Accurate Hemodynamic Response Estimation by Removal of Stimulus-Evoked Superficial Response in fNIRS Signals.

Alessandra Galli[1], **Sabrina Brigadoi**[1,2], **Giada Giorgi**[1], **Giovanni Sparacino**[1], **Claudio Narduzzi**[1]

[1] Department of Information Engineering University of Padova, I-35131 Padua, Italy

[2] Department of Developmental Psychology University of Padova, I-35131 Padua, Italy

E-mail: galliale@dei.unipd.it

Abstract. *Objective* We address the problem of hemodynamic response estimation when task-evoked extra-cerebral components are present in functional near-infrared spectroscopy (fNIRS) signals. These components might bias the hemodynamic response estimation, therefore careful and accurate denoising of data is needed.

Approach We propose a dictionary-based algorithm to process each single event-related segment of the acquired signal for both long separation and short separation channels. Stimulus-evoked components and physiological noise are modeled by means of two distinct waveform dictionaries.

For each segment, after removal of the physiological noise component in each channel, a template is employed to estimate stimulus-evoked responses in both channels. Then, the estimate from the short-separation channel is employed to correct for the evoked superficial response and refine the hemodynamic response estimate from the long-separation channel.

Main results Analysis of simulated, semi-simulated and real data shows that, by averaging single-segment estimates over multiple trials in an experiment, reliable results and improved accuracy compared to other methods can be obtained. While still far from the possibility of single-trial hemodynamic response estimation, a significant reduction in the number of averaged trials can also be obtained.

Significance This work proves that dedicated dictionaries can be successfully employed to model all different components of fNIRS signals. It demonstrates the effectiveness of a specifically designed algorithm structure in dealing with a complex denoising problem, enhancing the possibilities of fNIRS-based hemodynamic response analysis.

1. Introduction

Functional near-infrared spectroscopy (fNIRS) is a non-invasive, low cost neuroimaging technique, widely used to monitor cerebral activity and study healthy or pathological brain activation in response to a variety of cognitive tasks [1]. Neuroscience applications include the study of infant brain and cognitive development and of functional connectivity in the human brain [2, 3]. In clinical practice, fNIRS helps to investigate processes associated with neurological and psychiatric disorders, like Alzheimer disease, Parkinson disease, epilepsy, schizophrenia, and anxiety disorders [4]. More recently, fNIRS has also been considered in brain-computer interface [5, 6].

Because of neurovascular coupling, local changes in the concentrations of oxy- and deoxy-hemoglobin (respectively, HbO and HbR) occur in response to a particular stimulus or cognitive task. fNIRS monitors these changes by measuring absorbance variations, using a beam of near-infrared light that propagates through the brain between pairs of suitably positioned sources and detectors on the scalp. The estimated hemodynamic response (HR) can be interpreted as an indirect measurement of neural activity in the investigated brain area.

Since the light path includes the scalp, skull and cerebospinal fluid, the received signal is unavoidably contaminated by extra-cerebral and systemic components. Whereas HR amplitude is typically in the order of hundreds of nM, a much higher signal is recorded by the detector, where disturbance components are indeed the predominant part of the signal [7]. Physiological contributions due to extra-cerebral components associated to respiration, cardiac activity and Mayer waves [8] partially overlap the HR frequency band, but do not depend on neuronal activity, therefore they are present even without external stimuli. Extra-cerebral physiological contaminants can be also test-dependently, with increases in heart-rate, respiration and/or blood pressure time located to the stimulus [9].

In a typical fNIRS measurement set-up the "standard", or long-separation (LS) channel is supplemented by a shorter "reference", or short-separation (SS) channel, that relies on a second detector placed closer to the same light source. As photon penetration depth is related to source-detector distance, it is assumed that the SS channel can be referred almost exclusively to the extra-cerebral part while the LS channel probes deeper, providing information about both brain and extra-cerebral part [10]. The reference measurement provided by the SS channel can then be exploited to estimate disturbance contributions and correct for them, allowing significant improvements in accuracy.

Several approaches have been proposed in the literature [11]. For instance, the reference measurement can be used to regress superficial components from the standard channel, obtaining a scaling factor and, by subtraction of the component correlated with the SS channel, a "corrected" signal [12]. This and similar methods rely on the assumption that the superficial response measured by the SS channel represents only systemic noise with no task-related response, independent of the cerebral hemodynamic response. However, the assumption may not be always acceptable [13].

Experimental tasks in some fNIRS studies involve complex emotional and cognitive processing that, besides evoking a neural response within the brain, also induce changes in cutaneous vein blood volume due to an increase in heart rate and blood pressure [14]. This stimulus-evoked response is superficial, therefore both the LS and SS channels are affected. Furthermore, a superficial response may have similar features to HR, which consequently may be either enhanced or attenuated in the measured LS signal. These confounding effects can be considered, respectively, a false positive (FP) or a false negative (FN) as far as HR detection is concerned [9], [15]. As it can no longer be assumed that the reference channel is nearly uncorrelated with stimulus-evoked responses, the presence of a HR-correlated component has to be taken into account also in the denoising process, to prevent similar undesired modifications.

In this paper we present a dictionary-based denoising and waveform estimation technique that explicitly models the presence of a HR-correlated component in fNIRS signals. We propose the use of a combination of two dictionaries, which enable to deal with both physiological and task-evoked noise, leading to a more accurate determination of the HR.

The main features of our approach are the following:

- as a preliminary step, physiological background noise is estimated and removed from the two fNIRS channels independently. For this purpose we employ the dictionary developed in [17] to estimate parameters of a physiological noise model based on a sum of quasi-sinusoidal waves;

- we introduce a dedicated dictionary that allows estimation of task-related components in both LS and SS channels, where the latter contains information only about the stimulus-evoked superficial response;
- using information from resting state intervals, we determine a scaling factor to account for the different paths in the two channels, which allows us to use the superficial response contribution estimated in the SS channel to correct the denoised LS signal;
- finally, unreliable individual estimates are discarded before computing the averaged HR response estimate.

2. Method

2.1. Outline

In an experiment, a subject is presented with a series of stimuli referring to a given task, for a total of N_T trials. Continuously recorded variations in the concentrations of HbO and HbR produce a sequence of pulses $h_i(t)$, each representing the hemodynamic response to a single stimulus within the experiment. Indicating by T_i the onset of the i-th stimulus, the signal of interest is:

$$v(t) = \sum_{i=1}^{N_T} h_i(t - T_i). \tag{1}$$

For event-related analysis, fNIRS signals acquired during an experiment are partitioned into segments. Each single-trial response is time-locked to the onset of the corresponding stimulus, accordingly the signal of interest is assumed to be the single response $h(t)$.

