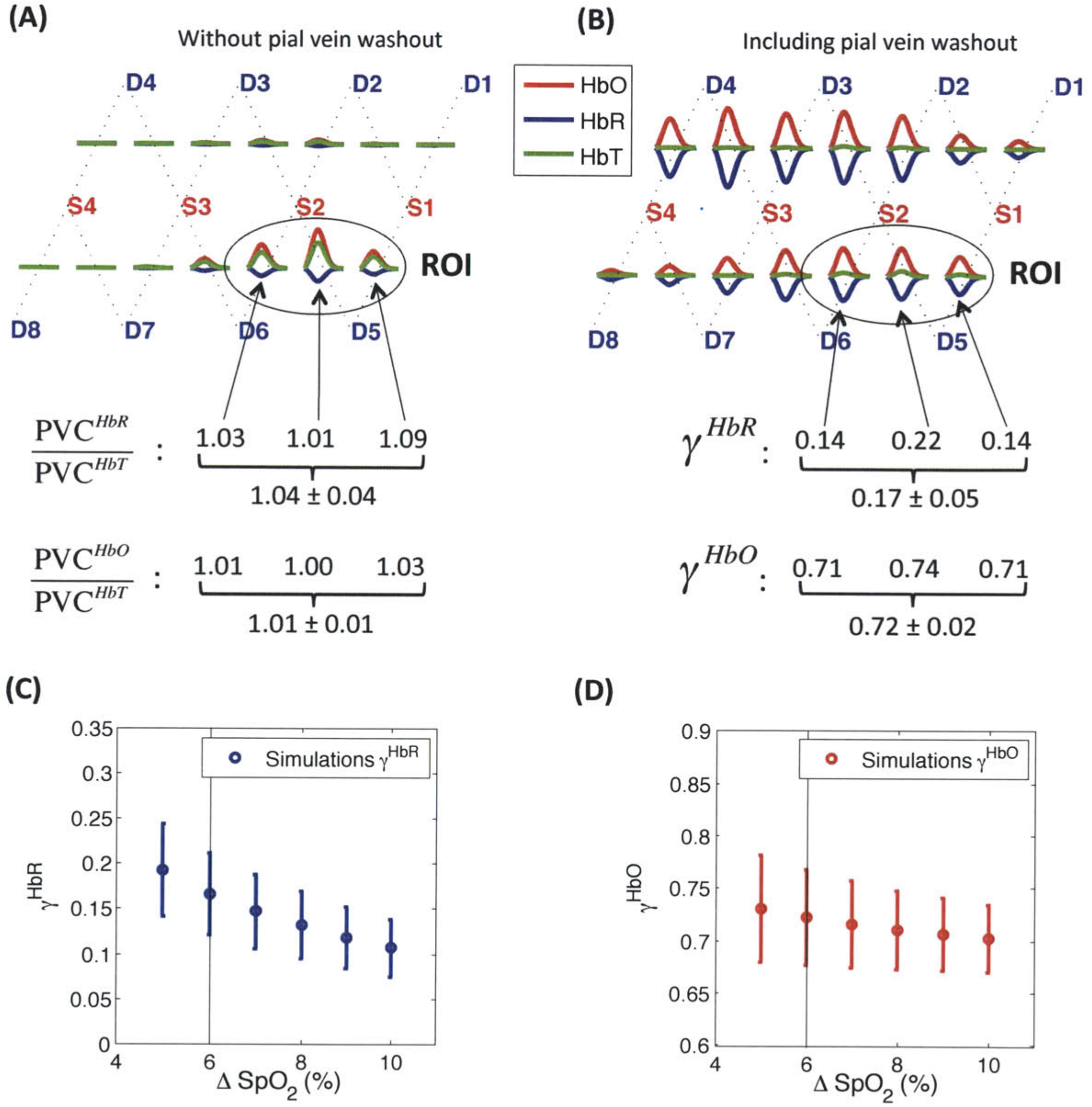

Figure 5-2: Results of the Monte Carlo simulations performed on the numerical volume shown in Fig. 5-1. (A) Simulation of a cortical activation with no pial vein washout ($\gamma^{HbR} = 1$ and $\gamma^{HbO} = 1$) to define the ROI and verify that $\text{PVC}^{HbR} = \text{PVC}^{HbT}$ and $\text{PVC}^{HbO} = \text{PVC}^{HbT}$. (B) Simulation of a pial veins oxygenation increase of 6 % (60 to 66%) on top of a cortical activation. No volume changes were simulated in the pial veins ($\gamma^{HbT} = 1$). (C-D) Sensitivity analysis showing (C) the γ^{HbR} and (D) the γ^{HbO} values recovered by simulating an oxygenation increase of 5 to 10 %.

5.4.2 Modeling analysis of *in vivo* measurements

The average experimental time course for ΔHbO, ΔHbR, ΔHbT and ΔBOLD during functional activation are shown in Fig. 5-3. The γ_r^{HbR} and γ_r^{HbO} values recovered with Eqs. (5.7) and (5.10) for each data set are also shown. The γ_r^{HbR} value was 18% for the 2 s finger tapping study and 23% for the 20 s finger tapping study. The γ_r^{HbO} value was 77% for both studies. The uncertainty for γ_r^{HbO} was larger than γ_r^{HbR} because its calculation using Eq. (5.10) requires additional experimental measurements, which introduced additional errors.

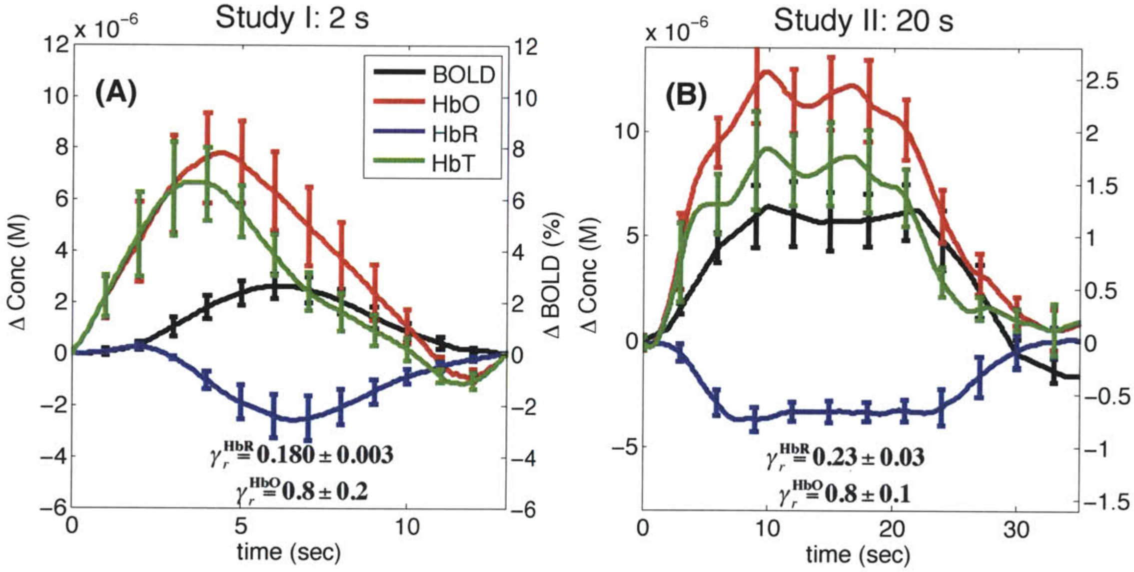

Figure 5-3: Finger tapping results. The average traces for ΔHbO, ΔHbR, ΔHbT and ΔBOLD are shown as well as the γ_r^{HbR} and γ_r^{HbO} values recovered with the NIRS-adapted Obata model. Error bars represent the standard error computed across the five subjects. (A) Study I: 2 s finger tapping. (B) Study II: 20 s finger tapping.

The value recovered in each case depends of the value assumed for six parameters: $\frac{\mathrm{PVC}^{HbR}}{\mathrm{PVC}^{HbT}}$, E_0, ϵ, ν, r_0 and SaO_2. A sensitivity analysis was performed for each of these parameters and the results are shown in Fig. 5-4 for γ_r^{HbR}. Sensitivity for γ_r^{HbO} is not shown since γ_r^{HbO} was computed directly from γ_r^{HbR} using Eq. (5.10). In each case, the value of a single parameter was varied while all the other parameters were kept constant. For each parameter, the black vertical line indicates the value assigned to this parameter while varying the value of another parameter. For all parameters except ϵ (the Intrinsic/Extrinsic fMRI signal ratio), the γ_r^{HbR}'s recovered ranged between 15 and 30% when varying the Obata model parameters over a reasonable range. However, the recovery was more sensitive to the value assumed for ϵ and γ_r^{HbR} ranged from 10 to 100% when ϵ varied from 0.5 to 1.4.

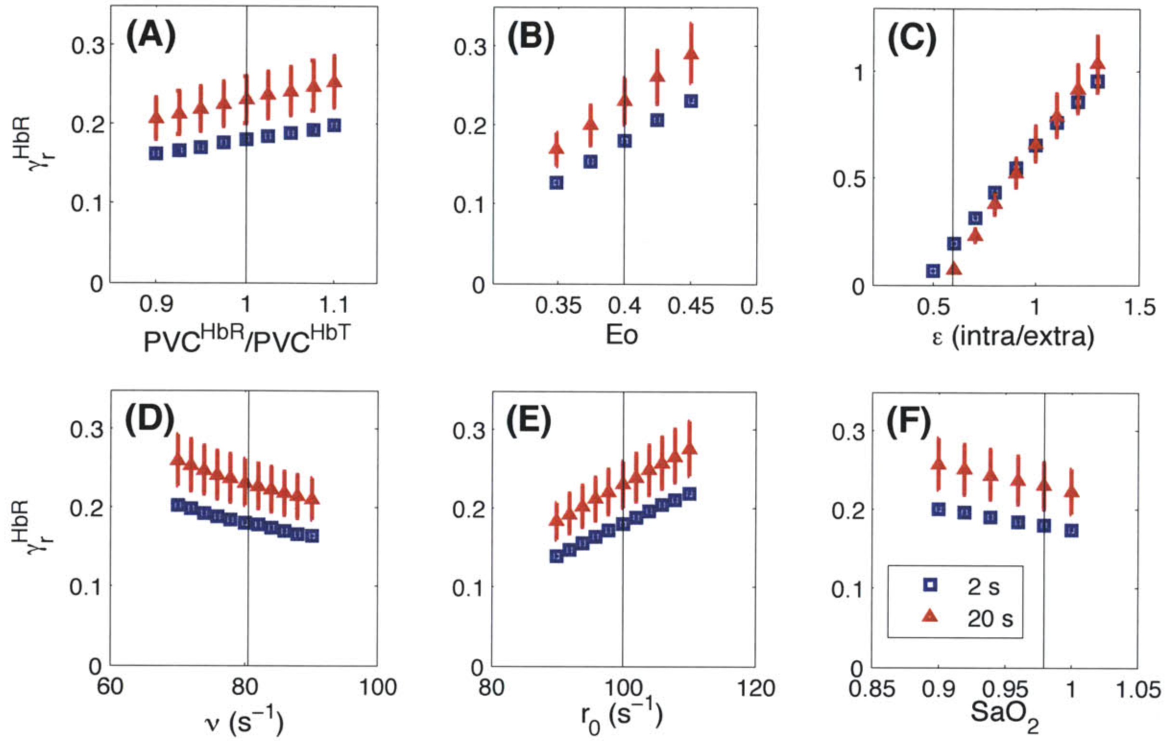

Figure 5-4: Sensitivity analysis for different parameters of the model. (A)-(F) Sensitivity of the recovered γ_r^{HbR} from our *in vivo* data for different values assumed for the Obata model parameters. Results are illustrated for both the 2 s (blue) and the 20 s (red) finger tapping studies. The vertical lines show the reference values indicated in Tables 5.1 and 5.2.

5.4.3 Combined results

The combined results from the simulations and the modeling of the *in vivo* measurements are summarized in Fig. 5-5. The average cortical weighting factor value was 19% ± 3% for HbR and 76% ± 3% for HbO.

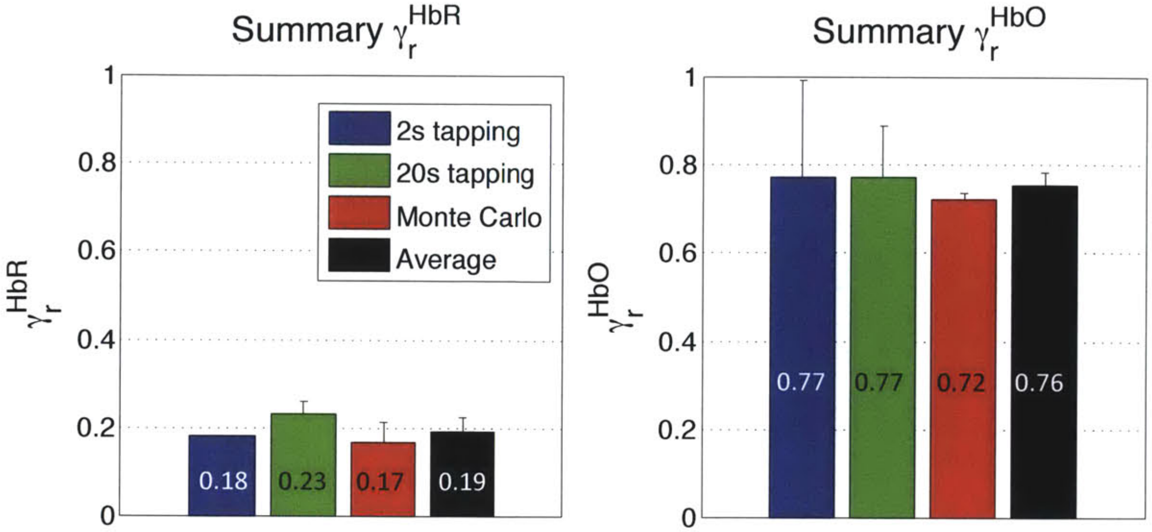

Figure 5-5: Summary of cortical weighting factors (γ_r^{HbR} and γ_r^{HbO}). Results are shown for the two *in vivo* studies, for the Monte Carlo simulations as well as the grand average.

5.5 Discussion

5.5.1 Cortical contribution to the NIRS signal

As shown in Fig. 5-5, our simulations agreed very well with our modeling results from the *in vivo* measurements, confirming the assumptions made in our computations. The good agreement between the γ^{HbR} values computed from simulations (where γ^{HbT} was forced to be 1) and the γ_r^{HbR} values computed from the experimental data indicates that γ^{HbT} modeled from the *in vivo* measurements was very close to 1. This result is strongly supported by exposed cortex animal imaging models [58, 30]. Therefore, our *in vivo* results can be interpreted on the same footing as our simulations results.

Our combined results (averaged from simulations and *in vivo* modeling) indicate that for a task-evoked response over the motor cortex, the cortical contribution of the detected ΔHbR signal corresponds to 19% of the cortical contribution of the ΔHbT signal e.g. that the cortical contribution is 5 times smaller for ΔHbR compared to

ΔHbT. For ΔHbO, this value was found to be 76%. Our results suggest that both the ΔHbO and ΔHbR signals contain a significantly higher pial vein contribution compared to ΔHbT. It is therefore likely that the ΔHbT signal will provide better spatial specificity [122, 21] and should be used instead of ΔHbO or ΔHbR to map cerebral activity with NIRS.

Since our results suggest that no pial volume change occurred, $\gamma_r^{HbR} \approx \gamma^{HbR}$ and thus our numbers indicate that 19% of the entire ΔHbR signal and 76% of the entire ΔHbO signal is coming from the cortical region. The remaining 81% and 24% of the signal for ΔHbR and ΔHbO respectively originate from the pial veins located at the surface of the motor cortex, where a change in oxygen saturation takes place following brain activation. This finding is not surprising since NIRS exhibits a sensitivity profile that exponentially decreases with depth [7] and the pial vasculature is located above the surface of the cortex.