We define a segment as a vector $\underline{\mathbf{y}} = [y(n_1T_s)\dots y(n_2T_s)]^T$ containing a set of $N = n_2 - n_1 + 1$ samples, with T_s as the sampling interval. Dropping trial index i for simplicity, the signals acquired by a dual-detector system can be described by the equations:

$$\begin{aligned} \underline{\mathbf{y}}_{\mathrm{L}} &= \underline{\mathbf{h}} + \underline{\mathbf{s}}_{\mathrm{L}} + \underline{\mathbf{p}}_{\mathrm{L}} + \underline{\mathbf{w}}_{\mathrm{L}} \\ \underline{\mathbf{y}}_{\mathrm{S}} &= \underline{\mathbf{s}}_{\mathrm{S}} + \underline{\mathbf{p}}_{\mathrm{S}} + \underline{\mathbf{w}}_{\mathrm{S}} \end{aligned} \tag{2}$$

where subscripts 'L' and 'S' refer, respectively, to the long- and short-separation channels. $\underline{\mathbf{w}}_L$ and $\underline{\mathbf{w}}_S$ are random noise vectors with finite variances that model acquisition noise introduced by the measurement system. Stimulus-evoked systemic variations are represented by $\underline{\mathbf{s}}_L$ and $\underline{\mathbf{s}}_S$ and the dominant physiological noise terms are $\underline{\mathbf{p}}_L$ and $\underline{\mathbf{p}}_S$. It is essential to suppress them by accurate denoising to enable the estimation of stimulus-evoked variations and, finally, of the vector of HR samples $\underline{\mathbf{h}}$.

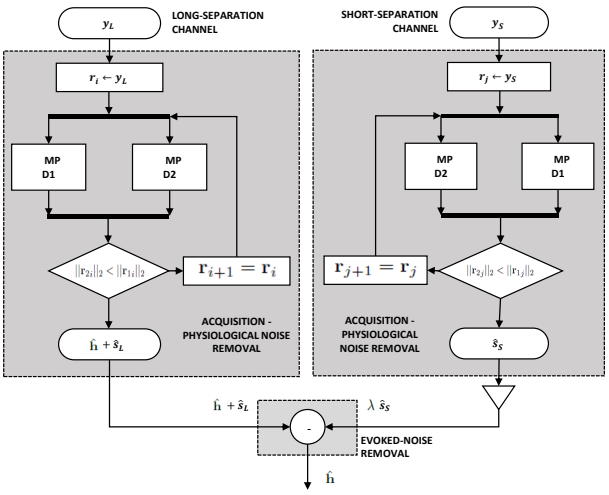


Figure 1: Logical structure of the proposed denoising and estimation algorithm.

An overcomplete dictionary $\mathbf{D}$ is a $N \times K$ matrix such that the linear relationship: $\underline{\mathbf{y}} = \mathbf{D}\underline{\mathbf{a}}$ holds, where $\underline{\mathbf{a}}$ is a *sparse* vector, which means only a small subset of its elements are non-zero. The corresponding indices form a subset of integers called the signal support $\mathcal{S}$. One may then define subvector $\underline{\mathbf{a}}_{\mathcal{S}} = [a_m]_{m\in\mathcal{S}}$, submatrix $\mathbf{D}_{\mathcal{S}} = [\mathbf{D}[\cdot, m]]_{m\in\mathcal{S}}$ and write:

$$\underline{\mathbf{y}} = \mathbf{D}_{\mathcal{S}}\underline{\mathbf{a}}_{\mathcal{S}} \tag{3}$$

showing that a signal vector $\underline{\mathbf{y}}$ can be described by the linear combination of a *small* number of columns of $\mathbf{D}$.

Dictionaries provide the signal modelling support enabling the separation of HR from physiological background noise and from evoked superficial components. For this purpose, we combine two dictionaries, $\mathbf{D}_1$ and $\mathbf{D}_2$. The former refers to physiological background components, while dictionary $\mathbf{D}_2$ is employed to estimate the cerebral HR and stimulus-evoked physiological responses. Accordingly, acquired fNIRS segments $\underline{\mathbf{y}}_{\mathrm{L}}$ and $\underline{\mathbf{y}}_{\mathrm{S}}$ are decomposed as:

$$\underline{\mathbf{y}} = \mathbf{D}_{\mathbf{1}\mathcal{S}_1}\underline{\mathbf{a}}_{1\mathcal{S}_1} + \mathbf{D}_{\mathbf{2}\mathcal{S}_2}\underline{\mathbf{a}}_{2\mathcal{S}_2}, \tag{4}$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are the supports for the two dictionaries. Only the second term on the right-hand side contains stimulus-evoked signal components.

As illustrated in Fig. 1, signals acquired from the LS and SS channels are processed separately. This allows to obtain an estimate of the task-evoked superficial response $\hat{\underline{\mathbf{s}}}_S$, whereas in the LS channel the hemodynamic response cannot be separated from the superposed evoked component, the resulting estimate being $\hat{\underline{\mathbf{x}}} = \underline{\mathbf{h}} + \underline{\mathbf{s}}_L$.

The two components $\underline{\mathbf{s}}_L$ and $\underline{\mathbf{s}}_S$ can be referred almost exactly to the same source, except for a scaling

factor that is needed to account for the different channel paths. Given the estimate $\hat{\underline{\mathbf{s}}}_S$, when this factor λ is known one has $\hat{\underline{\mathbf{s}}}_L = \lambda\hat{\underline{\mathbf{s}}}_S$ and a correction can then be applied to the LS channel estimate to finally yield the HR estimate:

$$\hat{\underline{\mathbf{h}}} = \hat{\underline{\mathbf{x}}} - \hat{\underline{\mathbf{s}}}_L = \hat{\underline{\mathbf{x}}} - \lambda\hat{\underline{\mathbf{s}}}_S. \tag{5}$$

Single-trial estimation variability is still high as a result of bias due to mismatch with actual HR shapes and of possible denoising inaccuracies. The final stage of the algorithm involves averaging over multiple trials within an experiment, which is essential to produce an accurate estimate of the hemodynamic response.

2.2. Sparse vector estimation

Dictionary-based signal analysis centers on finding a sparse solution to a matrix-vector equation. The problem can be formally expressed as:

$$\hat{\underline{\mathbf{a}}} = \arg\min_{\underline{\mathbf{a}}} ||\underline{\mathbf{a}}||_0 \quad \text{subject to:} \quad ||\underline{\mathbf{y}} - \mathbf{D}\underline{\mathbf{a}}||_2 < \epsilon \tag{6}$$

where ϵ is a threshold value associated to the energy of the residual $\underline{\mathbf{r}} = \underline{\mathbf{y}} - \mathbf{D}\hat{\underline{\mathbf{a}}}$. To solve this non-linear problem we employ a simple iterative "greedy" algorithm, known as *matching pursuit* (MP) [16].

After initializing the signal support to the empty set, $\mathcal{S} = \varnothing$ and the signal estimate to $\hat{\underline{\mathbf{y}}} = \underline{\mathbf{0}}$, the algorithm can be summarized as the iterative application of the following steps:

(i) compute $\underline{\mathbf{r}} = \underline{\mathbf{y}} - \hat{\underline{\mathbf{y}}}$, then find the dictionary index:

$$m^* = \arg\max_m |\underline{\mathbf{d}}_m^T\underline{\mathbf{r}}|^2 \quad \text{where:} \quad \underline{\mathbf{d}}_m = \mathbf{D}[\cdot, m]. \tag{7}$$

(ii) accordingly update the signal support: $\mathcal{S} = \mathcal{S} \cup m^*$ and the dictionary submatrix $\mathbf{D}_\mathcal{S}$;

(iii) compute a new amplitude estimate:

$$\hat{\underline{\mathbf{a}}}_\mathcal{S} = \left(\mathbf{D}_\mathcal{S}^T\mathbf{D}_\mathcal{S}\right)^{-1}\mathbf{D}_\mathcal{S}^T\underline{\mathbf{y}}\,; \tag{8}$$

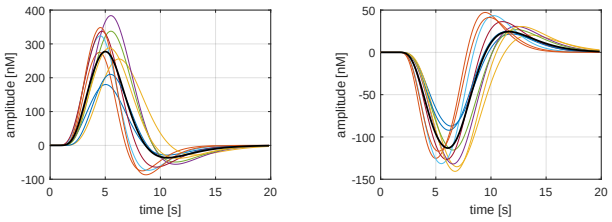
(iv) calculate the new signal estimate: $\hat{\underline{\mathbf{y}}} = \mathbf{D}_\mathcal{S}\hat{\underline{\mathbf{a}}}_\mathcal{S}$.

Search iterations can be stopped when the current approximation error satisfies condition $||\underline{\mathbf{r}}||_2 < \epsilon$ in (6), otherwise an *a priori* constraint on the cardinality of support $\mathcal{S}$ may be defined.

2.3. Overcomplete dictionaries

We describe the physiological noise components $\underline{\mathbf{p}}_L$ and $\underline{\mathbf{p}}_S$ as $p(t)$, by samples of a Taylor-Fourier multi-frequency model [20]:

$$p(t) = \sum_{j=1}^{J} \Re\left[\mathcal{X}_j(t) \cdot e^{2\pi f_j t}\right] \tag{9}$$



(a) HbO hemodynamic responses

(b) HbR hemodynamic responses

Figure 2: HR templates for HbO and HbR (black line) obtained by averaging multiple individual responses (coloured lines).

where $\mathcal{X}_j(t)$ is a second-order polynomial with complex-valued coefficients. Although more complex than the commonly employed sum-of-sinewaves, this model is more accurate and allows to track possible modulation effects induced by physiological variations, that may occur in a typical segment length. Its effectiveness in denoising was proven in [17], where construction of the relevant dictionary, indicated in this work as $\mathbf{D}_1$, is also discussed.

A hemodynamic response $\underline{\mathbf{h}}$ can be described by samples from a linear combination of gamma functions, in particular a double-gamma model is frequently employed:

$$h(t) = \kappa_1 \cdot \Gamma_{k_1}(t, \tau_1, \rho_1) - \kappa_2 \cdot \Gamma_{k_2}(t, \tau_2, \rho_2) \tag{10}$$

with:

$$\Gamma_k(t, \tau, \rho) = \frac{1}{k!\tau}\left(\frac{t-\rho}{\tau}\right)^k e^{-\left(\frac{t-\rho}{\tau}\right)} \cdot u(t-\rho) \tag{11}$$

where $u(t)$ is the unit-step function [18, 19]. Parameters τ and ρ refer, respectively, to the width of each gamma term and to the distance of its starting point from the stimulus onset. The factor k determines responsiveness in terms of peak latency and slope of edges.

Stimulus-evoked variations are known to be highly correlated with the hemodynamic response, therefore we model them by a similar expression, with possibly different parameter values. We assume that $\underline{\mathbf{x}}$ is still well described by the same model, which enables the use of a common dictionary for the SS channel and the LS channel.

The double-gamma expression (10) provides a six-parameter model, where a non-linear relationship exists between parameters $\tau_1, \tau_2, \rho_1, \rho_2$ and the corresponding waveform. The columns of dictionary matrix $\mathbf{D}_2$ represent possible changes in the basic waveform shape determined by variations of these non-linearly related parameters, for which a finite grid of values is considered. On the other hand parameters κ_1 and κ_2,

Table 1: Template parameters for double-gamma model (12).

	τ_1	τ_2	k_1	k_2	β	T_1	Δ_T
HbO	1 s	1 s	5	5	0.5	0.5 s	3 s
HbR	1 s	1 s	5	5	0.5	1.5 s	4 s

having a linear relationship, may be associated with vector $\underline{\mathbf{a}}$.

In general, a dictionary might have to account for a large number of combinations in the ranges of interest, making its column size considerably large. We simplified the construction of $\mathbf{D_2}$ by referring to an average shape template, relying on the assumption that HR shape does not vary much among subjects [21]. Figure 2 shows several HRs obtained from curve emulating those of real subjects and characterized by different combinations of parameter values, with average shapes for HbO and HbR shown by black lines. By referring to average shape as a template, width parameters τ_1 and τ_2 in (10) assume pre-determined values. Likewise, the time difference $\Delta_\rho = \rho_2 - \rho_1$ and the amplitude ratio $\beta = (\kappa_2/\kappa_1)$ are fixed.

HR can thus be described by an equation that only depends linearly on amplitude α and non-linearly on time delay ρ_1:

$$\begin{aligned} h(t) &= \alpha \cdot [\Gamma_{k_1}(t, \tau_1, \rho_1) - \beta \cdot \Gamma_{k_2}(t, \tau_2, \rho_1 + \Delta_\rho)] \\ &= \alpha \cdot d(t, \rho_1) \end{aligned} \tag{12}$$

Each column of $\mathbf{D_2}$ represents the basic shape defined by $d(nT_s, \cdot)$, time-shifted to a different position. Since template shapes may differ for HbO and HbR, two versions of the dictionary have been created, respectively $\mathbf{D}_{2,O}$ and $\mathbf{D}_{2,R}$. Relevant parameters are reported in Table 1. The resulting dictionary structure is shown for HbO in figure 3, where the time step between elements is equal to one sampling interval T_s. Only a limited set of time shifts needs to be considered and a maximum variation of ± 3 s from the expected HR peak position has been considered.

2.4. Single-trial denoising and reconstruction

Figure 1 shows that in our method two MP algorithms run in parallel, one using dictionary $\mathbf{D_1}$, the other employing $\mathbf{D_2}$. The two MP parts are run independently and their outcomes are compared at the end of each iteration. We distinguish the residuals of the two parts, that are the j-th MP iteration *outputs*, indicating them by $\underline{\mathbf{r}}_{1j}$ and $\underline{\mathbf{r}}_{2j}$, respectively. As evidenced in figure 1, the *input* to the next iteration will then be common for *both* parts and selected according to:

$$\underline{\mathbf{r}}_{j+1} = \arg\min \left[||\underline{\mathbf{r}}_{1j}||_2,\ ||\underline{\mathbf{r}}_{2j}||_2 \right] \tag{13}$$

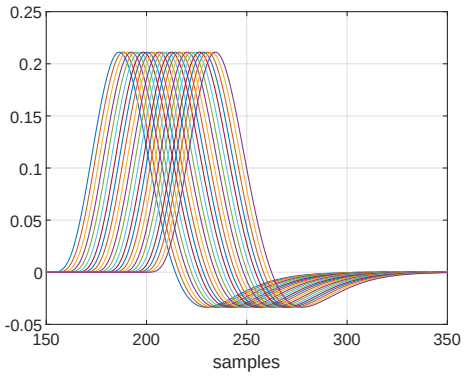


Figure 3: Structure of the HbO dictionary. Fixed template parameters are: $\tau_1 = \tau_2 = 1$ s; $\Delta_\rho = 3$ s; $\beta = 0.5$. Elements are normalized to have unit area, with time shifts covering a range of 6 s.