Previous work by [66] using the same data set used in our first study (2 s tapping) found that 44% of the NIRS signal (accouting for both HbO and HbR) was coming from the cortical region. Taking the average of our γ_r^{HbR} and γ_r^{HbO} value gives 48% which agrees well with this value.

Our sensitivity analysis revealed that the γ_r^{HbR} values found from the *in vivo* data were not very sensitive to the parameters assumed in the Obata model, except for the ϵ parameter representing the Intrinsic/Extrinsic fMRI signal ratio. Assuming larger ϵ values would change our results by increasing γ_r^{HbR} as shown in Fig. 5-4C. However, recent work by [52] supports the accuracy of our results by reporting similar ϵ values to those used in our study. Using a detailed analysis, Griffeth and Buxton showed that for TE=32 ms, the ϵ value is close to 0.5 in veins and goes up to 1 and 1.3 in the capillaries and the arteries respectively. In our work, the arterial compartment was assumed negligible since most of the BOLD signal is coming from the veins. The capillary compartment was combined with the veins in a single compartment model as in [97]. Using the equations in Obata *et al*, we computed an effective ϵ value of 0.58

and 0.7 for TE=30 ms and TE=20 ms respectively, which agrees well with Griffeth and Buxton's findings.

5.5.2 Impact on NIRS-fMRI $CMRO_2$ estimation

Our results suggest that caution must be taken in quantifying cortical cerebral activity from NIRS measurements acquired over a region where large pial veins are located. If concurrent fMRI data are available, the NIRS-adapted Obata model (Eq. 5.3) allows us to compute the cortical contribution under some realistic assumptions.

The Cerebral Metabolic Rate of Oxygen ($CMRO_2$) can be estimated from concurrent NIRS-fMRI recordings [62, 66]. The pial compartment was already taken into account in [66]'s computation. In Hoge *et al* [62], $CMRO_2$ was computed from the NIRS data without any corrections for pial vein washout. $CMRO_2$ was given by the product of the change in oxygen extraction fraction (E) and the change in CBF: $rCMRO_2 = rE \times rCBF$, where r denotes a quantity normalized by its baseline value. The change in CBF was estimated from the ASL data while the change in oxygen extraction fraction was computed from the NIRS data: $rE = rHbR/rHbT$. Following activation, CBF increases ($rCBF > 1$) and the oxygen extraction fraction decreases ($rE < 1$) [15]. Correcting rHbR for the pial contamination lowers the decrease in HbR resulting in a higher rE and a higher $rCMRO_2$. Depending on the baseline hemoglobin concentrations, this correction can be significant. With the values given in [62], the pial vein correction increases $rCMRO_2$ by 15% (from 1.40 to 1.61) while this value increased from 1.26 to 1.49 assuming a lower baseline oxygen saturation measured by time-resolved spectroscopy [45].

5.5.3 Limitations and future studies

The values for the oxygenation increase induced in the pial veins in our Monte Carlo simulation were taken from the literature [13, 142]. This was necessary since we did not monitor the response in the pial veins during our fMRI scans. Future studies could use magnetic resonance susceptometry-based oximetry [72] to monitor oxygen saturation in the pial veins following activation. These values would be more accurate than the ones computed from a simplified model and could be used in the Monte Carlo simulations to give more accurate results.

In our analysis, we have ignored any global systemic flow variations in the skin that could be phase-locked with the hemodynamic response in the brain. These flow changes in the skin potentially contaminate the NIRS signal because of its high sensitivity to superficial tissues. If true, these contributions would have simply been integrated with the pial vein washout $(1-\gamma)$ in our *in vivo* analysis without changing the conclusion of this paper. However, we have completely ignored any skin contributions in our MC simulations. Such skin artefact should also be considered in future work. Preliminary work by [79] analysed the fMRI response in the skin located under a NIRS probe and was able to relate these responses to artefacts in the NIRS signal acquired on the forehead.

Our setup did not allow us to acquire a pial angiogram and functional data on the same subjects. Even though our simulation results matched our *in vivo* analysis, we could not assess inter-subject vascular-based variability. Future multimodal studies will be required to quantify how the geometry of the vasculature impacts the NIRS signal. Preliminary work by [106] has shown that the inter-subject variability of NIRS can be partially attributed to anatomical vasculature differences.

Our study was performed over the motor cortex where pial vasculature is known to be present at the surface. Other regions might exhibit stronger contamination due to larger vessels, such as the sagittal sinus in the visual cortex [23], or weaker contamination due to smaller vessels as in the forehead. Future studies will be required to map the contribution of the pial vasculature in the NIRS signal across different regions of the brain.

5.6 Summary

We have shown that the NIRS signal collected over the motor cortex during an evoked task contained a smaller cortical contribution for ΔHbR and ΔHbO compared to ΔHbT. The cortical contribution to ΔHbR was equal to 19% of the cortical contribution to ΔHbT. Similarly, the cortical contribution to ΔHbO was equal to 76% of the cortical contribtion to ΔHbT. Our results suggest that the pial contamination is less important for ΔHbT, and therefore, the ΔHbT signal should be used rather than ΔHbO or ΔHbR to map cerebral activity with NIRS. This pial vein contamination has a significant impact on the quantification of cerebral activity using NIRS and correction factors must be applied in order to compute $CMRO_2$ from concurrent NIRS-fMRI measurements.

Chapter 6

Modeling the fMRI signal from two-photon measurements

The main contribution of this chapter is to develop a modeling approach to predict the fMRI signals from two-photon vascular stacks filled with oxygen distribution. We performed Monte Carlo simulations over real microvascular networks and oxygen distributions measured *in vivo* on rodents, at rest and during forepaw stimulation, using two-photon microscopy. Our model reveals for the first time the specific contribution of individual vascular compartment to the BOLD signal, for different field strengths and different cortical orientations.

6.1 Introduction

Over the last 20 years, functional magnetic resonance imaging (fMRI) has been driving a revolution in the brain sciences, providing new insights into the brain's functional organization. The celebrated Blood Oxygen Level Dependent (BOLD) contrast [101, 82] is still the most widely used fMRI method [89]. Despite its widespread utilization, the physiological mechanism that gives rise to the BOLD signal measured during cerebral activation is still poorly understood [89, 78]. One of the reasons for this challenge is the difficulty of measuring blood oxygenation in the microvascular compartments during cerebral activation and linking these physiological changes to the BOLD signal measured [135, 16]. Understanding how individual vascular compartments are reflected in the BOLD signal for different pulse sequences and for different B-field strengths is critical for the choice of sequence parameters, for the development of new quantitative fMRI methods, for the development of high-field fMRI technologies and for the interpretation of neuroimaging data in the context of vascular diseases. Moreover, quantifying how the signal varies across regions with different spatial orientations and correcting for this potentially confounding factor will positively impact our interpretation of BOLD images.

Recent advances in multiphoton microscopy have opened the door to measure microvascular oxygenation *in vivo* during cerebral activation in rodents [119, 83, 25, 103]. Here, we take advantage of this technology to measure microvascular networks with oxygen distribution *in vivo* on rodents, at rest and during forepaw stimulation. These tridimensional physiological measurements are then used to predict fMRI signals from first principles. Monte Carlo simulations of nuclear spins diffusing over the measured vascular networks placed in an external magnetic field are performed. The resulting BOLD signal is validated against experimental BOLD data measured under the same physiological conditions. Using this modeling framework, we compute the individual contribution of arteries, capillaries and veins to the BOLD signal for different B-field strengths and for both gradient echo (GRE) and spin echo (SE) pulse sequences.

We also study the effect of spatial orientation in the MRI scanner and validate our prediction against experimental BOLD measurements.

6.2 Reconstruction of baseline oxygen distribution across real vascular networks and validation against experimental pO_2 measurements

Two-photon microscopy was carried out first on a set of anesthetized mice (n=6) as described in the Supplementary Information. Briefly, an intravascular oxygen-sensitive nanoprobe (PtP-C343) was injected for the pO_2 measurements followed by the injection of FITC for angiography. The angiogram for a representative animal is shown in Fig. 6-1a.

In order to reconstruct microvascular oxygenation with sufficient spatio-temporal resolution to accurately model the BOLD signal, a vascular anatomical network (VAN) model was created for each animal [8, 35, 5] (see Supplementary Information for details). Angiograms were graphed using a suite of custom-built computer programs [35, 134] and a mesh of the vasculature was created for each animal [35]. Each vessel segment was identified as an artery, a capillary or a vein as shown in Fig. 6-1b. Blood flow distribution as shown in Fig. 6-1c was obtained after computing the resistance for each vascular segment on the graphs. The oxygen distribution was simulated with the VAN by using the experimental pO_2 measurements as boundary conditions and using a finite element approach described in [35]. An example of the resulting distribution of pO_2 obtained is shown in Fig. 6-1d and compared against the experimental pO_2 measurements in Fig. 6-1e. The agreement between the simulated and the experimental pO_2 measurements is demonstrated in Fig. 6-1f where both pO_2 and oxygen saturation (SO_2) are compared as a function of branching order from pial vessels.

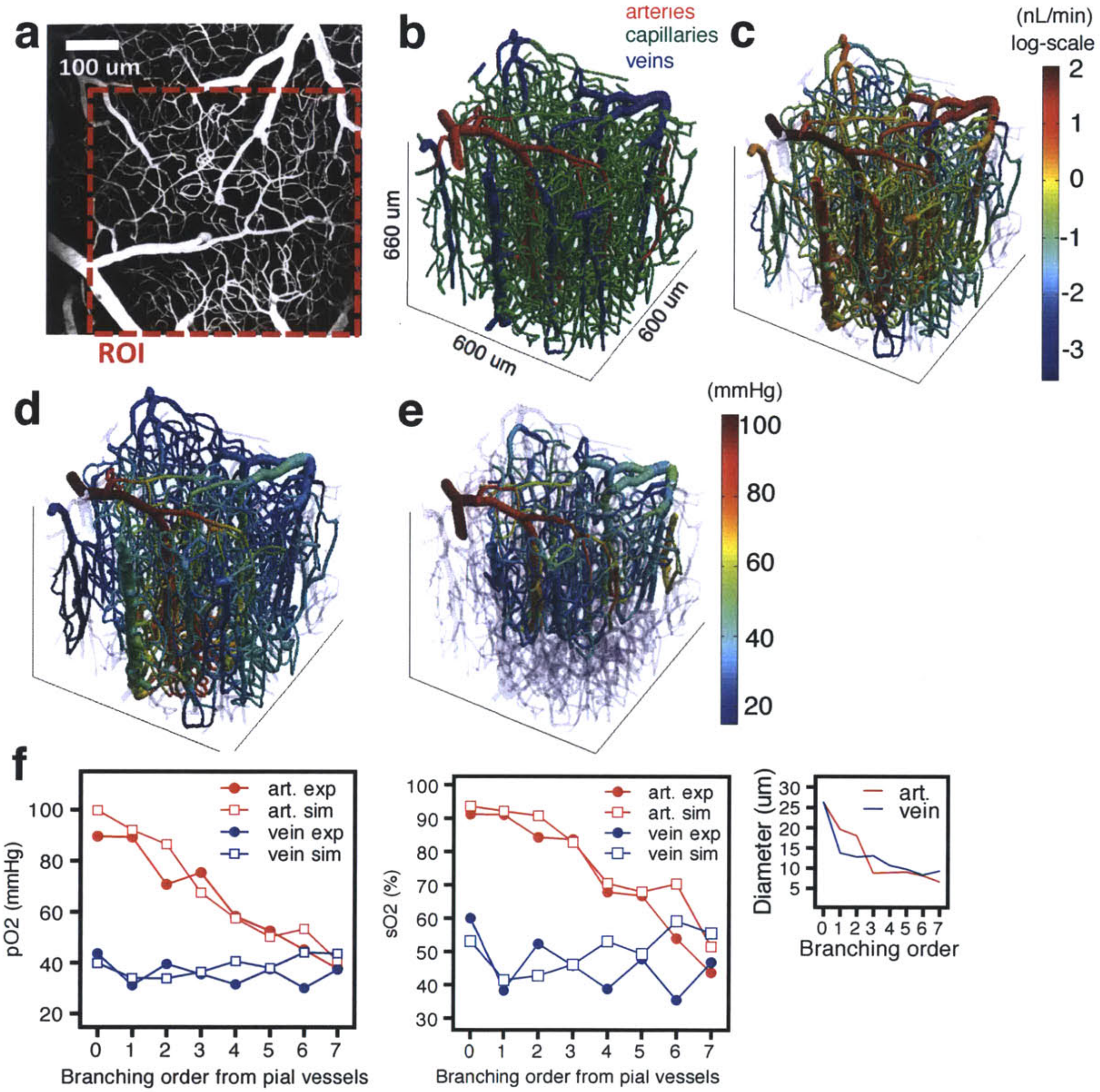

Figure 6-1: Construction of realistic vascular networks. (a) TPM FITC angiogram of the mouse cortex. (b) FEM mesh of the vasculature displaying arteries, capillaries and veins. (c) Blood flow distribution simulated across the vascular network assuming a global perfusion value of 100ml/min/100g. (d) Distribution of the partial pressure of oxygen (pO_2) simulated across the vascular network using the FEM model. (e) TPM experimental measurements of pO_2 in vivo using PtP-C343 dye. (f) Quantitative comparison of simulated and experimental pO_2 and SO_2 distributions across the vascular network.

6.3 Computation of physiological changes during forepaw stimulation and validation against experimental measurements

The stimulus used for our functional measurements was a 2-sec long electrical stimulation of the forepaw. The dilation of the vasculature following the stimulus was measured on a separated set of anesthesized rats (n=19) using two-photon microscopy as described in [?] and in the Supplementary Information. Time courses of the dilation were averaged for the surface pial arteries, the arterial diving trunk as well for the 1st and 2nd branching pre-capillary arterioles. These traces are shown in Fig. 6-2a. These time courses were used as inputs to the VAN model to compute the resulting changes in flow and volume across the entire network. The averaged flow changes and volume changes (both relative to baseline) for individual compartments are shown in Fig. 6-2b and Fig. 6-2c respectively. Note that no venous dilation was observed in our two-photon measurements, a fact consistent with other studies involving a cranial window [59, 86]. However, a venous dilation of 1-2% was computed from the VAN model and used in the fMRI model, which is consistent with two-photon measurements during short stimulus [30] under a reinforced thinned skull window [31].

Changes in oxygen saturation were computed from the VAN model given changes in flow and volume and assuming a ΔFlow/ΔCMRO2 ratio of 3, which is the typical value measured in rodents for short stimulations [68, 32]. Simulated SO_2 changes are shown in Fig. 6-2d for different vascular compartments. To validate these simulations, SO_2 measurements during functional stimulation were performed in pial vessels with confocal microscopy on a separate of set of rats (n=10) under the same experimental conditions [142]. The agreement between the simulations and the experimental measurements for both arteries and veins are demonstrated in Fig. 6-2e. Changes in SO_2 across the entire vasculature (Fig. 6-2f) are shown at different time points following the forepaw stimulus in Fig. 6-2g.