Comparison of residuals may yield the following cases:

- $||\underline{\mathbf{r}}_{1j}||_2 < ||\underline{\mathbf{r}}_{2j}||_2$ – part of the physiological noise is effectively modeled by $\mathbf{D_1}$, therefore $\underline{\mathbf{r}}_{j+1} = \underline{\mathbf{r}}_{1j}$ and the algorithm moves on to the next iteration. $\mathbf{D_2}$ is usually ineffective as long as physiological noise remains the dominant component in $\underline{\mathbf{r}}_j$, therefore support $\mathcal{S}_1$ has cardinality j, whereas it is still $\mathcal{S}_2 = \varnothing$;
- $||\underline{\mathbf{r}}_{2j}||_2 \leq ||\underline{\mathbf{r}}_{1j}||_2$ – it can be assumed that in the first $j-1$ iterations $\underline{\mathbf{y}}$ was denoised from the physiological component, to the point that $\underline{\mathbf{r}}_j$ has been better modeled by a column of $\mathbf{D_2}$. Both supports $\mathcal{S}_1$ and $\mathcal{S}_2$ are now non-empty sets and the latter has column index m_2^* as its single element.
  The algorithm stops, yielding the stimulus-related component estimate:

$$\begin{aligned} &\underline{\mathbf{d}}_2(m_2^*) \cdot \hat{\alpha}, \qquad \text{with:} \quad \underline{\mathbf{d}}_2(m_2^*) = \mathbf{D}_2[\cdot, m_2^*] \\ &\text{and:} \quad \hat{\alpha} = \frac{1}{\underline{\mathbf{d}}_2(m_2^*)^T \underline{\mathbf{d}}_2(m_2^*)} \cdot \underline{\mathbf{d}}_2(m_2^*)^T \underline{\mathbf{y}}. \end{aligned} \tag{14}$$

  For an SS channel, (14) represents the evoked superficial response estimate $\hat{\underline{\mathbf{s}}}_S$. For a LS channel it yields $\hat{\underline{\mathbf{x}}}$, the estimated superposition of the desired hemodynamic response and evoked superficial response;
- if condition $||\underline{\mathbf{r}}_{2j}||_2 \leq ||\underline{\mathbf{r}}_{1j}||_2$ never occurs, the algorithm stops, still with $\mathcal{S}_2 = \varnothing$, when constraint $||\underline{\mathbf{r}}_j||_2 < \epsilon$ is satisfied. It is then assumed that a stimulus-related response is not present within the segment.

This implementation has been preferred to a classic hybrid dictionary approach for the main reason that waveform template (12), being based on gamma functions, is general enough to be applicable also

to physiological noise modelling. With a single hybrid dictionary $\mathbf{D} = [\mathbf{D_1}\ \mathbf{D_2}]$ there would be a non-negligible probability that, based on the single criterion (7), a column from $\mathbf{D_2}$ is occasionally selected as more effective than the Taylor-Fourier model for physiological noise. Of course, this would prevent reliable denoising and HR estimation. In this algorithm the actual choice between the two candidate MP iteration outcomes is based instead on (13). This significantly reduces the chance of crosstalk between the two dictionaries.

Tuning of parameter ϵ is very important for the success of the algorithm. If too high, algorithm iterations may terminate before a double-gamma can be found, even when an evoked response is actually present. This causes a false negative (FN) error, whereas if ϵ is too low a false positive (FP) may occur, for instance, if $\mathbf{D_2}$ happens to model the residual physiological noise better than $\mathbf{D_1}$. Analysis for tuning parameters showed that the threshold value that yields the best trade off between FP and FN errors is $\epsilon = 0.45$ corresponding to a peak double-gamma value of 0.5 nM, that is very low compared to typical HR magnitudes but suitable for evoked response detection, in particular of the smaller component $\underline{\mathbf{s}}_S$.

2.5. Evoked-noise removal

The single-trial HR estimate $\underline{\hat{\mathbf{h}}}$ is obtained from estimates of stimulus-evoked responses in both channels, according to (5). As noted in 2.1, this requires the determination of a scale factor λ. For this purpose we consider the part of the signal acquired during a subject *resting state*, that is when no stimulus-related responses are present in the data and determine λ by linear regression of $y_{\mathrm{L}}(t)$ on $y_{\mathrm{S}}(t)$.

Although the computation is the same as in [12], our choice of a specific portion of the acquired signals ensures that good correlation exists between the LS and SS channel, improving the estimation of λ. The effect of correction (5) is demonstrated by figure 4, that shows HbO and HbR single-trial estimates obtained from one of the simulations described in Section 3. Results are representative of typical behaviour and the improvement is particularly significant in the HbO case, where relative peak amplitude deviation from the reference shape improves from 33% to 4%. Poorer accuracy is shown for the HbR response, that is smaller, yet relative peak amplitude deviation is still nearly halved, dropping from 57% to 28%.

2.6. Multi-trial averaging

A number of factors combine in producing a rather high variability of single-trial estimates. In first place, since even a single subject response may indeed vary slightly

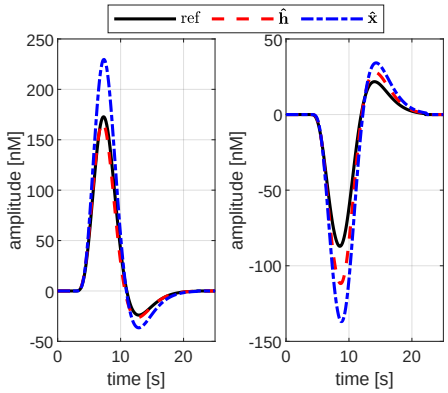


Figure 4: Effect of correction for evoked superficial responses in single-trial estimates for HbO (left) and HbR (right).

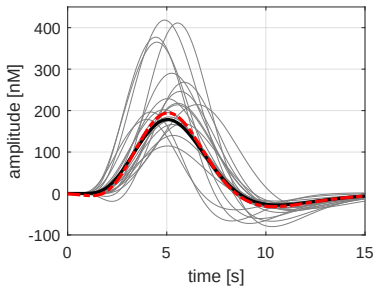
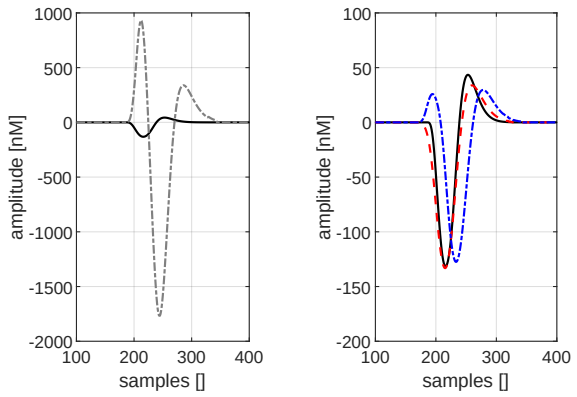


Figure 5: Comparison between the nominal HbO (black solid line) and the estimated one (red dashed line). The other lines are the single-trial estimates.

during multiple trials, the use of a single representative waveform shape in dictionary $\mathbf{D}_2$ implies some degrees of bias in many estimates. Furthermore, physiological noise may not be totally removed by the algorithm and this may contribute to inaccuracies. Typical results of a simulated experiment involving N_T trials are presented in figure 5, where single estimates are plotted together with their average (red line), that is compared to the reference waveform shape. This shows how averaging of the N_T single-trial estimates is essential for the accurate determination of a hemodynamic response.