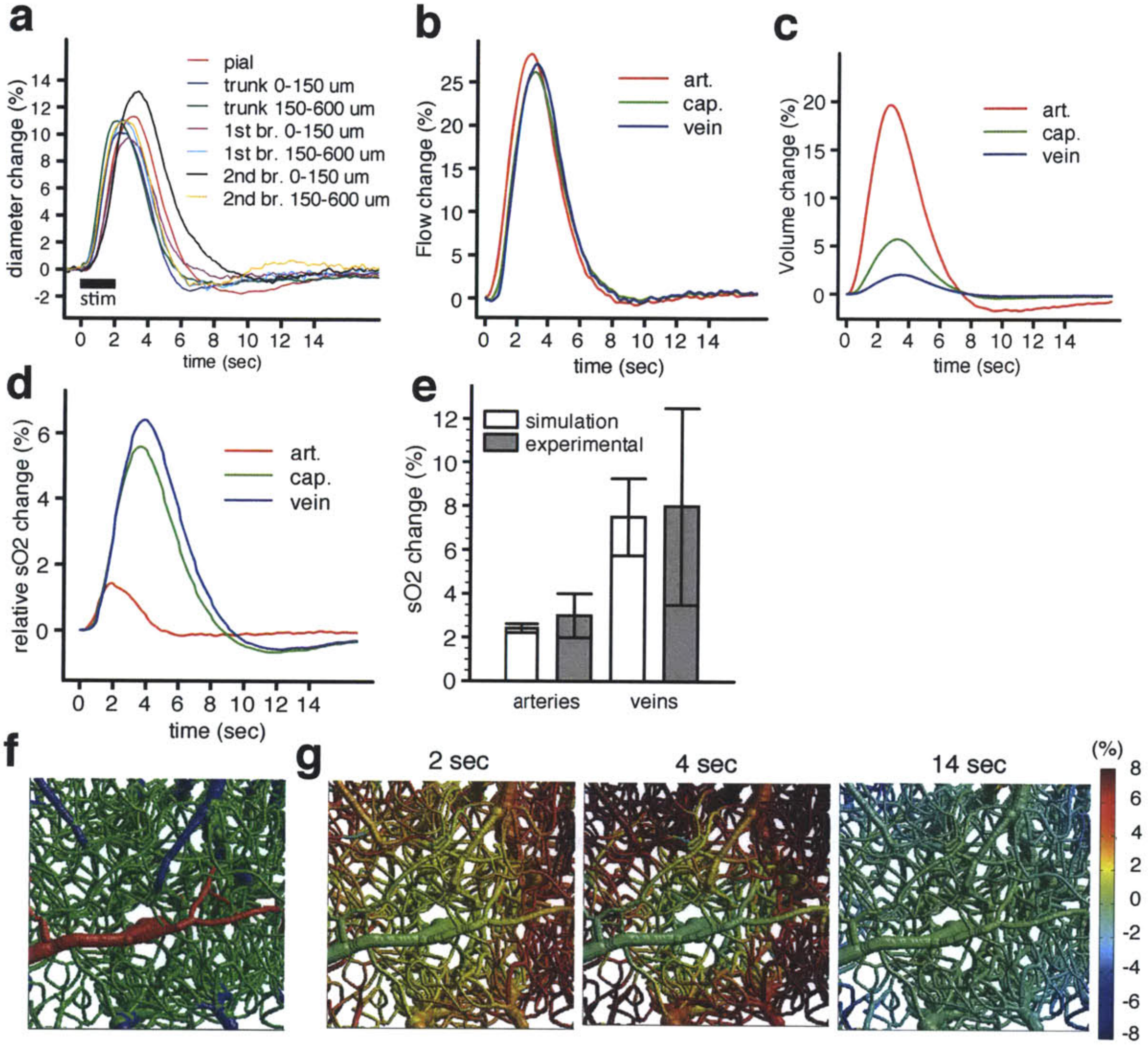

Figure 6-2: Modeling the physiological response to forepaw stimulus. (a) TPM experimental measurements or arterial dilation following forepaw stimulus. (b) Simulated flow changes, (c) simulated volume changes and (d) simulated SO_2 changes (all relative to baseline) in the different vascular compartments. (e) Comparison of simulated SO_2 changes with experimental SO_2 changes measured in pial vessels during a forepaw stimulus with confocal microscopy. (f) Vessel type. (g) Spatio-temporal evolution of simulated SO_2 changes following forepaw stimulus.

6.4 Prediction of the BOLD response to forepaw stimulation using Monte Carlo simulations of proton diffusion within the VAN and validation against experimental data

The BOLD signal is a measure of the transverse magnetization of nuclear spins. In gradient echo (GRE) BOLD, the signal decays to zero due to spin-spin interactions as well as dephasing induced by magnetic field inhomogeneities. In spin echo (SE) BOLD, most of the later process is reversed with the use of a 180 refocusing pulse. The presence of deoxyhemoglobin in the vasculature gives rise to microscopic magnetic fields perturbations within the cortical tissue (upon its introduction in the strong field of the MR scanner) and therefore contributes to local magnetic field inhomogeneities. During functional activation, variations in vessel size and oxygenation level affect the geometry and the amplitude of these magnetic field inhomogeneities and therefore affect the GRE signal. The oxygenation level in the vessels also affects spin-spin coupling and therefore the SE signal.

The magnetic field inhomogeneities were calculated from the SO_2 volumes at each time point using a numerical perturbative method [104, 17]. An example of SO_2 volumes and the resulting magnetic field inhomogeneities are shown in Fig. 6-3a and Fig. 6-3b respectively. BOLD was computed by simulating the random walk of proton nuclear spins within these volumes as shown in Fig. 6-3c and accounting for irreversible spin-spin dephasing (see Supplementary Information). A caveat of this approach is that the perturbative method produces relatively uniform fields inside the vasculature, which is not the case in reality as strong dipole fields arise around red blood cells that are tumbling around. To overcome this difficulty, the intravascular protons were treated separately as described in the Supplementary Information. Both GRE and SE signals were computed by applying different sets of spatial gradients as illustrated in Fig. 6-3d. An example of simulated GRE and SE BOLD are shown in Fig. 6-3e.

Typical features of the BOLD signal can be appreciated including a large overshoot followed by a post-stimulus undershoot. To test the accuracy of our simulations, we performed experimental BOLD measurements on a separate set of rats (n=6) during the same forepaw stimulations [131]. The agreement between the simulated and the experimental BOLD response is demonstrated in Fig. 6-3f where the amplitudes of the signals are compared.

6.5 Contribution of individual compartments for different field strengths

The individual contribution of arteries, capillaries and veins computed over 6 animals and two cortical orientations are shown in Fig. 6-4a . We found that 50-60% of the GRE signal (depending on the cortical orientation) originates from oxygenation changes occurring in the capillaries at 1.5T and that this number increases with field strength to plateau at 65-75% at higher fields. For SE, the inversion pulse refocuses the signal around larger vessels increasing the capillary contribution to 70% at 1.5T. We note that the arterial contribution is negative for GRE as previously reported [135]. The individual contribution of the intravascular and the extravascular compartment were also computed for two cortical orientations and displayed in Fig. 6-4b. For GRE, the intravascular contribution starts at 30-40% at 1.5T and decreases rapidly to become negligible at 7T and higher fields. This result is in good agreement with previous studies [10, 135]. For SE, the intravascular contribution starts higher (70-80%) and persists up to 14T, again, consistent with [135]. Finally, the total BOLD response increases linearly with the field strength for GRE but slightly saturates for SE as shown Fig. 6-4c, which is also consistent with [135].

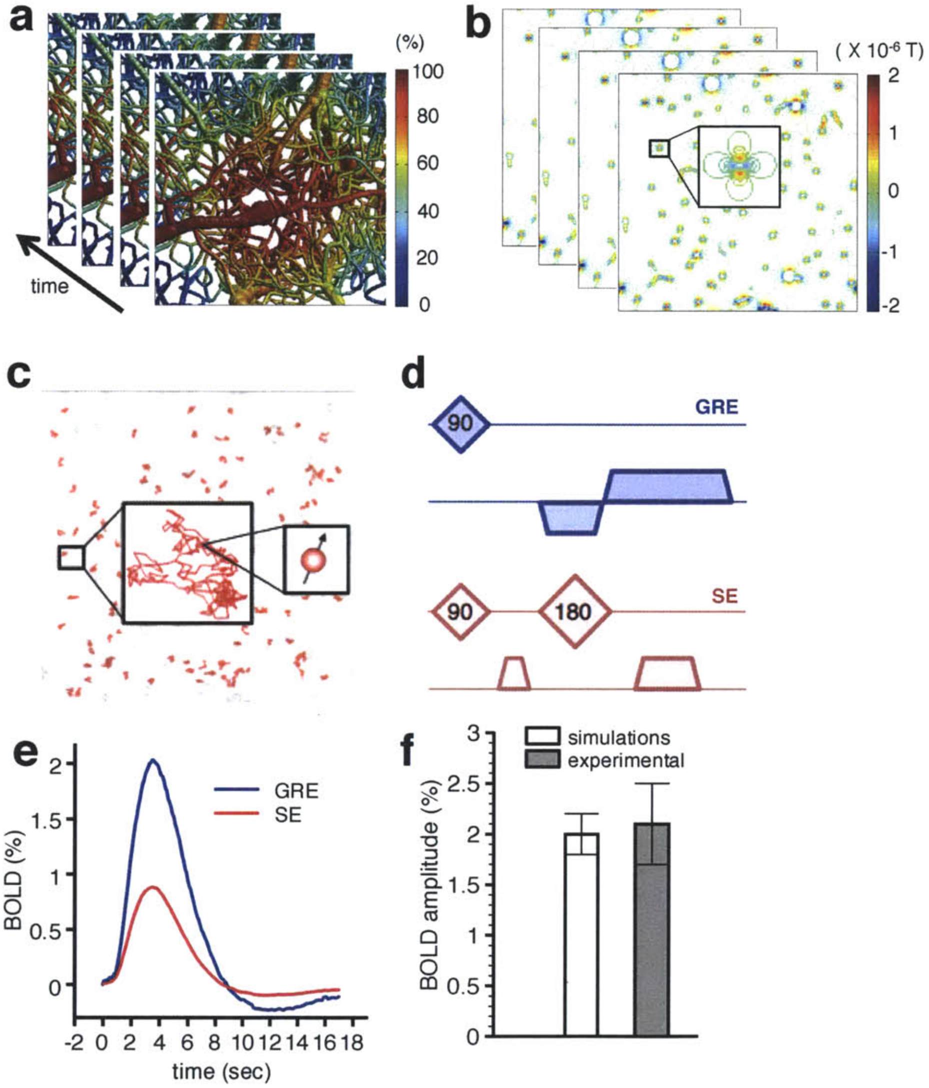

Figure 6-3: Figure 3. Modeling the fMRI signals from realistic vascular networks. (a) Time series of SO_2 volumes. (b) Resulting time series of magnetic field perturbation volumes computed from the SO_2 volumes at each time instant. (c) Simulation of nuclear spins ($n=10^7$) diffusing in the magnetic field perturbation volume. (d) Spatial gradients applied to simulate gradient echo (GRE) and spin echo (SE) signals. (e) Time series of the simulated GRE and SE signals. (f) Comparison of simulated GRE signal against experimental GRE signal measured during the same forepaw stimulus.

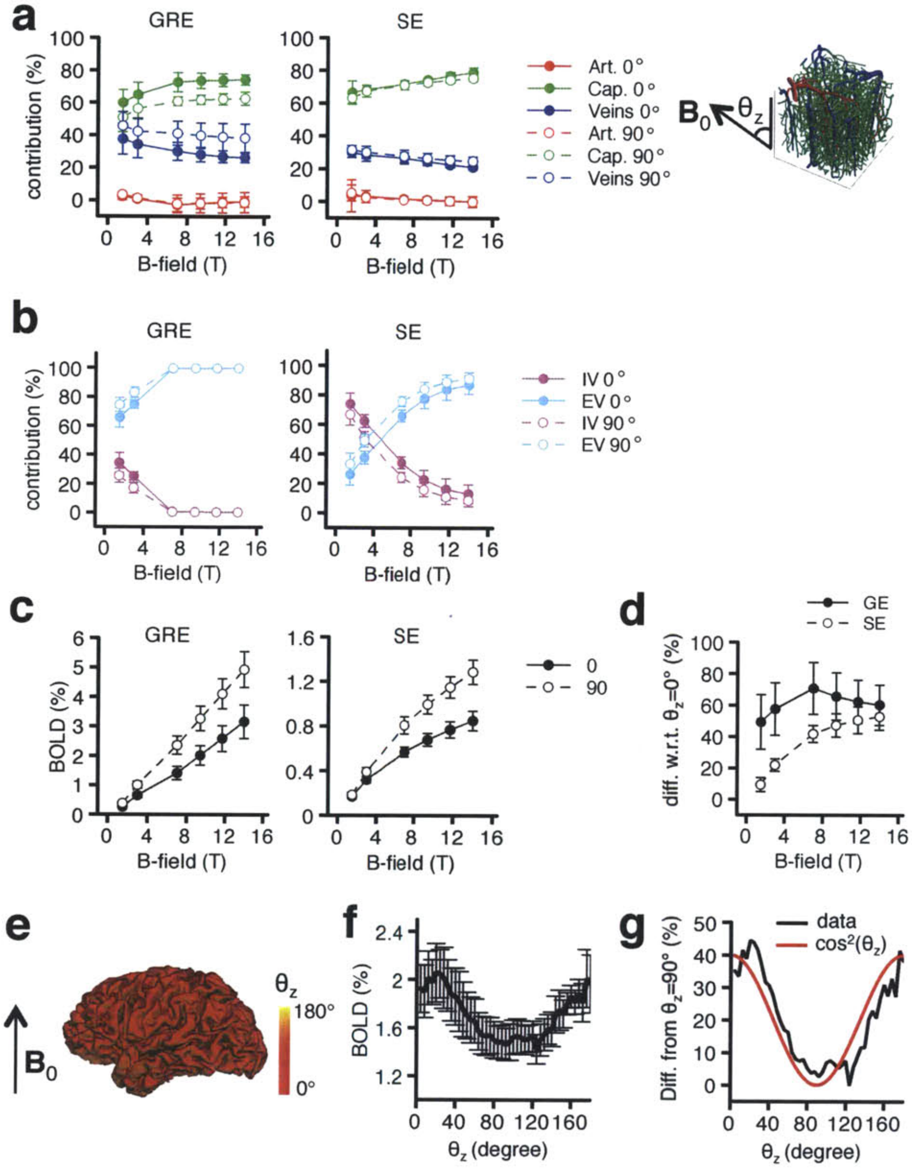

Figure 6-4: Compartment-specific contributions to the BOLD effect for different cortical orientations. a, Contributions of individual vascular compartment computed from our simulations b, Specific contributions of the intravascular and extravascular compartments. c, Total BOLD response simulated for different B-field strengths and for two cortical orientations. d, Difference in the BOLD response obtained for θ_z=90° with respect to θ_z=0° in percent (normalized to θ_z=0°) for different B-field strengths. e, Illustration of θ_z values computed in the gray matter of the human brain. f, Averaged BOLD responses measured in the gray matter of human subjects during a hypercapnic challenge as a function of θ_z. g, Variation of the BOLD response with respect to θ_z=90°. (a-d, n=6 animals, e-f, n=5 human subjects)

6.6 Preferential orientation of veins gives rise to an angular dependence of the BOLD effect

The angle between the cortical surface and the external magnetic field of the scanner (θ_z) is shown to influence the contribution of individual compartments for GRE (Fig. 6-4a) as well as the total BOLD response for both GRE and SE (Fig. 6-4c). The difference in BOLD signal simulated between θ_z=90 and θ_z=0 (as a percent change normalized to θ_z=0) is shown in Fig. 6-4d.

Arteries and veins are mostly oriented either perpendicular or parallel to the cortical surface while the capillary bed has very little preferential orientation and is more uniform (Fig. 6-5a). This observation is consistent with previous studies [5]. All three compartments will exhibit an angular dependence as shown in Fig. 6-5b, but this angular dependence will be stronger for the arterial and venous compartments because of their larger preferential orientation. However, the individual contribution of the arterial compartment to the total BOLD response is weak as shown in Fig. 6-5c, and therefore the descending veins remain the principal cause of the angular dependence of the BOLD effect. Moreover, the angular dependence is exclusively extravascular as shown in Fig. 6-5d since the intravascular BOLD signal is totally isotropic. This is due to the random nature of red blood cell distribution (and magnetic field distortion) inside arteries and veins. In the capillaries, the red blood cells are oriented linearly with the vessels but the distribution of the capillary orientations is mostly uniform as shown in Fig. 6-5a. Accounting for all these facts results in an isotropic intravascular signal globally over an entire voxel.

These results suggest that the same physiological change will produce BOLD responses with different amplitudes depending on the spatial orientation of each specific voxel in the scanner. To confirm this prediction, BOLD-fMRI was measured in humans (n=5) during a hypercapnic challenge, which would be expected to produce a relatively uniform change in deoxygenated hemoglobin over the cortical surface.

The angle θ_z was computed for every voxel as shown in Fig. 6-4e [18]. A scatter plot of the BOLD changes versus θ_z for all voxels of the gray matter was generated and the mean BOLD change for each θ_z is shown in Fig. 6-4f. A variation of 40% in the amplitude of the BOLD response was observed between voxels at 0° compared to voxels at 90° (Fig. 6-4g). We note that the maximum BOLD response in our simulations was obtained for θ_z=90° while in the experimental measurements the maximum was obtained for θ_z=0°. The position of the maximum in human data indicates that surface pial veins contribution is higher in human data while the ascending veins contribution is higher in our simulations. Simulations over larger two-photon stacks containing more surface pial veins should confirm this hypothesis in future studies.

The angular dependence of the BOLD effect produces a confounding effect when comparing BOLD response from different subjects with different brain morphologies or with different spatial orientations of the head in the MRI scanner. The method used to calculate θ_z (Fig. 6-4e) will help to better characterize this phenomenon in human BOLD data and to develop methods to correct for this confounding effect in future studies.

6.7 Supplementary Information

6.7.1 Baseline measurements of pO_2 and angiography

All experimental procedures were approved by the Massachusetts General Hospital Subcommittee on Research Animal Care. We anesthetized C57BL/6 mice (male, 2530 g, n=6) by isoflurane (12% in a mixture of O_2 and air) under constant temperature (37°C). A cranial window with the dura removed was sealed with a 150-m-thick microscope coverslip. During the experiments, we used a catheter in the femoral artery to monitor the systemic blood pressure and blood gases and to administer the two-photon dyes. During the measurement period, mice breathed a mixture of O_2,

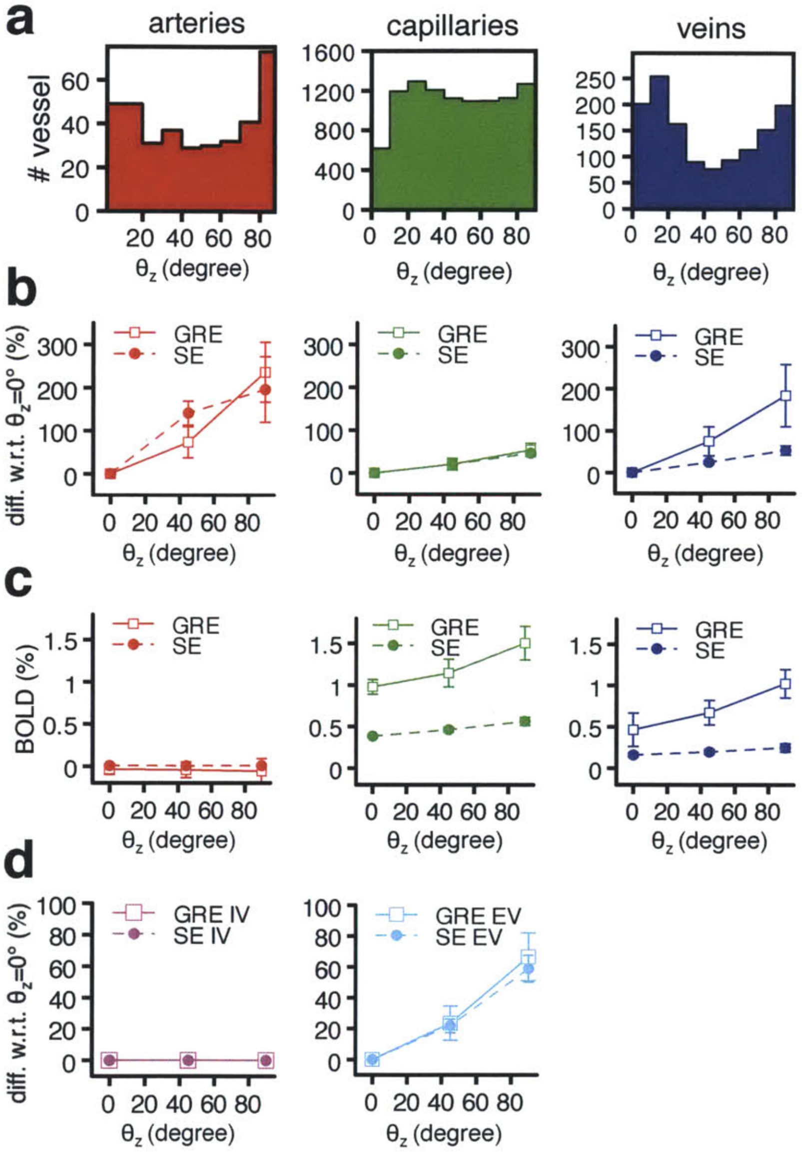

Figure 6-5: Contribution of different vascular compartments to the angular dependence of the BOLD effect. (a) Histogram of vessel orientation with respect to the z-axis for arteries, capillaries and veins. (b) Normalized BOLD response with respect to θ_z=0°. (c) BOLD responses from specific vascular compartments for different cortical orientations. (d) Intravascular and Extravascular BOLD responses normalized with respect to θ_z=0° for different cortical orientations.

air, and CO_2 under the same isoflurane anesthesia. Imaging was performed using a custom built two-photon microscope (Sakadi et al. 2010) and two-photon enhanced oxygen sensitive phosphorescent dye PtP-C343 [37]. The time-domain measurements of phosphorescence lifetimes were performed following the procedures outlined in [119, 25, 103]. Approximately 400 pO_2 measurements were collected in various microvascular segments down to 450 m from the cortical surface. The conversion between pO_2 and oxygen saturation of hemoglobin (SO_2) was performed using the Hill equation with Hill coefficients specific for C57BL/6 mice ($h = 2.59$ and $P50 = 40.2$) (Uchida et al. 1998). After collecting the pO_2 measurements, we obtained structural images of the cortical vasculature by labeling the blood plasma with dextran-conjugated fluorescein (FITC) at 500nM concentration. We acquired 600 x 600 x 662 um stacks of the vasculature with 1.2 x 1.2 x 2.0 um voxel sizes under a 20X Olympus objective (NA=0.95).

6.7.2 Functional measurements on rodents

Rationale for using both rats and mice

All baseline measurements were performed on mice while all functional measurements were performed on rats. This was motivated by recent works performing a detailed topological analysis of the cortical microvasculature of rodents and suggesting that the topology of cortical vessels is very similar for mice and rats [4]. Therefore, our analysis should lead to very similar results and conclusions whether we used rats or mice. However, from a technical perspective, using mice for baseline measurements allowed us to take advantage of the powerful two-photon pO_2 measurement technology [119] while using rats for functional data resulted in higher signal-to-noise ratio for the parameters measured (especially arterial dilation and BOLD-fMRI).

Stimulus

All experimental procedures were approved by the University of California at San Diego Institutional Animal Care and Use Committee. Sprague-Dawley rats (130200 g) were anesthetized as described in [27, 26]. The stimulation lasted 2 s and consisted of a train of six electrical pulses (3 Hz, 300 s, 1 mA) with an interstimulus interval of 2025 s. The intensity was adjusted to provide stimulation below the movement threshold. Stimulation was presented using a separate computer that also acquired transistortransistor logic (TTL) timing signals for data acquisition (trigger out TTLs for each line or frame during two-photon acquisition and for each slice during fMRI) using a National Instruments IO DAQ interface controlled by a home-written software in Matlab. The TTL data were used to determine the timing of each line/frame/ slice relative to the stimulus onset during data analysis performed in Matlab.

Two-photon measurements of arterial dilation

Two-photon microscopy was performed on rats (n=19) as described in [26]. Fluorescein-conjugated dextran (FD-2000; Sigma) in physiological saline was injected i.v [96]. Images were obtained with an Ultima two-photon microscopy system from Prairie Technologies using 4 (Olympus XLFluor4/340, NA = 0.28) and 40 (Olympus, NA = 0.8) objectives. Line scans up to 1 mm long were acquired across multiple vessels (up to six) with a scan rate of 80170 Hz. The scan resolution was 0.5 m or less. Diving arterioles were measured in the frame mode at five to eight frames/s.

Confocal measurements of pO_2

Confocal microscopy was performed on rats (n=10) as described in [142]. A solution of Oxyphor R2 in saline was administered through the femoral vein to yield a concentration of 40 mmol/L in the bloodstream. At each location, phosphorescence was excited for 100 ms. The resultant phosphorescence emission decay profile was collected at 50MHz sampling rate for 500ms. Fifty decay profiles were averaged for each measurement (30ms per point measurement). The decay lifetime was calculated and converted to pO_2 [120, 143]. The temporal resolution was 0.5 to 1 sec depending on the number of points measured per interval.

Functional MRI

MRI was performed on rats (n=10) on a 7T/21-cm BioSpec 70/30 USR horizontal bore scanner (Bruker) as described in [131]. BOLD functional data were acquired using a single-shot gradient-echo echo planar imaging (EPI) pulse sequence with the following parameters: TE = 10 ms, flip angle = 30, matrix = 80 80, slice thickness = 1 mm, TR = 1 s, five adjacent slices. The laminar analysis procedure used was previously described in [131].

Graphing the angiograms

In order to estimate vessel diameters, to label vessel types and to compute statistics across the angiogram such as branching order from pial vessels, a mathematical representation of the vasculature must be obtained. This mathematical representation in termed a graph, which consists of nodes interconnected by segments. Structural images based on FITC-labeled blood plasma were used to construct a graph of the microvascular network for each animal. A 3x3x3 median filter was used to enhance vessel contrast. We created the graphs and performed image processing using a suite of custom-designed tools in Matlab (MathWorks, Inc). Initial steps involved running the VIDA suite [134]. The graphs were then inspected and manual corrections were applied until all segments in the field of view become interconnected (single group). Vessel diameter was estimated at each graph node by thresholding the image at a low value of approximately 2% of the maximum image intensity, considering lines through the node point oriented every 3 degrees in the local plane perpendicular to the vessel axis, and finding the minimum distance from vessel edge to vessel edge [35].

Vessel Type identification

Capillary labeling was done automatically by setting all segments with diameter equal or less than 8 um as capillaries. Arteries were then manually identified as the upstream segments and veins as the downstream segments. Surface pial arteries tend to be straighter and can be easily distinguish from surface pial veins, which are more curvy. The pO_2 measurements were used in cases of ambiguity. The Floyd-Warshall algorithm was then used to calculate branching orders of individual vascular segments with respect to main pial vessels, which were manually identified.

6.7.3 Vascular Anatomical Network model

Rationale

The complete set of measurements required to acquire the angiogram and the pO_2 distribution at rest takes about 45 minutes. It is therefore not possible to measure all vessel sizes and the entire pO_2 distribution at every time point during functional activation with short stimulus, which are the typical stimulus length used in human event-related fMRI experiments. To model the BOLD signal accurately, we would like to reconstruct these changes in vessel size and oxygenation with a temporal resolution of 1 sec, which is 3 orders of magnitude too fast for the actual technology. To overcome this limitation, we rather focused on measuring a single parameter at a time (i.e. vessel diameters or pO_2 changes) during functional activation. Together with a vascular anatomical network (VAN) model, these functional measurements allow to reconstruct oxygenation changes everywhere in the graph with very good temporal resolution (0.1 sec), which is not achievable with the actual microscope technology alone.

Overview

Vessel dilation in the arterial compartment is an active process mediated by the release of vasodilatory agents and therefore these measurements were used as inputs to perturb the VAN model from steady-state during functional activation. Flow changes and volume changes can then be computed everywhere assuming a passive compliance model for the capillary and the venous bed. Knowing flow and volume changes allows to compute the changes in oxygenation in all the vessels assuming a ΔFlow/ΔCMRO2 coupling ratio. To ensure that this gives accurate and realistic results, the simulated oxygenation changes are then compared with partial pO_2 measurements during functional activation. For all six vascular networks constructed, this approach gave really good agreements between the simulated and the measured functional changes as shown in Fig. 6-2e.

Steady-state VAN

The goal here is to reconstruct the resting distribution of oxygen in all vessels. This distribution will then be compared with the pO_2 distribution measured experimentally to make sure it is realistic. This steady-state distribution will then be perturbed during functional activation.

The oxygenation level in the vasculature is globally determined by two competing parameters, which are blood flow and the cerebral metabolic rate of oxygen (CMRO2). Higher blood flow increases oxygenation while higher CMRO2 decreases it. In steady-state, these two parameters are related by:

$$CMRO_2 = CBF \times OEF \times C_a \tag{6.1}$$

where OEF is the oxygen extraction fraction and C_a is the arterial blood oxygen content given by

$$C_a = p_aO_2 \cdot \alpha + 4 \cdot Hct \cdot C_{Hb} \cdot S_aO_2 \tag{6.2}$$

where $alpha = 1.27x10\char`~3$ umol/mL/mmHg is the solubility of oxygen, $C_{Hb} = 5.3$ umol/mL is the hemoglobin content of blood and $Hct = 0.4$ is the hematocrit in arteries. OEF can be computed directly for each animal using our two-photon measurements:

$$OEF = \frac{S_aO_2 - S_vO_2}{S_aO_2} \tag{6.3}$$

CBF in rodents is well documented in the literature [149, 140]. We therefore fixed CBF to obtain a perfusion of 100 mL/100g/min in our volumes. CMRO2 was then computed for each animal using the above equation and values obtained for each animal are shown below:

Table 6.1: Physiological parameters measured in mice

ID	p_aO_2 (mmHg)	p_vO_2 (mmHg)	S_aO_2 (%)	S_vO_2 (%)	OEF (%)	CMRO2 umol/ml/min)
20100203	109	44	92	58	36.9	2.28
20110408	94	43	88	54	38.6	2.32
20120404	107	60	93	67	28.0	1.72
20120409	116	57	93	66	29.0	1.81
20120625	118	60	94	68	27.7	1.73
20120626	109	41	91	58	36.2	2.23
Mean	108.8	50.8	91.8	61.8	32.7	2.02
Std	8.5	9.1	2.1	5.9	5.0	0.26

Capillary segments cut by the limits of the field-of-view were removed to obtain a closed graph between the 1-2 pial arteries and the 4-6 pial veins. This procedure was previously used by [90, 91] and was shown to results in accurate flow distributions. BOLD simulations were performed on regions of interest (Fig. 6-1a) excluding regions of tissue were capillary end-segments have been removed. The resistance for each segment was calculated using Poiseuilles law corrected for hematocrit as described in [108]. Flow speeds in inflowing pial arteries were calculated based on the perfusion assumed (100 mL/100g/min) and the arterial diameters. Blood pressure boundary conditions for pial veins were set using values from [87] and the blood flow distribution was finally computed using the matrix equations given in [8] together with velocity boundary conditions for inflowing arteries and the blood pressure boundary conditions for outflowing veins.

A tridimensional mask of the vasculature was obtained from the angiogram and the graph and a mesh of the vasculature was generated using the software iso2mesh [34]. Oxygen advection was then performed individually for each animal using the computed blood flow distribution and the inflowing arterial pO_2 given in the table above for each animal. The pO_2 was initialized everywhere to 10 mmHg and oxygen advection was run with constant inputs until steady-state was achieved (typically after 15 sec in the model time). The details of the finite elements algorithm used can be found in our previous paper [35].

VAN model during functional activation

The arterial dilation traces shown in Fig. 6-2a were used as inputs to compute changes in blood flow and blood volume as described in [8]. An intra-cranial pressure of 10 mmHg was assumed and the compliance parameter β was set to 1 for both capillaries and veins [8]. Oxygenation changes during functional activation was then computed using the same advection code [35] by keeping pO_2 in the arterial inflowing nodes constant and using the updated flow and volume values at each time point. CMRO2 was increased following a temporal profile corresponding the averaged arterial dilation trace with a peak amplitude corresponding to a relative change 3 times lower compared to the relative change in blood flow, giving a ΔFlow/ΔCMRO2 coupling ratio of 3 [68, 32].