It is important to remember that single estimates might be occasionally unreliable and deviate significantly from the actual signal shape. One such occurrence is presented in figure 6(a), where the reference shape of a HbR response (black line) is compared with a single-trial estimate (dashed grey line). The averaged estimate is significantly improved by discarding the inaccurate single-trial estimate, as figure 6(b) shows by comparing the average over N_T trials (blue line), for which $\mathrm{E_{HR}} = 125\%$, with the average over $N_T - 1$ trials from which the inaccurate estimate has been deleted



(a) Wrong single-trial HR estimate (grey dashed line)

(b) Averaged estimates (dashed lines): including all trials (blue), less the discarded trial (red)

Figure 6: Effect of a wrong estimate on averaging: a) single HbR hemodynamic response estimate; b) averaged estimate. The reference response is the continuous black line in both cases.

(red line), where $\mathrm{E_{HR}} = 6\%$.

Averaging over multiple trials should therefore be applied selectively, preliminarily discarding inaccurate individual estimates. To decide when a single-trial estimate can be considered acceptable, a criterion is needed. We refer to the *median* of the maximum absolute peak values of single-trial HR estimates, m_h, that is considered because, in the presence of occasional large deviations, it may provide a more robust statistic.

Let $\hat{\underline{\mathbf{h}}}_i$ be the estimate from the i-th trial. We retain only the single-trial estimates for which:

$$\max_n |h_i(nT_s)| < \gamma \cdot m_h \qquad i = 1, \ \ldots, \ N_T \qquad (15)$$

where coefficient γ determines the acceptance range. We verified that, by setting $\gamma = 2.5$, acceptance ratio is approximately 90% for HbO, independent of the number of trials. The ratio drops to 80% for HbR, where signal-to-noise ratio is poorer.

2.7. Simulated Data Set

Simulated data are introduced with the main purpose of validating the algorithm by determining its robustness to variations in the levels and composition of the noise part of the signal. While a synthesized signal must replicate as best as possible the physiological characteristics of a real fNIRS signal, it should be remembered that it is created in accordance with the signal model, thereby satisfying by default our assumptions about relationships among its components.

Nevertheless, testing on simulated data is essential to understand possibilities and basic limitations, properly tune the algorithm and characterize its performance. We took advantage of the fact that synthetic signals allow to test the algorithm in a greater variety of conditions, simulating in particular intra-subject and inter-subject variability to an extent that would be harder to observe in real situations. To ensure statistical stability of our performance indications, Monte Carlo simulations reproduced a total of 1000 experiments, emulating as many different subjects.

Table 2: Mean±standard deviation of physiological component frequencies and amplitudes.

Component	**Frequency [Hz]**	**Amplitude [nM]**
Very low frequency	$f_1 = .002 \pm .0001$	$A_1 = 700 \pm 100$
Low frequency	$f_2 = .01 \pm .001$	$A_2 = 700 \pm 100$
Vasomotor	$f_3 = .07 \pm .04$	$A_3 = 400 \pm 10$
Respiratory	$f_4 = .2 \pm .03$	$A_4 = 200 \pm 10$
Cardiac	$f_5 = 1.1 \pm .1$	$A_5 = 400 \pm 10$

Real experiments in Section 2.8 refer to a finger-tapping task where up to 40 trials can be averaged. Accordingly, all simulated experiments also reproduce a sequence of $N_T = 40$ trials, where LS and SS signals are generated for both HbO and HbR. Following the guidelines in [21], the peak amplitude for HbO is in the range 160-380 nM and latency varies between 4 s and 6 s. For HbR, simulated peak amplitude is between 80 nM and 140 nM, with latency in the range 5-7 s.

To create inter-subject variability, different parameters were employed in the double-gamma model (10), while intra-subject variability was represented by the introduction between trials of small variations in peak amplitude, latency and shape. This allowed to validate the applicability of the template approach in HR estimation.

Evoked systemic noise, that is correlated with HR, was modeled using the same shape template $d(nT_s, \cdot)$ defined in (12). In simulations it is further assumed that both $\underline{\mathbf{s}}_L$ and $\underline{\mathbf{s}}_S$ are time-aligned to HR.

Physiological noise was synthesized according to (9), with the frequency, amplitude and phase of each component varying between and within acquisitions. To simulate the non-stationarity of disturbances, random points in an acquisition are chosen to change the parameter values. The transition is not abrupt but gradual, to better replicate real behavior. Ranges for model parameters, based on estimates from real data, are reported in Table 2.

Finally, measurement noise was generated as white Gaussian noise with mean value 400 nM and standard deviation 180 nM. A different random sequence of noise samples is employed for each experiment. Figure 7 shows an example of synthesized acquisition from an LS channel, together with the "true" reference

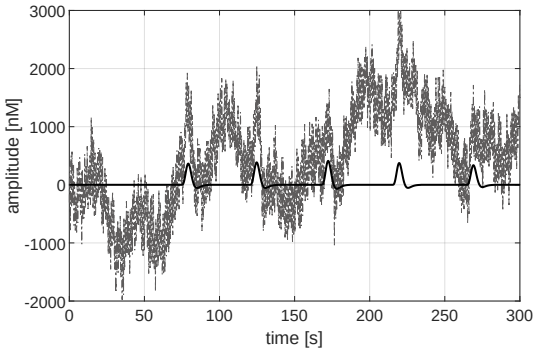


Figure 7: Comparison between $y_{\mathrm{LS}}(t)$, in grey, true HR $v(t)$ in black.

HR trace, evidencing how physiological noise is predominant and much larger than HRs.

It should be noted that signal preconditioning is applied to eliminate components outside the hemodynamic response frequency band, before the algorithm starts. We employ a 100th-order bandpass Butterworth digital filter with cut-off frequencies 0.01 Hz and 0.3 Hz, that also removes possible signal offsets.

2.8. Real and Semi-Simulated Data Set

Real data were acquired at the Department of Developmental Psychology of the University of Padova using a multichannel frequency domain NIR spectrometer (ISS Imagent, Champaign, Illinois, USA) equipped with 40 laser diodes (20 emitting light at 690 nm and 20 at 830 nm) and 4 photo-multiplier tubes. The 4 detectors and 16 sources are arranged in two patches, one located in the frontal area and the other located in the parietal area, both centered on the midline, which means the disposition is symmetrical for the two hemispheres. The measuring instrument provided a total of 8 frontal and 8 parietal LS channels (length: 3 cm) and 4 SS channels (one for each hemisphere and patch – length: 0.7cm). Each channel measures the concentration changes of HbO and HbR with a sampling frequency of 7.8125 Hz. Further information about the positioning of optical sensors and the experimental set up are provided in [7].

The real data set consists of 10 acquisitions referred to the same number of participants, that refer to a finger-tapping task. This is probably one of the most popular paradigms, generally employed in studies aimed to validate signal processing algorithms since it provides robust and localized HR. The task is to perform a right (t1) or left (t2) finger tapping, according to the visual stimulus presented on a monitor in the form of an arrow pointing to the right or left. Each participant performed a total of 40 trials for t1 and 40 for t2. Subjects were instructed to relax during the first part of the acquisition, that corresponds to a resting state interval and were then presented with a random series of stimuli. An interstimulus interval (ISI) ranging from 12 s to 15 s elapsed between consecutive trials.

The semi-simulated data set was created by combining synthesized hemodynamic and evoked superficial responses with real physiological noise data, obtained during the acquisition of resting state intervals only.

It is important to notice that both hemodynamic response and evoked superficial response need to be simulated, since neither is supposedly present in the long and short channels when a subject is in resting state. As only this part of the signal is synthesized according to our modelling assumptions, testing by a semi-simulated data set allows to check that denoising of the physiological component using the Taylor-Fourier dictionary $\mathbf{D}_1$ is accurate enough to enable the subsequent HR response estimation.

The synthesized part of the signals in the semi-simulated data set was again generated by Monte Carlo simulation, while noise components are randomly selected segments of resting state recordings, also referring to 10 different subjects, that only contain physiological noise.

2.9. Metrics

Performance metrics referred to the synthesized reference HR provide a quantitative assessment for both simulated and semi-simulated data sets. For each experiment we consider the averaged HR estimate $\hat{\underline{\mathbf{h}}}_{\mathrm{HR}}$, its peak amplitude $\hat{A}$ and the latency $\hat{L}$. Given the reference HR shape $\underline{\mathbf{h}}_{\mathrm{HR}}$, with the corresponding peak amplitude A and latency L, reconstruction accuracy is assessed by means of the following quantitative metrics:

$$\mathrm{E}_{\mathrm{HR}} = \frac{\|\hat{\underline{\mathbf{h}}}_{\mathrm{HR}} - \underline{\mathbf{h}}_{\mathrm{HR}}\|^2}{\|\underline{\mathbf{h}}_{\mathrm{HR}}\|^2} \cdot 100 \tag{16}$$

$$\mathrm{E}_A = \frac{|\hat{A} - A|}{|A|} \cdot 100 \tag{17}$$

$$\mathrm{E}_L = |\hat{L} - L| \tag{18}$$

The first index E_{HR} gives the percent root-mean-square deviation of the HR estimate from the reference and is an indication of overall agreement between pulse shapes. Indices E_A and E_L quantify the error in estimating the two most important HR parameters.

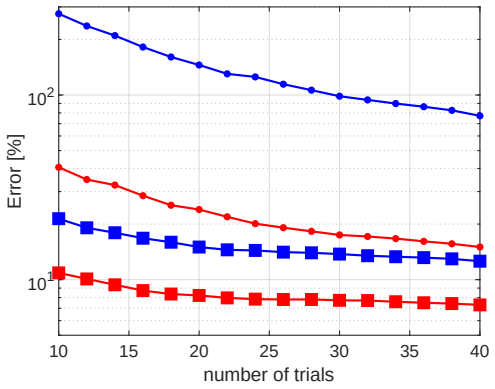


Figure 8: Dependence of mean E_{HR} on the number of trials in an experiment for HbO (red) and HbR (blue) response estimates. Smaller dots refer to averaging over N_T trials, squares indicate selective averaging over N_T^* trials.

2.10. Comparison algorithms

We compare our method with two other well-known approaches in the literature: Standard Block Averaging (SBA) and the dual-detector least-squares approach (DDLS) presented in [12].

With SBA the signal is pre-filtered by a bandpass filter with cutoff frequencies of 0.01 and 0.5 Hz to remove high frequency noise. The filtered signal is segmented according to the occurrence of the stimuli, in order to have only one HR in each segment, then segments are averaged to reduce superposed noise. Physiological noise is assumed to be independent and uncorrelated in different segments.

DDLS aims at removing both physiological and evoked noise by subtracting the SS channel signal from the LS channel output, after scaling by a coefficient obtained by least-squares regression between the two channels.

3. Results

3.1. Performance analysis with simulated data

We compare results obtained by unrestricted averaging (that is, considering all N_T trials) with those obtained by selective averaging, that is, considering only the N_T^* trials that passed (15), to investigate the effect of the trials selection on accuracy. Furthermore, we assess the dependence on the number of trials in an experiment, starting from $N_T = 40$ for comparison with the real case and progressively lowering that number to a minimum of 10. Below this number of trials result variability becomes too large.

Plots in figure 8 show mean values of E_{HR} for HbO and HbR responses, with averages computed over either N_T or N_T^* trials. To allow a fair comparison, the abscissa always reports the total number of trials in an experiment. Therefore, even when results are reported for selective averaging this is always the value of N_T. It can be noticed that by discarding inaccurate single estimates a very significant improvement is achieved. In particular, mean E_{HR} improves by about an order of magnitude for HbR.

In the following we shall then refer to the averaged HR estimate defined by:

$$\hat{\underline{\mathbf{h}}}_{HR} = \frac{1}{N_T^*} \sum_{i=1}^{N_T^*} \hat{\underline{\mathbf{h}}}_i \tag{19}$$

where $N_T^* \leq N_T$ is the actual number of accepted single-trial estimates. Acceptance rates are high enough to make the variability of N_T^* negligible when the outcomes of all 1000 simulated experiments are analyzed.

For selective averaging, plots of mean E_{HR} are repeated in figure 9, with vertical bars added to show intervals of ± 1 standard deviation. Particularly for HbO it is interesting to note that curves tend to level-off, in terms of both mean value and standard deviation, with N_T lower than 40. This suggests the possibility to consider shorter experiments involving fewer trials, which might be of interest when subject fatigue can become an issue.

Similar plots are reported in figure 10 for E_A and E_L. They also refer to results of Monte Carlo simulation involving 1000 synthesized experiments, therefore are considered to provide reliably stable statistical information. With all metrics, performance for HbO estimation is better than for HbR which is expected, as already remarked, since for a given noise level the response amplitude of HbR is lower, making estimation more challenging. A levelling-off of curves similar to that already noted in figure 9 is apparent here.

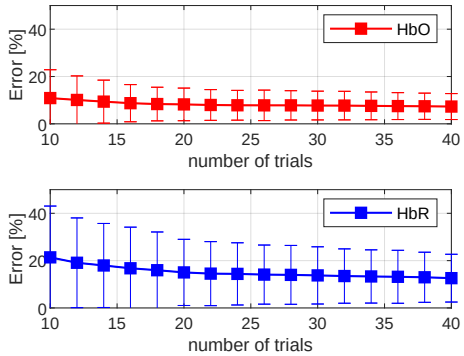
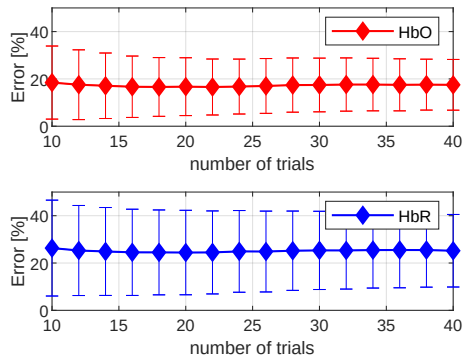


Figure 9: Mean $E_{HR} \pm 1\times$std. dev. versus number of trials in an experiment.

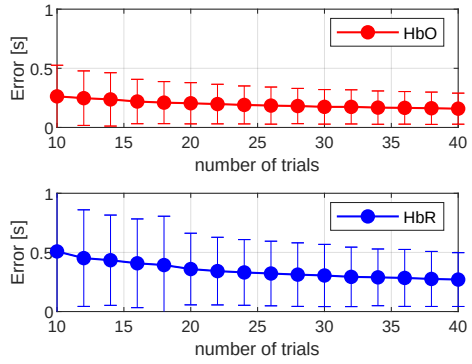
Table 3: Performance metrics mean ±standard deviation – Simulated data set

	trials	$\mathbf{E_{HR}}$ [%]	$\mathbf{E}_A$ [%]	$\mathbf{E}_L$ [s]
HbO	**20**	8.2 ± 6.7	16.7 ± 12.2	0.2 ± 0.2
	40	7.2 ± 5.5	17.5 ± 10.7	0.2 ± 0.1
HbR	**20**	15.0 ± 14.0	24.4 ± 17.8	0.3 ± 0.3
	40	12.5 ± 10.0	25.2 ± 15.3	0.3 ± 0.2

Numerical values are summarized for convenience in Table 3, where results are reported for experiments comprising 40 trials, as in the experimental part of this work, as well as for experiments involving half that number. Simulated results evidence no significant improvements in accuracy with the longer experiments. This confirms that the single-trial denoising and estimation algorithm effectively enhances HR estimation.