6.7.4 fMRI simulations

Overview

The BOLD signal is a measure of the transverse magnetization of nuclear spins. In gradient echo (GRE) BOLD, two processes contributes to the decay of the signal: dipole-dipole coupling (spin-spin interactions) as well as dephasing induced by the inhomogeneities in the local magnetic field. The relaxation constant embedding these two processes is termed T_2^*. In spin echo (SE) BOLD, the second process is reversed around larger vessels (veins) with the use of a 180 refocusing pulse. The relaxation constant in this case is termed T_2. A contributor to the local magnetic field inhomogeneities is the presence of deoxyhemoglobin in the vasculature, which is paramagnetic. During functional activation, variations in vessel size and oxygenation level affect the geometry and the amplitude of these magnetic field inhomogeneities and therefore affect T_2^*. Furthermore, the oxygenation level affects spin-spin coupling and therefore T_2.

The challenges in modeling BOLD are (1) to compute the magnetic field inhomogeneities at every time point and (2) to keep track of spin-spin decay (T_2 effect). These tasks both require to know exactly the microvascular geometry and the deoxyhemoglobin content in each vessel segments at every time point.

Computing magnetic field inhomogeneities

We used a numerical method previously described by [104] to compute the magnetic field inhomogeneities. The SO_2 volumes (like the one shown in Fig. 6-3a were resampled to 1x1x1 um and converted to a susceptibility shift volume $\Delta\chi$ using:

$$\Delta\chi = \Delta\chi_0 \cdot Hct \cdot (1 - SO_2) \tag{6.4}$$

where $\Delta\chi_0 = 4 \cdot 0.264x10^{-6}$ is the susceptibility difference between fully oxygenated and fully deoxygenated hemoglobin [17] and Hct is the hematocrit that was assumed to be 0.3 is capillaries and 0.4 in arteries and veins [52]. Assuming that the magnetic field inhomogeneities are small, the method uses perturbation theory and the inhomogeneities across the entire volume are computed by convolving the susceptibility shift volume $\Delta\chi$ with the geometrical factor for the magnetic field inhomogeneity induced by a unit cube:

$$\Delta B_{cube} = \frac{6}{\pi} \cdot \frac{1}{3}\frac{a^3}{r^3}\left(3cos^2\theta - 1\right) \cdot B_0 \tag{6.5}$$

This procedure allowed to compute the magnetic field inhomogeneities across the entire vascular volume $\Delta B_{inhom}(x)$Fig. 6-3b.

T_2 and T_2^* volumes

T_2 and T_2^* volumes are required on top on magnetic field inhomogeneity volumes to accurately model the fMRI signals. T_2 and T_2^* values (in sec) along the vasculature were computed using the formulas obtained by fitting experimental measurements and given in [135]

$$T_{2,vessel} = \left(12.67 \cdot B_0^2 \cdot (1 - SO_2)^2 + 2.74B_0 - 0.6\right)^{-1} \tag{6.6}$$

$$T^*_{2,vessel} = \left(A + C(1 - SO_2)^2\right)^{-1} \tag{6.7}$$

where A and C are constants (in sec^{-1}) which depend on the external magnetic field B_0 and given below.

Table 6.2: Constants for T_2 and T_2^* in the vasculature

$B0 \leq 1.5T$	A=6.5	C=25
$1.5T < B0 \leq 3T$	A=13.8	C=181
$3T < B0 \leq 4T$	A=30.4	C=262
$4T < B0 \leq 4.7T$	A=41	C=319
$B0 > 4.7T$	A=100	C=500

In the tissue (outside the vessels), T_2 and T_2^* (in sec) were computed using the formulas given in [135]

$$T_{2,tissue} = (1.74 \cdot B_0 + 7.77)^{-1} \tag{6.8}$$

$$T^*_{2,tissue} = (3.74 \cdot B_0 + 9.77)^{-1} \tag{6.9}$$

Monte Carlo simulation of nuclear spins

Water protons experience diffusion in cortical tissue and this process was simulated with Monte Carlo simulations. The positions of 107 protons were initialized uniformly in the three-dimensional volume. Each proton experienced a random walk for a period of TE sec as shown in Fig. 6-3c. The diffusion coefficient was set to 1x10^{-5} cm^2/sec [104] and the time step dt was set to 0.2x10^{-3} sec. At each time step, the position $x = (x_1 x_2 x_3)$ of each proton was updated using

$$x_1' = x_1 + N(0, 2 \cdot D \cdot dt) \tag{6.10}$$

$$x_2' = x_2 + N(0, 2 \cdot D \cdot dt) \tag{6.11}$$

$$x_3' = x_3 + N(0, 2 \cdot D \cdot dt) \tag{6.12}$$

Protons reaching a vessel wall were bounced-back such that all intravascular protons stayed inside the vessels for the duration of the simulation and vice versa for the extravascular protons. The MR signal was computed at each time step by averaging the contribution of all N protons:

$$S(t) = \mathrm{Re}\left(\frac{1}{N} \cdot \sum_{i=1}^{N} \exp(\phi_n(t))\right) \tag{6.13}$$

where the phase was updated every time step using:

$$\phi_{n,intra}(t) = \sum_{k=1}^{t/dt} -T^*_{2,blood}(x(k)) \tag{6.14}$$

$$\phi_{n,extra}(t) = \sum_{k=1}^{t/dt} \gamma \cdot j \cdot \Delta B(x(k)) - T_{2,tissue}(x(k)) \tag{6.15}$$

where $\Delta B = \Delta B_{inhom} + \Delta B_{gradients}$ with ΔB_{inhom} the magnetic field homogeneity computed above and $\Delta B_{gradients}$ is the field homogeneity introduced by the spatial gradient shown in Fig. 6-3d. For spin echo (SE) signal, the imaginary part of the phase was inverted at TE/2

$$\phi_n\,(\mathrm{TE}/2) = \mathrm{conj}\,(\phi_n\,(\mathrm{TE}/2)) \qquad (6.16)$$

We note that the intravascular and the extravascular protons were treated differently. This method was adopted since the numerical method produces relatively uniform magnetic fields inside the vasculature, which is definitively not the case in reality where very strong dipole fields arise around red blood cells that are tumbling around and water molecules are exchanged between red blood cells and the plasma. There is no current microscopic way of modeling accurately the intravascular signal so we relied on empirical formulas based on experimental measurements to model the intravascular portion of the signal.

This procedure is repeated at each desired time point during the functional activation. The relative signal changes was computed by comparing the signal obtained at each time point to the signal obtained at t=0 and converted to a percent change.

6.7.5 Comparison of simulations against experimental BOLD

For the simulations traces presented in Fig. 6-3e, the direction of the external magnetic field B_0 (7T) was set perpendicular (90°) to the z-axis. This matches the B_0 amplitude and direction of our experimental BOLD data which were collected with a 7T horizontal bore magnet which leads to an angle of 90° between B_0 and the cortical surface of the forepaw area. The experimental data presented in Fig. 6-3f are averaged over the first 600 um of the cortex (first three slices as described in [131] to match the volume sampled in our Monte Carlo simulations (layers I-IV).

6.7.6 Individual contributions to the BOLD signal

The individual contributions to the BOLD signal shown in Fig. 6-4a were computed at the peak of the activation, which occurred between 3.5 and 4 sec (depending on the animal) after the start of the stimulus. For each field strength, TE was set to $T^*_{2,tissue}$ for GRE and to $T_{2,tissue}$ for SE.

Arteries, capillaries and veins contributions

To compute the contribution of an individual vascular compartment (e.g. the capillaries), two different simulations were performed. In the first one, the oxygenation volume (with dilated vessels) computed with the VAN at the peak of the functional activation was used. The signal obtained was compared to the signal obtained using the oxygenation volume (with baseline vessel size) at t=0. This will be referred as the total BOLD response. In the second simulation, we constructed a new volume by using baseline (values at t=0) oxygenation and vessel size for arteries and veins, but peak values for oxygenation and vessel sizes for capillaries. The signal obtained in this case was also compared to the same baseline signal computed using baseline oxygenation and vessel size everywhere. This signal change will be referred to the capillary BOLD signal change. To compute the individual contribution of the capillaries, the capillary BOLD signal change was divided by the total BOLD signal change and converting to percentage. This procedure was repeated for arteries and veins, and for all field strengths.

Intravascular and extravascular contributions

From a practical point of view, the extravascular BOLD signal is the signal one would get by applying strong crusher gradients to null the signal of fast moving protons inside blood vessels [10]. We used this definition for our simulations. Again, two simulations were performed. The first one is identical to the one described previously and termed total BOLD signal change. In the second one, we simulated only extravascular protons both at the peak activation and at baseline (t=0). The second one is referred to the extravascular BOLD signal change. The extravascular contribution was computed by dividing the extravascular BOLD signal change by the total BOLD signal change and converting to percentage. The intravascular contribution was then computed by subtracting the extravascular contribution in percentage from 100%. This way of calculating the intravascular and extravascular contributions takes into account both blood volume and blood oxygenation and is consistent with the practical definition given above.

Calculation of the angular dependence of BOLD

To quantify the angular dependence of the BOLD response, we compared the BOLD signal changes simulated at different θ_z values with the BOLD signal changes simulated at $\theta_z = 0°$. The difference was converted to a percentage with respect to $\theta_z = 0°$:

$$\mathrm{diff}\left(\theta_z\right) = 100 \times \frac{\Delta S_{\theta_z} - \Delta S_{\theta_z=0}}{\Delta S_{\theta_z=0}} \tag{6.17}$$

With this definition, a variation of 60% indicates that the BOLD signal change for $\theta_z = 90°$ is 60% stronger compared to the BOLD signal change for $\theta_z = 0°$.

6.7.7 Experimental BOLD measurements on human

BOLD during Hypercapnia

All experimental procedures were approved by the Massachusetts General Hospital. Healthy subjects (n=5) were enrolled in the study. Written consent was received from each subject before the experiment. Each subject received 2 blocks of 2-min hypercapnia during which his or her end-tidal pCO2 increased by 8 mmHg with respect to their baseline pCO2 value. The hypercapnic blocks were interleaved with

3-min blocks at normocapnia. Combined ASL-BOLD data were collected during the gas manipulations. Only the control images were used in the BOLD analysis. Sequence parameters were TR=3000ms, IR-1=600ms, IR-2=1800ms, TE=13ms, flip angle=90, Res=3.4x3.4x6.0 mm, 6 slices. An anatomical T1-weighted scan (MPRage) was also collected (Res=1x1x1.2 mm).

Angular analysis

BOLD data were analyzed using Freesurfer. Motion correction and slice-timing correction were applied. No smoothing was used. BOLD signal changes between normocapnic and hypercapnic conditions were computed across the 6 slices. A complete cortical surface reconstruction of the anatomical scan was performed with Freesurfer using the recon-all function. An additional cortical surface was generated mid-way in the gray matter and the angle between the normal to this surface and B_0 was computed as previously described in [18]. The BOLD signal volumes were then interpolated on this surface, leading to a series of voxels containing both θ_z and BOLD change values. The data were pruned by selecting only voxel with a positive BOLD response.

The pruned data point were binned based on θ_z at every four degrees between 0 and 180° and the average BOLD change for each bin was computed. The variation in BOLD changes with respect to BOLD change at $\theta_z = 90°$ was computed using

$$\mathrm{diff}\,(\theta_z) = 100 \times \frac{\Delta S_{\theta_z} - \Delta S_{\theta_z=90}}{\Delta S_{\theta_z=90}}. \tag{6.18}$$

6.8 Summary

In conclusion, the modeling approach developed here permits study of the BOLD effect in its finest details. This framework will help to validate and refine simplified

BOLD models used in the calibrated BOLD approaches [22], or any other quantitative T_2^* fMRI method to infer physiological parameters. Finally, this framework has great potential in the new field of MRI fingerprinting [92].

Chapter 7

Multimodal reconstruction of cerebral blood flow using combined two-photon microscopy and optical coherence tomography

The main contribution of this chapter is to investigate whether adding flow constraints to vessel segments can help in computing cerebral blood flow (CBF) in incomplete microvasculature networks. Simulations were performed on a synthetic network and showed that flow constraints reduce the impact of imperfect pressure boundary conditions and vessel resistance resulting in a more robust reconstruction of flow. We then combine two-photon laser scanning microscopy (TPLSM) angiography with Doppler optical coherence tomography (DOCT) to apply our method for calculating microvascular blood flow within the rodent brain.

7.1 Introduction

There is increasing interest in modeling blood flow from the perspectives of neuroimaging [16, 60] and cerebrovascular physiology [75]. Several groups have performed the computation of cerebral blood flow in realistic vascular networks [108, 35, 8, 114, 2, 53, 110, 100, 90, 91, 121]. Several difficulties arise when the computation takes place over a real vascular network measured *in vivo*. This includes vessel diameter estimation and therefore estimation of the resistance, as well as red blood cell (RBC) rheology. Moreover, the truncation of the network at the edges of the imaging volume inevitably increases the number of boundary conditions required for the computation of flow[114, 90]. All these drawbacks increase the uncertainty of the flow distribution computed with these models and surely limit their application.

Doppler Optical Coherence Tomography (DOCT) applied to the rodent cerebral microvasculature has recently been demonstrated [125] and validated [124] to evaluate flow in a small number of vessels. Such measurements could in principle help the reconstruction of flow in truncated microvascular networks.

Using simulations, we investigate the error reduction obtained by using measurements to constrain the reconstruction of flow within the vascular network. We then apply our method to *in vivo* data by combining DOCT for quantitative measurement of CBF in specific vessels and two-photon laser scanning microscopy (TPLSM) for precise measurement of the three-dimensional vessel morphology.

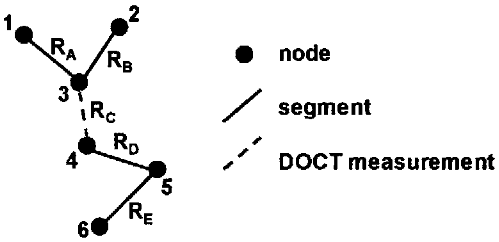

Figure 7-1: Graph representation of a vessel network.

7.2 Theory

We represent a vessel network as a graph consisting of nodes joined by segments as illustrated in Fig. 7-1. The flow in a given segment of length ℓ and diameter d is computed according to Poiseuille's law

$$F = \frac{128\eta\ell\Delta P}{\pi d^4} \tag{7.1}$$

where the term $128\eta\ell/\pi d^4$ is the resistance R of the segment and ΔP is the pressure gradient between the two nodes of the segment. η is the viscosity of blood and was set to 15 x 10^{-6} mmHg $\cdot$ s in our work [107, 87]. More complex models described in the literature [107, 109, 114] could be used similarly. Conservation of flow at branching nodes implies that

$$F_A + F_B + F_C = 0 \tag{7.2}$$

$$F_C + F_D = 0 \tag{7.3}$$

$$F_D + F_E = 0. \tag{7.4}$$

If we first ignore the DOCT measurement in segment C, we substitute Eq. (7.1) into Eqs. (7.2-7.4) to get

$$\frac{P_3 - P_1}{R_A} + \frac{P_3 - P_2}{R_B} + \frac{P_3 - P_4}{R_C} = 0 \tag{7.5}$$

$$\frac{P_4 - P_3}{R_C} + \frac{P_4 - P_5}{R_D} = 0 \tag{7.6}$$

$$\frac{P_5 - P_4}{R_D} + \frac{P_5 - P_6}{R_E} = 0. \tag{7.7}$$

On top on flow conservation, boundary conditions for blood pressure need to be established at the graph end points. In this work, these values were taken from literature [87] based on the diameter of the vessel. For the simple graph of Fig. 7-1, one gets

$$P_1 = P_1^{lit} \tag{7.8}$$

$$P_2 = P_2^{lit} \tag{7.9}$$

$$P_6 = P_6^{lit}. \tag{7.10}$$

Eqs. (7.5-7.10) can be written in matrix form

$$\mathbf{M_R} \cdot \mathbf{P} = \mathbf{y_R} \tag{7.11}$$

and the pressure for all nodes is solved simultaneously using a least-squared procedure.

If we now take into account the DOCT measurement in segment C (labeled F_C^{OCT}) and use this value as a hard constraint, we get

$$\frac{P_3 - P_1}{R_A} + \frac{P_3 - P_2}{R_B} = F_C^{DOCT} \tag{7.12}$$

$$\frac{P_4 - P_3}{R_C} = -F_C^{DOCT} \tag{7.13}$$

$$\frac{P_5 - P_4}{R_D} + \frac{P_5 - P_6}{R_E} = 0 \tag{7.14}$$

which together with Eqs. (7.8-7.10) can be written in matrix form

$$\mathbf{M_C} \cdot \mathbf{P} = \mathbf{y_C} \tag{7.15}$$

where the subscript **C** stands for constrained. In this case, neither the resistance nor the length of the segments containing a DOCT measurement is taken into account in the computation of flow.

7.3 Material and methods

7.3.1 Animal preparation

Animal protocols were approved by the Massachusetts General Hospital Subcommittee on Research Animal Care. We anesthetized C57BL/6 mice (male, 25-30 g) with isoflurane (1-2% in a mixture of O_2 and air) under constant temperature (37°C). We then opened a cranial window with dura removed and sealed it with a 150-μm-thick microscope coverslip. Blood plasma was labeled with fluorescein isothiocyanate-conjugated dextran (FITC) at 500 nM concentration. Physiological parameters such as blood pressure, heart rate and pCO_2 were monitored and maintained within the normal ranges.

7.3.2 Multimodal microscopy

Our multimodal microscope, described in Refs. [142, 119], has independent TPLSM and OCT scanning arms which share a common imaging objective, allowing easy switching between sequential TPLSM and OCT measurements.

We first performed a DOCT scan protocol (with a 856 nm laser source) for quantitative flow measurements. The DOCT scan generated a three-dimensional map of the axial (z) projection of velocity, $v_z\left(x, y, z\right)$ [125]. Flow at specific locations in vessels was obtained from DOCT velocity axial projections by calculating the velocity flux through the en face (also known as transverse or xy) plane

$$F = \int\int_{\mathrm{ROI}} v_z\left(x, y, z_0\right) dxdy. \tag{7.16}$$

The X-Y resolution of our OCT was 10-15 μm which did not permit accurate flow estimation in capillaries . Moreover, the system could only accurately measure velocity projections in the range ±3.6 mm/s before producing phase wrappings, and thus the higher velocities in large arteries could not be realiably measured. Because of these two limitations, DOCT measurements were only considered in larger veins in this work.

We then acquired a 600x600x660μm stack of the FITC-labeled vasculature using TPLSM (1.17x1.17x 2.0μm voxel size) as described in ref [119]. Although OCT can also be used to generate angiograms, the spatial resolution and contrast-to-noise ratio is generally higher with FITC TPLSM, which makes it easier to graph (i.e. vectorize) the vascular stack in the subsequent steps.

7.3.3 Data processing

The TPLSM angiogram was graphed with a custom software using a combination of automatic segmentation procedures and manual corrections [35]. Vessel diameter was estimated at each graph node by thresholding the image at a low value of approximately 2% of the maximum image intensity, considering lines through the node point oriented every 3 degrees in the local plane perpendicular to the vessel axis, and finding the minimum distance from vessel edge to vessel edge. We set the blood pressure boundary condition for the arteries and venules crossing the sides of the imaged volume using tabulated values from ref [87] given the vessel diameter and the vessel type (artery or vein). This set of boundary conditions have been shown to result in accurate and realistic flow distributions in previous studies [35, 90].

7.4 Results and discussion

7.4.1 Simulations

We investigated the improvement obtained by adding DOCT measurements to the recovery procedure using a synthetic vessel network. The diameter of the vessels were fixed to 8 μm and flow was first computed given a set of boundary conditions at the end nodes. Noise was then introduced by varying randomly the values for the pressure boundary conditions and the resistance of the vessels. Flow was then recovered using the erroneous pressure boundary conditions and resistance values without DOCT priors using Eq. (7.11). A fraction of the vessels were then randomly selected and DOCT priors were set for these vessels. Flow was then reconstructed using these priors from Eq. (7.15). The fraction of vessels with DOCT priors ranged from 5 to 25 %. The value for the DOCT measurements was set to the true (simulated) flow value with noise added, ranging from 5 to 25%. The performance was quantified by computing the mean squared error (MSE) between the simulated and the recovered flow in both cases (i.e. with and without DOCT priors). To insure reproducible results, the simulations were run 30 times with different noise instances for boundary conditions and resistance values and MSE values were averaged across all simulations.

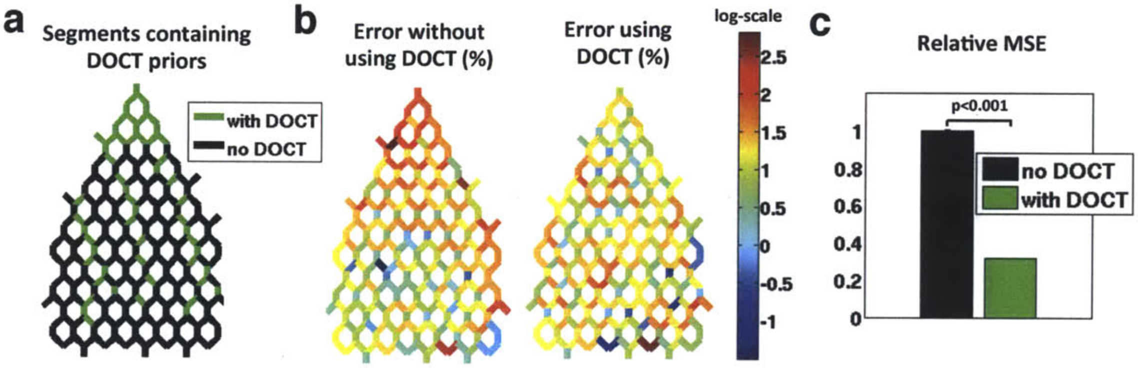

Figure 7-2: Simulations showing the improvement in accuracy obtained by integrating DOCT measurements in the computation of flow. The fraction of vessel with DOCT was 20% and noise in the DOCT measurement was 25%. a) Segments containing DOCT priors. b) Error map in percent for flow recovered with and without priors. c) Comparison of the MSE.

A detailed output of the simulation is shown in Fig. 7-2 with 20% of the vessels containing DOCT priors and for a DOCT noise level of 25%. Averaged results over all simulations are presented in Fig. 7-3. As seen in Fig. 7-3, the improvement in MSE varies almost linearly with the fraction of vessels with DOCT priors while noise has

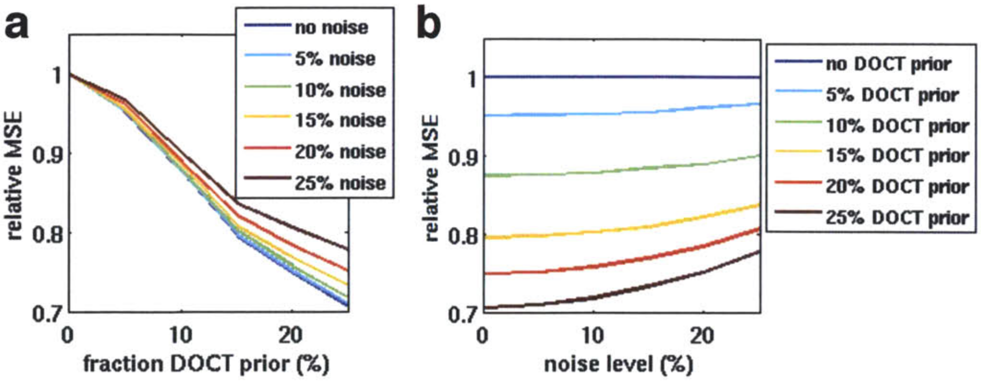

Figure 7-3: Averaged results of the simulations. a) Improvement in MSE for different fractions of vessels containing DOCT priors. b) Improvement in MSE for different level of noise in the DOCT priors.

less effect. Improvement in MSE of the order of 25% are obtained when 25% of the vessel contained DOCT priors, even in measurements contain 25% noise. This result highlights the importance of getting DOCT measurements in the largest possible number of vessels to reduce the error in the flow computation.

Another point worth mentioning is the influence of DOCT prior locations in the network. For the specific case illustrated in Fig. 7-2, the MSE was reduced by 60% which is more than twice the averaged value obtained from all simulations (shown in Fig. 7-3). This is explained by the location of the DOCT priors which are concentrated around one of the inflowing vessels. Constraining the flow value in an inflowing vessel will null the impact of the pressure boundary condition for this specific vessel and several downstream vessels since flow must be conserved along the vascular tree. This result indicates that getting DOCT measurements in large inflowing arteries or large outflowing veins will have the most impact on the flow computation.

7.4.2 Experimental measurements

The TPLSM angiogram and the DOCT velocity volume are shown in Fig. 7-4a) and b) respectively. The reconstructed flow distribution is illustrated in Fig. 7-4c). We observe a large range of flow across the network with higher flow in larger pial

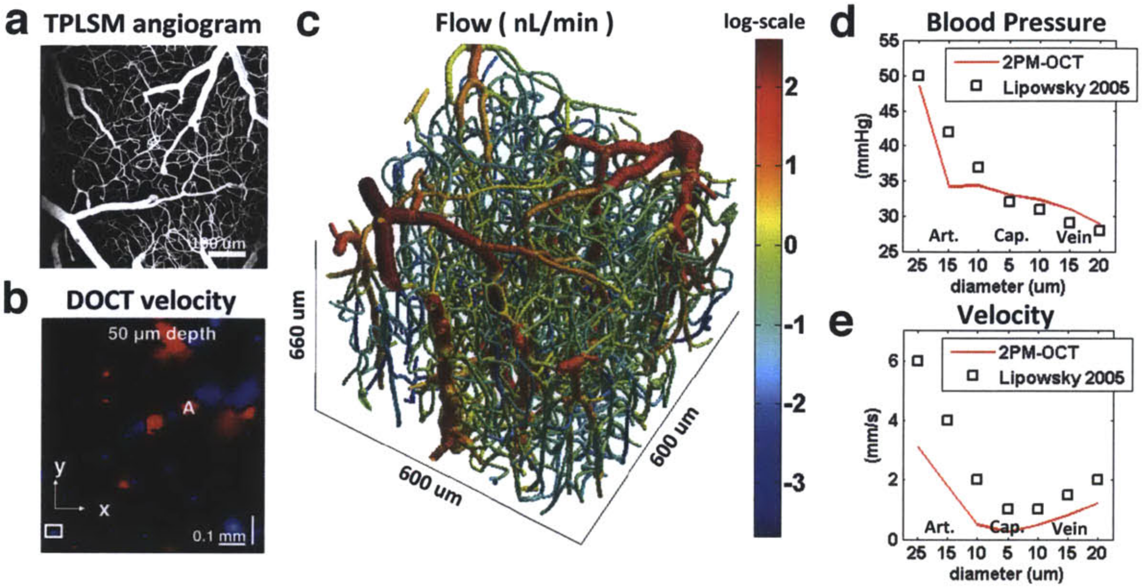

Figure 7-4: Reconstruction of flow in the mouse cortex. a) TPLSM FITC angiogram. b) DOCT velocity image. c) CBF computed from the the model with constrains from DOCT. d) Blood pressure and e) blood velocity computed from the model as a function of vessel type and diameter.

vessels and lower flow in smaller capillaries. The blood pressure drop and the velocity profile from arteries to veins are plotted in Fig. 7-4d) and e) respectively and are superimposed with experimental values measured on cats [87]. Our velocities for capillaries are also in good agreement with values measured in the mouse cortex [80].

7.5 Summary

In conclusion, we demonstrate that adding flow constraints results in more accurate reconstruction of blood flow in incomplete or truncated microvascular networks. Simulations show that measuring flow in 25% of the vessels can reduce the error by 25%. We finally use TPLSM angiography together with quantitative flow measurements from DOCT to apply this method for reconstruction of cerebral blood flow in the mouse cortex down to a depth of 660μm.

Bibliography

[1] A. F. Abdelnour and T. J. Huppert. Real-time imaging of humain brain function by near-infrared spectroscopy using an adaptive general linear model. *NeuroImage*, 46:133–143, 2009.

[2] Tomer Anor, Leopold Grinberg, Hyoungsu Baek, Joseph R Madsen, Mahesh V Jayaraman, and George E Karniadakis. Modeling of blood flow in arterial trees. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2(5):612–623, 2010.

[3] L Becerra, W Harris, D Joseph, T Huppert, and D A Boas. Diffuse optical tomography of pain and tactile stimulation: Activation in cortical sensory and emotional systems. *NeuroImage*, 2008.

[4] Pablo Blinder, Andy Y Shih, Christopher Rafie, and David Kleinfeld. Topological basis for the robust distribution of blood to rodent neocortex. *Proceedings of the National Academy of Sciences of the United States of America*, 107(28):12670–12675, July 2010.

[5] Pablo Blinder, Philbert S Tsai, John P Kaufhold, Per M Knutsen, Harry Suhl, and David Kleinfeld. The cortical angiome: an interconnected vascular network with noncolumnar patterns of blood flow. *Nature Neuroscience*, pages 1–11, June 2013.

[6] D.A. Boas, J.P. Culver, J.J. Stott, and A.K. Dunn. Three dimensional monte carlo code for photon migration through complex heterogenous media including the adult human head. *Opt. Express*, 10(3):159–170, 2002.

[7] D.A. Boas, A. M. Dale, and M. A Franceschini. Diffuse optical imaging of brain activation: Approaches to optimizing image sensitivity, resolution and accuracy. *NeuroImage*, 23:S275–S288, 2004.

[8] D.A. Boas, S. R. Jones, A. Devor, T.J. Huppert, and A. M. Dale. A vascular anatomical network model of the spatio-temporal response to brain activation. *NeuroImage*, 40(3):1116–1129, 2008.

[9] D S Bolar, B R Rosen, A G Sorensen, and E Adalsteinsson. QUantitative Imaging of eXtraction of oxygen and TIssue consumption (QUIXOTIC) using venular-targeted velocity-selective spin labeling. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 2011.

[10] J.L. Boxerman, P A Bandettini, K K Kwong, J R Baker, T L Davis, B R Rosen, and R M Weisskoff. The intravascular contribution to fmri signal change: monte carlo modeling and diffusionweighted studies in vivo. *Magnetic Resonance in Medicine*, 34(1):4–10, 1995.

[11] D. H. Brainard. The psychophysics toolbox. *Spatial Vision*, 10:433–436, 1997.

[12] K. Buchheim, H. Obrig, W. Pannwitz, A. Muller, H. Heekeren, A. Villringer, and H. Meierkord. Decrease in haemoglobin oxygenation during absence seizures in adult humans. *Neuroscience Letters*, 354:119–122, 2004.

[13] R. B. Buxton. *Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques.* Cambridge University Press, 2002.

[14] R. B. Buxton, K. Uludag, D. J. Dubowitz, and T. T. Liu. Modeling the hemodynamic response to brain activation. *NeuroImage*, 23:S220–S233, 2004.

[15] R.B Buxton, E.C Wong, and L.R Frank. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn. Reson. Med.*, 39(6):855–864, 1998.

[16] Richard B Buxton. Interpreting oxygenation-based neuroimaging signals: the importance and the challenge of understanding brain oxygen metabolism. *Frontiers in Neuroenergetics*, 2:1–16, 2010.