(a) mean $E_A \pm 1\times$std. dev.



(b) $E_L \pm 1\times$std. dev.

Figure 10: Accuracy of HR amplitude and latency estimates versus number of trials in an experiment.

3.2. Comparisons and discussion for semi-simulated data

The analysis we presented so far allowed to optimize and test our method under a wide variety of realistically simulated test conditions. We now discuss experimental results obtained with semi-simulated data, for which metrics can still be used to summarize results.

Comparison results obtained on the semi-simulated data set are reported in Table 4: for each experiment the HR is estimated from the standard LS channel signal, both for HbO and HbR. SBA does not employ information from the reference SS channel, whereas both DDLS and our proposed method make use of it.

As SBA employs only the LS channel, the stimulus-evoked superficial response is not suppressed from the estimates. This is the reason why reported values for E_{HR} and E_A are particularly high. Since the error is systematic, varying the number of trials has little effect and the difference between 20 and 40 trials is nearly irrelevant for E_A.

With the DDLS approach HR amplitude estimation is more precise, but still affected by part of the evoked noise. Table 4 shows that the method perfoms better than SBA, underlining the importance of the SS channel for noise suppression. On the other hand, DDLS assumes that correlation between the LS and SS channel is high *within* the segment of interest and its accuracy degrades if this assumption tends to fail.

With the proposed method the contribution due to the stimulus-evoked superficial response is effectively rejected, therefore results related to shape and amplitude accuracy (respectively, E_{HR} and E_A) outperform the other methods. A different consideration lies to be made for latency error E_L. Evoked noise is correlated and time-aligned with the hemodynamic response, since both have the same source. For this reason it has minimal influence on the latency estimate. In this case, the good results achieved by the proposed method can be explained mainly as a consequence of very effective denoising of the physiological component.

It should be noticed that with our approach performance metrics do not change considerably if the number of trials is reduced from 40 to 20. This confirms the possibility of reducing acquisition time, as already mentioned.

3.3. Discussion for real data

Due to the absence of a reference hemodynamic response, quantitative assessment of algorithm performances is not feasible with real data. However, it is still possible to compare estimates obtained by dif-

Table 4: Performance metrics mean ±standard deviation – Semi-Simulated data set.

	trials	$\mathbf{E_{HR}}$ [%]		$\mathbf{E}_A$ [%]		$\mathbf{E}_L$ [s]	
		HbO	HbR	HbO	HbR	HbO	HbR
SBA	**20**	667 ± 895	490 ± 392	129 ± 31	135 ± 40	2.3 ± 6.6	0.1 ± 0.1
	40	400 ± 292	330 ± 190	123 ± 39	140 ± 27	0.6 ± 3.4	0.1 ± 0.1
DDLS	**20**	339 ± 1004	121 ± 110	51 ± 31	59 ± 34	4.6 ± 6.3	1.2 ± 2.8
	40	288 ± 909	105 ± 131	53 ± 31	54 ± 31	3.3 ± 5.4	0.9 ± 1.8
Proposed Method	**20**	59 ± 90	88 ± 77	41 ± 29	72 ± 45	0.6 ± 1.1	0.2 ± 0.2
	40	34 ± 35	78 ± 75	37 ± 28	65 ± 48	0.5 ± 0.9	0.1 ± 0.2

ferent methods. Further considerations are possible, based on the expected cerebral response behavior for the assigned task. Specifically, a finger tapping task should produce a higher response in the contralateral brain hemisphere, and a lower or even absent response on the ipsilateral hemisphere. Thus, an assessment can be obtained by comparing results for couples of fNIRS channels referred to either brain hemisphere of the same subject.

As discussed in 2.4, the algorithm has been designed to minimize chances that a double-gamma template could be erroneously associated to a physiological noise component. Nevertheless this may occur in particular, referring to (13), when $||\underline{\mathbf{r}}_{1_j}||_2 \cong ||\underline{\mathbf{r}}_{2_j}||_2$. This situation may present itself when there is actually no stimulus-related component in the analyzed signal segment. In this case the averaged estimate might produce a false positive (FP).

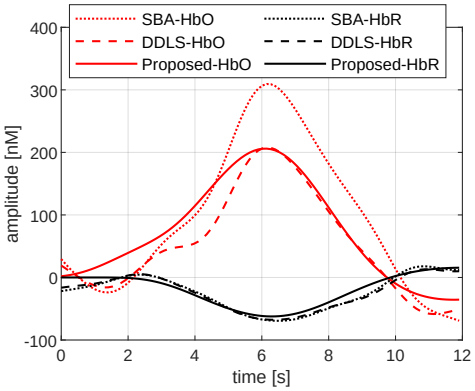
To analyze the algorithm response under this condition, another set of tests involving Monte Carlo simulations was carried out with no HRs present in synthesized fNIRS signals. The algorithm was expected to produce at most very small estimates and the analysis yielded the following indications (mean ±1×standard deviation) for the peak values of the averaged estimates:

HbO: 24 ± 13 nM HbR: 17 ± 9 nM

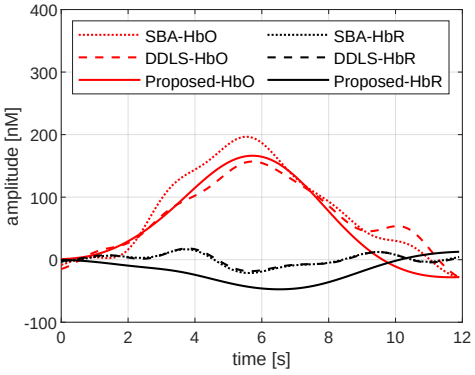
Determining these values allowed to rationally set a detection limit, below which it can be assumed that brain activation did not occur and a hemodynamic response is actually not present.

Results presented in figures 11 and 12 refer to two of the subjects that took part in the finger-tapping task experiments. They differ because selected fNIRS channels for the two subjects employ different sensor dispositions on the scalp. Similar results can be obtained for all other subjects and are not reported for conciseness.

Figure. 11 presents the averaged HR responses obtained from the two measurement channels in Subject 3. In this case optical sensors were placed over different brain hemispheres but close to the central



(a) HR, contralateral brain hemisphere



(b) HR, ipsilateral brain hemisphere

Figure 11: Real data set, Subject 3, two brain hemispheres – fNIRS channels close to central septum. Estimated HbO (red) and HbR (black) responses.

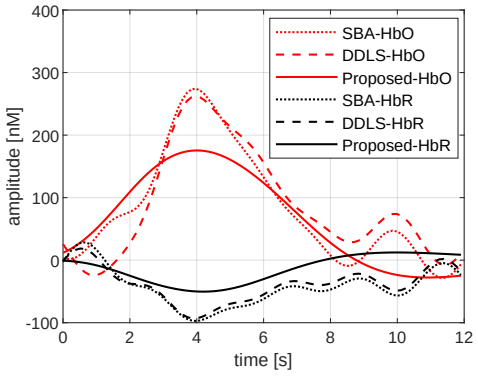
septum, so that significant activation is expected to be seen in both channels. In fact, HbO and HbR response estimates obtained by the proposed method exceed the minimum thresholds denoting brain activation for both hemispheres, with lower responses from the ipsilateral one. For the contralateral hemisphere figure 11(a) shows good agreement with the DDLS method for both HbO (red) and HbR (black) estimates, whereas SBA tends to overestimate HbO response amplitude. Such

differences agree with the results obtained from the Semi-Simulated data.