[17] Thomas Christen, Greg Zaharchuk, Nicolas Pannetier, Raphaël Serduc, Nicolas Joudiou, Jean-Claude Vial, Chantal Rémy, and Emmanuel L Barbier. Quantitative MR estimates of blood oxygenation based on T2*: A numerical study of the impact of model assumptions. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 67(5):1458–1468, December 2011.

[18] J Cohen-Adad, J R Polimeni, K G Helmer, T Benner, J A McNab, L L Wald, B R Rosen, and C Mainero. T2* mapping and B0 orientation-dependence at 7T reveal cyto- and myeloarchitecture organization of the human cortex. *NeuroImage*, 60(2):1006–1014, April 2012.

[19] M. Cope and D. T. Deply. System for long-term measurement of cerebral blood flow and tissue oxygenation on newborn infants by infrared transillumination. *Med. Biol. Eng. Comput.*, 26:289–294, 1998.

[20] X. Cui, S. Bray, and A. L. Reiss. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImages*, 49:3039–3046, 2010.

[21] Joseph P Culver, Andrew M Siegel, Maria Angela Franceschini, Joseph B Mandeville, and David A Boas. Evidence that cerebral blood volume can provide brain activation maps with better spatial resolution than deoxygenated hemoglobin. *NeuroImage*, 27(4):947–59, 2005.

[22] T.L Davis, K.K Kwong, R.M Weisskoff, and B.R Rosen. Calibrated functional mri: mapping the dynamics of oxidative metabolism. *Proc Natl Acad Sci USA*, 95(4):1834, 1998.

[23] M. Dehaes, L. Gagnon, F. Lesage, M. Pelegrini-Issac, A. Vignaud, R. Valabregue, R. Grebe, F. Wallois, and H. Benali. Quantitative investigation of the effect of the extra-cerebral vasculature in diffuse optical imaging: a simulation study. *Biomedical Optics Express*, 2(3):680–695, 2011.

[24] D. T. Deply, M. Cope, P. van der Zee, S. Arridge, S. Wray, and J. Wyatt. Estimation of optical pathlength through tissue from direct time of flight measurement. *Phys. Med. Biol.*, 33:1433–1442, 1988.

[25] A Devor, S Sakadžić, and PA Saisan. "Overshoot" of O2 Is Required to Maintain Baseline Tissue Oxygenation at Locations Distal to Blood Vessels. *The Journal of Neuroscience*, 31(38):13676–13681, 2011.

[26] Anna Devor, Elizabeth M C Hillman, Peifang Tian, Christian Waeber, Ivan C Teng, Lana Ruvinskaya, Mark H Shalinsky, Haihao Zhu, Robert H Haslinger, Suresh N Narayanan, Istvan Ulbert, Andrew K Dunn, Eng H Lo, Bruce R Rosen, Anders M Dale, David Kleinfeld, and David A Boas. Stimulus-induced changes in blood flow and 2-deoxyglucose uptake dissociate in ipsilateral somatosensory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(53):14347–14357, December 2008.

[27] Anna Devor, Peifang Tian, Nozomi Nishimura, Ivan C Teng, Elizabeth M C Hillman, S N Narayanan, Istvan Ulbert, David A Boas, David Kleinfeld, and Anders M Dale. Suppressed neuronal activity and concurrent arteriolar vasoconstriction may explain negative blood oxygenation level-dependent signal. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(16):4452–4459, April 2007.

[28] S. Diamond, T. J. Huppert, V. Kolehmainen, M. A. Franceschini, J. P. Kaipo, S. R. Arridge, and D. A. Boas. Dynamic physiological modeling for functional diffuse optical tomography. *NeuroImage*, 30:88–101, 2006.

[29] S. G. Diamond, K. L. Perdue, and D. A. Boas. A cerebrovascular response model for functional neuroimaging including dynamic cerebral autoregulation. *Math. Biosci.*, 220(2):102–117, 2009.

[30] Patrick J Drew, Andy Y Shih, and David Kleinfeld. Fluctuating and sensory-induced vasodynamics in rodent cortex extend arteriole capacity. *Proc Natl Acad Sci USA*, 108(20):8473–8478, 2011.

[31] Patrick J PJ Drew, Andy Y AY Shih, Jonathan D JD Driscoll, Per Magne PM Knutsen, Pablo P Blinder, Dimitrios D Davalos, Katerina K Akassoglou, Philbert S PS Tsai, and David D Kleinfeld. Chronic optical access through a polished and reinforced thinned skull. *Nature Methods*, 7(12):981–984, December 2010.

[32] S Dubeau, G Ferland, P Gaudreau, E Beaumont, and F Lesage. Cerebrovascular hemodynamic correlates of aging in the Lou/c rat: A model of healthy aging. *NeuroImage*, 56(4):10–10, June 2011.

[33] Audrey P Fan, Thomas Benner, Divya S Bolar, Bruce R Rosen, and Elfar Adalsteinsson. Phase-Based Regional Oxygen Metabolism (PROM) using MRI. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, pages 1–10, 2011.

[34] Qianqian Fang and D A Boas. Tetrahedral mesh generation from volumetric binary and grayscale images. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Proceedings*, pages 1142–1145, January 2009.

[35] Qianqian Fang, Sava Sakadžić, Lana Ruvinskaya, Anna Devor, Anders M. Dale, and David A. Boas. Oxygen Advection and Diffusion in a Three Dimensional Vascular Anatomical Network. *Optics Express*, 16(22):17530–17541, 2008.

[36] M. Ferrari and V. Caresima. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, 2012.

[37] Olga S Finikova, Artem Y Lebedev, Alexey Aprelev, Thomas Troxler, Feng Gao, Carmen Garnacho, Silvia Muro, Robin M Hochstrasser, and Sergei A Vinogradov. Oxygen microscopy by two-photon-excited phosphorescence. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 9(12):1673–1679, August 2008.

[38] M. Firbank, E. Okada, and D. T. Delpy. A theoretical study of the signal contribution of regions of the adult head to near-infrared spectroscopy studies of visual evoked responses. *NeuroImage*, 8(1):69–78, 1998.

[39] Peter T Fox, Marcus E Raichle, Mark A Mintun, and Carmen Dence. Nonoxidative glucose consumption during focal physiologic neural activity. *Science (New York, NY)*, 241(4864):462–464, 1988.

[40] M.A. Franceschini, S. Fantini, J.H. Thompson, J.P. Culver, and D.A. Boas. Hemodynamic evoked response of the sensorimotor cortex measured noninvasively with near-infrared optical imaging. *Psychophysiology*, 40(4):548–560, 2003.

[41] M.A. Franceschini, D. K. Joseph, T. J. Huppert, S. G. Diamond, and D.A. Boas. Diffuse optical imaging of the whole head. *J. Biomed. Opt.*, 11(5):054007, 2006.

[42] K Friston. Nonlinear responses in fMRI: The balloon model, volterra kernels, and other hemodynamics. *NeuroImage*, 12(4):466–477, 2000.

[43] L. Gagnon, R. J. Cooper, M. A. Yücel, K. L. Perdue, D. N. Greve, and D. A. Boas. Short separation channel location impacts the performance of short channel regression in NIRS. *NeuroImage*, 59(3):2518–2528, 2012.

[44] L. Gagnon, K. Perdue, D. N. Greve, D. Goldenholz, G. Kaskhedikar, and D. A. Boas. Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling. *NeuroImage*, 56(3):1362–1371, 2011.

[45] Louis Gagnon, Claudine C Gauthier, Rick D RD Hoge, Frédéric F Lesage, Juliette J Selb, and David A DA Boas. Double-layer estimation of intra- and extracerebral hemoglobin concentration with a time-resolved system. *Journal of Biomedical Optics*, 13(5):054019–054019, September 2008.

[46] A. Gelb. *Applied Optimal Estimation.* MIT Press, 1974.

[47] A.P. Gibson, J.C. Hebden, and S.R. Arridge. Recent advances in diffuse optical imaging. *Phys. Med. Biol.*, 50:R1–R43, 2005.

[48] G. H. Glover. Deconvolution of impulse response in event-related bold fmri. *NeuroImage*, 9(4):416–429, 1999.

[49] G. Gratton and P. M. Corballis. Removing the heart from the brain: compensation for the pulse artifact in the photon migration signal. *Psychophysiology*, 32(3):292–299, 1995.

[50] Henry Gray. *Anatomy of the Humain Body.* Lea & Febiger, Philadelphia, 2000.

[51] N. M. Gregg, B. R. White, B. W. Zeff, A. J. Berger, and J. P. Culver. Brain specificity of diffuse optical imaging: improvements from superficial signal regression and tomography. *Font. in Neuroenergetics*, 2(14):1–8, 2010.

[52] Valerie E M Griffeth and Richard B Buxton. A theoretical framework for estimating cerebral oxygen metabolism changes using the calibrated-BOLD method: modeling the effects of blood volume distribution, hematocrit, oxygen

extraction fraction, and tissue signal properties on the BOLD signal. *NeuroImage*, 58(1):198–212, September 2011.

[53] Romain Guibert, Caroline Fonta, and Franck Plouraboué. Cerebral blood flow modeling in primate cortex. *Journal of cerebral blood flow and metabolism*, 30(11):1860–1873, July 2010.

[54] K. Haginoya, M. Munakata, R. Kato, H. Yokoyama, M. Ishizuka, and K. Iinuma. Ictal cerebral haemodynamics of chilhood epilespsy measured with near-infrared spectroscopy. *Brain*, 125:1960–1971, 2002.

[55] Simon Haykin. *Adaptive Filter Theory.* Prentice-Hall, Upper Saddle River, NJ, 2001.

[56] Simon Haykin. *Kalman filtering and neural Networks.* John Wiley & Sons, New York, 2001.

[57] Xiang He, Mingming Zhu, and Dmitriy A Yablonskiy. Validation of oxygen extraction fraction measurement by qBOLD technique. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 60(4):882–888, 2008.

[58] E Hillman, A Devor, M.B Bouchard, A.K Dunn, GW Krauss, J Skoch, B.J Bacskai, A.M Dale, and D.A Boas. Depth-resolved optical imaging and microscopy of vascular compartment dynamics during somatosensory stimulation. *NeuroImage*, 35(1):89–104, 2007.

[59] E. M. C. Hillman. Optical brain imaging in vivo: techniques and applications from animal to man. *J. Biomed. Opt.*, 12(5):051402, 2007.

[60] Sven Hirsch, Johannes Reichold, Matthias Schneider, Gábor Székely, and Bruno Weber. Topology and hemodynamics of the cortical cerebrovascular system. *Journal of Cerebral Blood Flow and Metabolism*, 32(6):952–967, April 2012.

[61] R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. B. Pike. Investigation of bold signal dependence on cerebral blood flow and oxygen consumption: The deoxyhemoglobin dilution model. *Mag. Res. in Med.*, 42:849–863, 1999.

[62] R. D. Hoge, M. A. Franceschini, R. J. M. Covolan, T. Huppert, J. B. Mandeville, and D. A. Boas. Simultaneous recording of task-induced changes in blood oxygenation, volume, and flow using diffuse optical imaging and arterial spin-labeling mri. *NeuroImage*, 25:701–707, 2005.

[63] Y. Hoshi. Functional near-infrared spectroscopy: current status and future prospects. *J. Biomed. Opt.*, 12(6):062106, 2007.

[64] X.-S. Hu, K.-S. Hong, S. S. Ge, and M.-Y. Jeong. Kalman estimator- and general linear model-based on-line brain activation mapping by near-infrared spectroscopy. *BioMedical Engineering OnLine*, 9(82), 2010.

[65] T. J. Huppert. *Hemodynamic-Based Inference of Cerebral Oxygen Metabolism.* PhD thesis, Harvard University, 2007.

[66] T. J. Huppert, M. S. Allen, S. G. Diamond, and D. A. Boas. Estimating cerebral oxygen metabolism from fmri with a dynamic multicompartment windkessel model. *Hum. Brain Mapp.*, 30(5):1548–1567, 2009.

[67] T. J. Huppert, S. G. Diamond, and D. A. Boas. Direct estimation of evoked hemoglobin changes by multimodality fusion imaging. *J. Biomed. Opt.*, 13(5):054031, 2008.

[68] T.J. Huppert, M. S. Allen, H. Benay, P. B. Jones, and D.A. Boas. A multi-compartment vascular model for inferring baseline and functional changes in cerebral oxygen metabolism and arterial dilation. *J. Cereb. Blood Flow Metab.*, 27:1262–1279, 2007.

[69] T.J. Huppert, R.D. Hoge, A. M. Dale, M.A. Franceschini, and D.A. Boas. Quantitative spatial comparison of diffuse optical imaging with blood oxygen

level-dependent and arterial spin labeling-based functional magnetic resonance imaging. *J. Biomed. Opt.*, 11(6):064018, 2006.

[70] T.J. Huppert, R.D. Hoge, S.G. Diamond, M.A. Franceschini, and D.A. Boas. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *NeuroImage*, 29:368–382, 2006.

[71] Farzin Irani, Steven M. Platek, Scott Bunce, Anthony C. Ruocco, and Douglas Chute. Functional near infrared spectroscopy (fNIRS): An emerging neuroimaging technology with important applications for the study of brain disorders. *The Clinical Neuropsychologist*, 21(1):9–37, 2010.

[72] Varsha Jain, Michael C Langham, and Felix W Wehrli. MRI estimation of global brain oxygen consumption rate. *J. Cereb. Blood Flow Metab.*, 30(9):1598–1607, 2010.

[73] K. E. Jang, S. Tak, J. Jung, J. Jang, Y. Jeong, and J. C. Ye. Wavelet minimum description length detrending for near-infrared spectroscopy. *J. Biomed. Opt.*, 14(3):034004, 2009.

[74] G. Jasdzewski, G. Strangman, J. Wagner, K. K. Kwong, R. A. Poldorack, and D. A. Boas. Differences in the hemodynamic response to event-related motor and visual paradigms as measured by near- infrared spectroscopy. *NeuroImage*, 20(1):479–488, 2003.

[75] Sune N Jespersen and Leif Ostergaard. The roles of cerebral blood flow, capillary transit time heterogeneity, and oxygen tension in brain oxygenation and metabolism. *Journal of Cerebral Blood Flow and Metabolism*, 32:264–277, November 2012.

[76] C. Julien. The enigma of Mayer waves: facts and models. *Cardiovascular Research*, 70:12–21, 2007.

[77] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[78] Seong-Gi Kim and Seiji Ogawa. Biophysical and physiological origins of blood oxygenation level-dependent fMRI signals. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, March 2012.

[79] E. Kirilina, A. Jelzow, A. Heine, M. Niessing, H. Wabritz, R. Bruehl, B. Ittermann, A. M. Jacobs, and I. Tachtsidis. The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy. *NeuroImage*, 61:70–81, 2011.

[80] D D Kleinfeld, P P PP Mitra, F F Helmchen, and W W Denk. Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex. *PNAS*, 95(26):15741–15746, December 1998.

[81] V. Kolehmainen, S. Prince, S. R. Arridge, and J. P Kaipo. State- estimation approach to the nonstationary optical tomography problem. *J. Opt. Soc. Am. A*, 20(5):876–889, 2003.

[82] K K Kwong, J W Belliveau, D A Chesler, I E Goldberg, R M Weisskoff, B P Poncelet, D N Kennedy, B E Hoppel, M S Cohen, and R Turner. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, 89(12):5675–5679, June 1992.

[83] Jérôme Lecoq, Alexandre Parpaleix, Emmanuel Roussakis, Mathieu Ducros, Yannick Goulam Houssen, Sergei A Vinogradov, and Serge Charpak. Simultaneous two-photon imaging of oxygen and blood flow in deep cerebral vessels. *Nature Medicine*, 17(7):893–898, 2011.

[84] J-M. Lina, M. Dehaes, C. Matteau-Pelletier, and F. Lesage. Complex wavelets applied to diffuse optical spectroscopy for brain activity detection. *Optics Express*, 16(2):1029–1050, 2008.

[85] J-M. Lina, C. Matteau-Pelletier, M. Dehaes, M. Desjardins, and F. Lesage. Wavelet-based estimation of the hemodynamic responses in diffuse optical imaging. *Med. Imag. Anal.*, 14:606–616, 2010.

[86] Liis L Lindvere, Rafal R Janik, Adrienne A Dorr, David D Chartash, Bhupinder B Sahota, John G JG Sled, and Bojana B Stefanovic. Cerebral microvascular network geometry changes in response to functional stimulation. *NeuroImage*, 71:248–259, May 2013.

[87] Herbert Lipowsky. Microvascular Rheology and Hemodynamics. *Microcirculation*, 12(1):5–15, February 2005.

[88] S. Lloyd-Fox, A. Blasi, and C. E. Elwell. Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy. *Neuroscience & Biobehavioral Reviews*, 34:269–284, 2010.

[89] Nikos K Logothetis. What we can do and what we cannot do with fMRI. *Nature*, 453(7197):869–878, June 2008.

[90] S Lorthois, F Cassot, and F Lauwers. Simulation study of brain blood flow regulation by intra-cortical arterioles in an anatomically accurate large human vascular network: Part I: Methodology and baseline flow. *NeuroImage*, 54(2):12–12, January 2011.

[91] S Lorthois, F Cassot, and F Lauwers. Simulation study of brain blood flow regulation by intra-cortical arterioles in an anatomically accurate large human vascular network. Part II: flow variations induced by global or localized modifications of arteriolar diameters. *NeuroImage*, 54(4):2840–2853, February 2011.

[92] Dan Ma, Vikas Gulani, Nicole Seiberlich, Kecheng Liu, Jeffrey L Sunshine, Jeffrey L Duerk, and Mark A Griswold. Magnetic resonance fingerprinting. *Nature*, 495(7440):187–192, March 2013.

[93] A. Machado, J-M. Lina, J. Tremblay, M. Lassonde, D. K. Nguyen, F. Lesage, and C. Grova. Detection of hemodynamic responses to epileptic activity us-

ing simultaneous electro-encephalography (EEG)/near infra red spectroscopy (NIRS) acquisitions. *NeuroImage*, 56(1):114–125, 2010.

[94] C. Matteau-Pelletier, M. Dehaes, F. Lesage, and J-M. Lina. 1/f noise in diffuse optical imaging and wavelet-based response estimation. *IEEE Trans. Med. Imag.*, 28:415–422, 2009.

[95] P. Moran. A flow velocity zeugmatographic interlace for nmr imaging in humans. *Mag. Res. Med.*, 1:197–203, 1982.

[96] Nozomi Nishimura, Chris B Schaffer, Beth Friedman, Philbert S Tsai, Patrick D Lyden, and David Kleinfeld. Targeted insult to subsurface cortical blood vessels using ultrashort laser pulses: three models of stroke. *Nature Methods*, 3(2):99–108, February 2006.

[97] T Obata, T.T Liu, K.L Miller, W.M Luh, E.C Wong, L.R Frank, and R.B Buxton. Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas: application of the balloon model to the interpretation of BOLD transients. *NeuroImage*, 21(1):144–153, 2004.

[98] H. Obrig, M. Neufang, R. Wenzel, M. Kohl, J. Steinbrink, K. Einhaupl, and A. Villringer. Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults. *NeuroImage*, 12(6):623–639, 2000.

[99] Y. Obrig and A. Villringer. Beyond the visible: imaging the human brain with light. *J. Cereb. Blood Flow Metab.*, 23:1–18, 2003.

[100] Dominik Obrist, Bruno Weber, Alfred Buck, and Patrick Jenny. Red blood cell distribution in simplified capillary networks. *Philosophical transactions Series A, Mathematical, physical, and engineering sciences*, 368(1921):2897–2918, June 2010.

[101] S Ogawa, TM Lee, AR Kay, and DW Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(24):9868, 1990.

[102] S Ogawa, D W Tank, R Menon, J M Ellermann, S G Kim, H Merkle, and K Uğurbil. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*, 89(13):5951–5955, July 1992.

[103] Alexandre Parpaleix, Yannick Goulam Houssen, and Serge Charpak. Imaging local neuronal activity by monitoring PO2 transients in capillaries. *Nature Medicine*, 19(2):241–246, January 2013.

[104] AP Pathak and BD Ward. ScienceDirect.com - NeuroImage - A novel technique for modeling susceptibility-based contrast mechanisms for arbitrary microvascular geometries: The finite perturber method. *NeuroImage*, 2008.

[105] S. Payne, J. Selb, and D. A. Boas. Effects of autoregulation and CO2 reactivity on cerebral oxygen transport. *Ann. Biomed. Eng.*, 37(11):2288–2298, 2009.

[106] K. L. Perdue and S. G. Diamond. Anatomical vasculature modeling for functional near-infrared spectroscopy. 17th Annual Meeting of the Organization for Human Brain Mapping, Quebec City, 2011.

[107] A R Pries, D Neuhaus, and P Gaehtgens. Blood viscosity in tube flow: dependence on diameter and hematocrit. *Am J Physiol Heart Circ Physiol*, 263:H1170–H1778, 1992.

[108] A R Pries, T W Secomb, P Gaehtgens, and J F Gross. Blood flow in microvascular networks. Experiments and simulation. *Circulation Research*, 67:826–834, 1990.

[109] A R Pries, T W Secomb, T Gessner, M B Sperandio, J F Gross, and P Gaehtgens. Resistance to blood flow in microvessels in vivo. *Circulation*, 75:904–915, 1994.

[110] Axel R Pries, Michael Höpfner, Ferdinand le Noble, Mark W Dewhirst, and Timothy W Secomb. The shunt problem: control of functional shunting in nor-

mal and tumour vasculature. *Nature Reviews: Cancer*, 10(8):587–593, August 2010.

[111] S. Prince, V. Kolehmainen, J. P. Kaipo, M. A. Franceschini, D. A Boas, and S. R. Arridge. Time-series estimation of biological factors in optical diffusion tomography. *Phys. Med. Biol.*, 48:1491–1504, 2003.

[112] M E Raichle. Visualizing the mind. *Scientific American*, 1994.

[113] H. E. Rauch, , F. Tung, and C. T. Striebel. Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8):1445–1450, 1965.

[114] J Reichold, M Stampanoni, and A Keller. Vascular graph model to simulate the cerebral blood flow in realistic vascular networks. *Journal of Cerebral Blood Flow and Metabolism*, 29:1429–1443, 2009.

[115] N. Roche-Labarbe, B. Zaaimi, P. Berquin, A. Nehlig, R. Grebe, and F. Wallois. NIRS-measured oxy- and deoxyhemoglobin changes associated with EEG spike-and-wave discharges in children. *Epilepsia*, 49(11):1871–1880, 2008.

[116] Rolf B. Saager and Andrew J. Berger. Direct characterization and removal of interfering absorption trends in two-layer turbid media. *J. Opt. Soc. Am. A*, 22(9):1874–1882, 2005.

[117] Rolf B. Saager and Andrew J. Berger. Measurement of layer-like hemodynamic trends in scalp and cortex: implications for physiological baseline suppression in functional near-infrared spectroscopy. *J. Biomed. Opt.*, 13(3):034017, 2008.

[118] Rolf B. Saager, Nicole L. Telleri, and Andrew J. Berger. Two-detector corrected near infrared spectroscopy (C-NIRS) detects hemodynamic activation responses more robustly than single-detector NIRS. *NeuroImage*, 55(4):1679–1685, 2011.

[119] Sava Sakadžić, Emmanuel Roussakis, Mohammad A Yaseen, Emiri T Mandeville, Vivek J Srinivasan, Ken Arai, Svetlana Ruvinskaya, Anna Devor, Eng H

Lo, Sergei A Vinogradov, and David A Boas. Two-photon high-resolution measurement of partial pressure of oxygen in cerebral vasculature and tissue. *Nature Methods*, 7(9):755–759, September 2010.

[120] Sava Sakadžić, Shuai Yuan, Ergin Dilekoz, Svetlana Ruvinskaya, Sergei A Vinogradov, Cenk Ayata, and David A Boas. Simultaneous imaging of cerebral partial pressure of oxygen and blood flow during functional activation and cortical spreading depression. *Applied optics*, 48(10):D169–77, April 2009.

[121] Matthias M Schneider, Johannes J Reichold, Bruno B Weber, Gábor G Székely, and Sven S Hirsch. Tissue metabolism driven arterial tree generation. *Medical Image Analysis*, 16(7):1397–1414, October 2012.

[122] S.A Sheth, M Nemoto, M Guiou, M Walker, N Pouratian, N Hageman, and A.W Toga. Columnar specificity of microvascular oxygenation and volume responses: implications for functional brain mapping. *J. of Neuroscience*, 24(3):634, 2004.

[123] L Sokoloff. Relation between physiological function and energy metabolism in the central nervous system. *Journal of neurochemistry*, 29(1):13–26, July 1977.

[124] Vivek J Srinivasan, Dmitriy N Atochin, Harsha Radhakrishnan, James Y Jiang, Svetlana Ruvinskaya, Weicheng Wu, Scott Barry, Alex E Cable, Cenk Ayata, Paul L Huang, and David A Boas. Optical coherence tomography for the quantitative study of cerebrovascular physiology. *Journal of cerebral blood flow and metabolism*, 31(6):1339–1345, March 2011.

[125] Vivek J Srinivasan, Sava Sakadžić, Iwona Gorczynska, Svetlana Ruvinskaya, Weicheng Wu, James G Fujimoto, and David A Boas. Quantitative cerebral blood flow with optical coherence tomography. *Optics express*, 18(3):2477–2494, February 2010.

[126] G. Strangman, M. A. Franceschini, and D. A. Boas. Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters. *NeuroImage*, 18:865–879, 2003.

[127] S. Tak, J. Jang, K. Lee, Y. Jeong, and J. C. Ye. Quantification of CMRO2 without hypercapnia using simultaneous near-infrared spectroscopy and fMRI measurements. *Phy. Med. Biol.*, 55:3249–3269, 2010.

[128] S. Tak, S. J. Yoon, J. Jang, K. Yoo, Y. Jeong, and J. C. Ye. Quantitative analysis of hemodynamic and metabolic changes in subcortical vascular dementia using simultaneous near-infrared spectroscopy and fMRI measurements. *NeuroImage*, 55:176–184, 2011.

[129] T. Takahashi, Y. Takikawa, R. Kawagoe, S. Shibuya, T. Iwano, and S. Kitazawa. nfluence of skin blood flow on near-infrared spectroscopy signals measured on the forehead during a verbal fluency task. *NeuroImage*, 57:991–1002, 2011.

[130] F. Tian, H. Niu, B. Khan, G. Alexandrakis, K. Behbehani, and H. Liu. Enhanced functional brain imaging by using adaptive filtering and a depth compensation algorithm in Diffuse Optical Tomography. *IEEE Trans. Medical Imaing*, 30(6):1239–1251, 2011.

[131] Peifang Tian, Ivan C Teng, Larry D May, Ronald Kurz, Kun Lu, Miriam Scadeng, Elizabeth M C Hillman, Alex J De Crespigny, Helen E D'Arceuil, Joseph B Mandeville, John J A Marota, Bruce R Rosen, Thomas T Liu, David A Boas, Richard B Buxton, Anders M Dale, and Anna Devor. Cortical depth-specific microvascular dilation underlies laminar differences in blood oxygenation level-dependent functional MRI signal. *Proceedings of the National Academy of Sciences*, 107(34):15246–15251, August 2010.

[132] Yunjie Tong and Blaise deB Frederick. Time lag dependent multimodal processing of concurrent fMRI and near-infrared spectroscopy (NIRS) data suggests a global circulatory origin for low-frequency oscillation signals in human brain. *NeuroImage*, 53(2):553–564, 2010.

[133] V. Tonorov, M. Franceschini, M. Filiaci, S. Fantini, M. Wolf, A. Michalos, and E. Gratton. Near-infrared study of fluctuations in cerebral hemodynamics

during rest and motor stimulation: temporal analysis and spatial mapping. *Med. Phys.*, 27(4):801–815, 2000.

[134] P S Tsai, J P Kaufhold, P Blinder, B Friedman, P J Drew, H J Karten, P D Lyden, and D Kleinfeld. Correlations of Neuronal and Microvascular Densities in Murine Cortex Revealed by Direct Counting and Colocalization of Nuclei and Vessels. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(46):14553–14570, November 2009.

[135] Kâmil Uludağ, Bernd Müller-Bierl, and Kamil Ugurbil. An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *NeuroImage*, 48(1):150–165, October 2009.

[136] S. Umeyama and T. Yamada. Monte carlo study of global interference cancellation by multidistance measurement of near-infrared spectroscopy. *J. Biomed. Opt.*, 14(6):064025, 2009.

[137] S. Umeyama and T. Yamada. New method of estimating wavelength-dependent optical path length ratios for oxy- and deoxyhemoglobin measurement using near-infrared spectroscopy. *J. Biomed. Opt.*, 14(5):054038, 2009.

[138] A. Villringer, C. Hock, L. Schleinkofer, and U. Dirnagl. Near infrared spectroscopy (NIRS): A new tool to study hemodynamic changes during activation of brain function in human adults. *Neurosci. Lett.*, 154:101–104, 1993.

[139] F. Wallois, A. Patil, C. Heberle, and R. Grebe. EEG-NIRS in epilepsy in children and neonates. *Clinical Neurophysiology*, 40:281–292, 2010.

[140] H F Wehrl, M S Judenhofer, F C Maier, P Martirosian, G Reischl, F Schick, and B J Pichler. Simultaneous assessment of Perfusion with [15O] water PET and arterial spin labeling MR using a hybrid PET/MR device. *ISMRM*, 18:715, March 2010.

[141] T. Yamada, S. Umeyama, and K. Matsuda. Multidistance probe arrangement to eliminate artifacts in functional near-infrared spectroscopy. *J. Biomed. Opt.*, 16(4):06434, 2009.

[142] Mohammad A Yaseen, Vivek J Srinivasan, Sava Sakadžić, Harsha Radhakrishnan, Iwona Gorczynska, Weicheng Wu, James G Fujimoto, and David A Boas. Microvascular oxygen tension and flow measurements in rodent cerebral cortex during baseline conditions and functional activation. *J Cereb Blood Flow Metab*, 31:1051–1063, 2011.

[143] Mohammad A Yaseen, Vivek J Srinivasan, Sava Sakadžić, Weicheng Wu, Svetlana Ruvinskaya, Sergei A Vinogradov, and David A Boas. Optical monitoring of oxygen tension in cortical microvessels with confocal microscopy. *Optics express*, 17(25):22341–22350, December 2009.

[144] M. A. Yucel, T. J. Huppert, D. A. Boas, and L. Gagnon. Calibrating the BOLD signal during a motor task using an extended fusion model incorporating DOT, BOLD and ASL data. *NeuroImage*, 61:1268–1276, 2012.

[145] Q. Zhang, D. H. Brooks, M. A. Franceschini, and D. A. Boas. Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *J. Biomed. Opt.*, 10(1):011014, 2005.

[146] Q. Zhang, G. Strangman, and G. Ganis. Adaptive filtering to reduce global interference in non-invasive nirs measures of brain activation: How well and when does it work? *NeuroImage*, 45:788–794, 2009.

[147] Quan Zhang, Emery N. Brown, and Gary E. Strangman. Adaptative filtering for global interference cancellation and real-time recovery of evoked brain activity: a Monte Carlo simulation study. *J. of Biomed. Opt.*, 12(4):044014, 2007.

[148] Quan Zhang, Emery N. Brown, and Gary E. Strangman. Adaptative filtering to reduce global interference in evoked brain activity detection: a human subject case study. *J. of Biomed. Opt.*, 12(6):064009, 2007.

[149] Bingwen Zheng, Philip Teck Hock Lee, and Xavier Golay. High-sensitivity cerebral perfusion mapping in mice by kbGRASE-FAIR at 9.4 T. *NMR in biomedicine*, 23(9):1061–1070, July 2010.