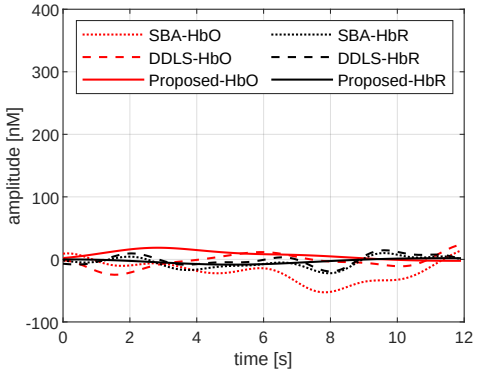
For the ipsilateral hemisphere only the proposed method enables to estimate the HbR response in agreement with expectations, that are based on the position of sensors and on the corresponding HbO response estimate (figure 11(b)).

In the fNIRS channels selected from Subject 9, sensors were placed laterally on the head, at greater distance from the septum so that each channel can be considered uniquely associated to one hemisphere of the brain. Well-differentiated responses are expected here, specifically, brain activation should be observed only in the contralateral channel. This is again confirmed by the plots of figure 12(a). In particular, whereas the proposed method correctly estimates a HbO response of approximately 180 nM and a HbR response of about 50 nM, the contralateral response appears to be overestimate with both SBA and DDLS.

In the ipsilateral hemisphere only residual oscillations are present, as shown in figure 12(b) and this means the response is absent, as expected. In this case



(a) HR, contralateral brain hemisphere



(b) HR, ipsilateral brain hemisphere

Figure 12: Real data set, Subject 9, two brain hemispheres – well-separated fNIRS channels. Estimated HbO (red) and HbR (black) responses.

the estimates obtained by the proposed method tend to be the smallest and consistently remain below the minimum thresholds for brain activation derived from our analysis. Thus, the result can be interpreted more reliably also in this case.

4. Conclusions

A tecnique to denoise fNIRS signals and accurately estimate HR has been presented in this paper. Its novelty derives from the introduction of a complete model of the acquisition system that separately accounts for each signal component in the standard LS channel and in the reference SS channel. This led to a dictionary-based estimation and denoising algorithm where two dictionaries are employed, respectively, to estimate and remove the physiological background components and to estimate task-evoked components.

The proposed method has been characterized on a synthetic data set and evaluated on both semi-simulated and real data sets showing very good results, that in many cases significantly outperform those obtained with more traditional approaches. Particular care has been taken to ensure reliability and prevent the occurrence of FN and FP errors.

The work has shown that a dictionary-based approach can significantly enhance the accuracy of single-trial estimates. This points to future developments in our research, particularly the refinement of dictionaries, either by the creation of an *ad hoc* dictionary for each subject, or by the introduction of an extended dictionary to accommodate further shape variations of the basic template. This should overcome the problem of bias due to the use of an average shape as the model template, allowing to pursue the goal of achieving reliable single trial estimation.

References

[1] Boas D A, Elwell C E, Ferrari M and Taga G 2014 Twenty years of functional near-infrared spectroscopy: introduction for the special issue *Neuroimage* **85** 1-5

[2] Wilcox T and Biondi M 2015 fNIRS in the developmental sciences *Wiley Interdisciplinary Rev.: Cogn. Sci.*, **6** 263-283

[3] Molavi B, May L, Gervain J, Carreiras M, J F Werker J Fand and Dumont G A 2014 Analyzing the resting state functional connectivity in the human language system using near infrared spectroscopy *Front. Hum. Neurosci.* **7** 1-9

[4] Irani F, Platek S M, Bunce S, Ruocco A C and Chute D 2007 Functional near infrared spectroscopy (fNIRS): an emerging neuroimaging technology with important applications for the study of brain disorders *The Clinical Neuropsychologist* **21** 9-37

[5] Hong, K. S., Naseer, N., and Kim, Y. H. 2015 Classification of prefrontal and motor cortex signals for three-class fNIRSBCI Neuroscience letters, 587, 87-92.

[6] Ge S *et al.* 2017 A brain-computer interface based on a few-channel EEG-fNIRS bimodal system *IEEE Access* **5** 208-218

[7] Scarpa F, Brigadoi S, Cutini S, Scatturin P, Zorzi M, Dell'Acqua R and Sparacino G 2013 A reference-channel based methodology to improve estimation of event-related hemodynamic response from fNIRS measurements *Neuroimage* **72** 106-119

[8] Ycel M A, Selb J, Aasted C M, Lin P Y, Borsook D, Becerra L and Boas D A 2016 Mayer waves reduce the accuracy of estimated hemodynamic response functions in functional near-infrared spectroscopy *Biomedical optics express* **7** 3078-3088

[9] Tachtsidis I and Scholkmann F 2016 False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward *Neurophotonics* **3** 031405

[10] Calderon-Arnulphi M, Alaraj A and Slavin K V 2009 Near infrared technology in neuroscience: past, present and future *Neurological research* **31** 60514

[11] Saager R B and Berger A J 2008 Measurement of layer-like hemodynamic trends in scalp and cortex: implications for physiological baseline suppression in functional near-infrared spectroscopy *J. of biomedical optics 13* 034017

[12] Saager R B and Berger A J 2005 Direct characterization and removal of interfering absorption trends in two-layer turbid media *J. Opt. Soc. Am.* **22** 1874-1882

[13] Kirilina E *et al.* 2012 The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy *Neuroimage* **61** 70-81

[14] Brigadoi S and Cooper R J 2015 How short is short? Optimum sourcedetector distance for short-separation channels in functional near-infrared spectroscopy *Neurophotonics* **2** 025005

[15] Tachtsidis I , Leung T S, Chopra A, Koh P H, Reid C B and Elwell C E 2009 False positives in functional nearinfrared topography *Oxygen Transport to Tissue XXX* 307314

[16] Mallat S G, Zhang Z 1993 Matching pursuits with time-frequency dictionaries *IEEE Trans. on Signal Processing* **41** 3397-3415

[17] Frigo G, Brigadoi S, Giorgi G, Sparacino G and Narduzzi C 2016 Measuring cerebral activation from fNIRS signals: An approach based on compressive sensing and TaylorFourier model. *IEEE Trans. on Instrum. Meas.* **65** 1310-1318

[18] Lindquist M A, Loh J M, Atlas L Y and Wager T D 2009 Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling *NeuroImage*, **45** S187S198

[19] Santosa H, Zhai X, Fishburn F and Huppert T 2018 The NIRS brain AnalyzIR toolbox *Algorithms* **11** 73

[20] de la O Serna J Synchrophasor measurement with polynomial phaselocked-loop 2015 Taylor-Fourier filters *IEEE Trans. Instrum. Meas.* **64** 328337

[21] Huppert T J, Hoge R D, Diamond S G, Franceschini M A and Boas D A 2006 A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans *Neuroimage* **29** 368-38.