bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Physiological noise modeling in fMRI based on the pulsatile component of photoplethysmograph

Michalis Kassinopoulos[1], Georgios D. Mitsis[2]

[1]Graduate Program in Biological and Biomedical Engineering, McGill University, Montreal, QC, Canada

[2]Department of Bioengineering, McGill University, Montreal, QC, Canada

Abstract

The blood oxygenation level-dependent (BOLD) contrast mechanism allows someone to non-invasively probe changes in deoxyhemoglobin content. As such, it is commonly used in fMRI to study brain activity since levels of deoxyhemoglobin are indirectly related to local neuronal activity through neurovascular coupling. However, the BOLD signal is severely affected by physiological processes as well as motion. Due to this, several noise correction techniques have been developed through the years to correct for the associated confounds. This study sought to refine model-based techniques that utilize the photoplethysmograph (PPG) signal. RETROICOR, a technique commonly used to model fMRI fluctuations induced by cardiac pulsatility was compared with a new technique proposed here, named cardiac pulsatility model (CPM), that is based on convolution filtering. Further, this study investigated whether the variations in the amplitude of the PPG pulses (PPG-Amp) covary with variations in amplitude of pulse-related fMRI fluctuations as well as with systemic low frequency oscillations (SLFOs) present in the global signal (i.e. mean fMRI timeseries averaged across all voxels in gray matter). Capitalizing on 3T fMRI data from the Human Connectome Project, CPM was found to explain significantly more variance in fMRI compared to RETROICOR, particularly for subjects that presented high variance in heart rate during the scan. The amplitude of the fMRI pulse-related fluctuations did not seem to covary with PPG-Amp. That said, PPG-Amp explained significant variance in the GS that did not seem to be attributed to variations in heart rate or breathing patterns. In conclusion, our results suggest that the techniques proposed here can model high-frequency fluctuations due to pulsation as well as low-frequency physiological fluctuations more accurately than model-based techniques commonly employed in fMRI studies.

Keywords: RETROICOR; CPM; fMRI artifacts; noise correction techniques; Cardiac pulsation; SLFOs; Global signal

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1. Introduction

Functional magnetic resonance imaging (fMRI) is a powerful neuroimaging modality that provides measurements of brain activity with great spatial coverage and resolution (Bandettini et al., 1992; Kwong et al., 1992; Ogawa et al., 1990). It has been extensively used in behavioral experiments to study brain function associated to a specific task or condition but also even during resting condition to examine the intrinsic brain functional architecture (Biswal et al., 1995; Van Dijk et al., 2010). The majority of fMRI experiments are based on the blood oxygen level dependent (BOLD) contrast that can detect changes in blood oxygenation, and specifically changes in concentration of deoxygenated hemoglobin (Hb). The main principle exploited in BOLD fMRI is that neuronal activity triggers changes in local cerebral blood flow (CBF) which, in turn, affects the concentration Hb (Iadecola, 2017; Kisler et al., 2017). As such, due to its sensitivity to levels of Hb, the BOLD signal provides an indirect measure of the underlying neuronal activity. A main challenge, however, when analyzing fMRI data is that the BOLD signal consists also of physiological-induced fluctuations as well as fluctuations due to motion which, if not accounted for, can severally diminish the detection of neural-induced signals or lead to artificial associations (Birn, 2012; Chang and Glover, 2009; Glasser et al., 2018; Power et al., 2015; Xifra-porxas et al., 2020).

The fMRI confounds induced by physiological processes and motion fall into two categories: (1) purely physiological blood-borne signals also known as systemic low-frequency oscillations (SLFOs), and (2) acquisition artifacts. Blood-borne signals are signals driven by changes in the levels of Hb in the sample being imaged which in principle can be influenced by several physiological factors. Experimentally, it has been shown that variations in heart rate (HR; Shmueli et al., 2007), levels of carbon dioxide (CO_2; Prokopiou et al., 2019; Wise et al., 2004), breathing patterns (Birn et al., 2006), as well as arterial blood pressure (Whittaker et al., 2019) give rise to low-frequency (~0.1 Hz) fluctuations in fMRI presumably due to their effects on the levels of Hb (Caballero-Gaudes and Reynolds, 2017; Liu, 2016; Murphy et al., 2013). On the other hand, acquisition artifacts are caused by any kind of motion that forces the imaged sample to move in space or perturbs the magnetic field, as these manipulations have a direct impact on the acquisition process (Caballero-Gaudes and Reynolds, 2017; Liu, 2016; Murphy et al., 2013). Acquisition artifacts may be related to bulk head motion and breathing-related chest expansion (Power et al., 2015) but also to cardiac contractions through vessel expansion in the brain vasculature and its associated tissue movement (Dagli et al., 1999).

To account for the effects of physiological processes and motion, several data-driven techniques have been proposed (Caballero-Gaudes and Reynolds, 2017). An important class of data-driven techniques involves the decomposition of the fMRI data into components either using principal or independent component analysis (Behzadi et al., 2007; Pruim et al., 2015; Salimi-Khorshidi et al., 2014) and the removal of components that are likely due to physiological fluctuations or acquisition artifacts prior to further analysis. These techniques have been shown to perform fairly well in the context of whole-brain functional connectivity, particularly for acquisition artifacts (Ciric et al., 2017; Kassinopoulos and Mitsis, 2019a; Parkes et al., 2018; Xifra-porxas et al., 2020). However, their performance on task-based studies or studies with limited field of view (e.g. Mitsis et al., 2009; Pattinson et al., 2009) has not been well addressed. Often the so-called global signal (GS), defined as the mean timeseries across all voxels in the brain (or gray matter (GM)), is removed from the data as it reflects to a large extent SLFOs (Birn et al., 2006; Chang and Glover, 2009; Falahpour et al., 2013; Kassinopoulos and Mitsis, 2019b; Shmueli et al., 2007). That said, this practice is been somewhat controversial as there is evidence that GS variations may also reflect neuronal activity (Liu et al., 2017; Murphy and Fox, 2017; Power et al., 2017). Due to this, several methods have been proposed that assess the effect of global signal regression (GSR) on connectivity measures (Carbonell et al., 2014; Falahpour et al., 2018; Nalci et al., 2019b, 2019a) as well as alternative approaches that avoid some of the limitations of GSR (Aquino et al., 2019; Carbonell et al., 2011; Glasser et al., 2018).

As the effectiveness and reliability of data-driven techniques is yet to be determined, fMRI studies often employ model-based techniques that utilize concurrent external recordings (e.g. levels of CO_2, continuous arterial blood pressure, etc) and, thus, are more certain to address a specific source of confound without the risk of removing signal of interest. The physiological recordings most commonly acquired are the photoplethysmograph (PPG) and the respiratory bellow signal due to the availability of the necessary equipment in the majority of MR units and due to

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that the transducers are well tolerable for participants. PPG and the respiratory bellow capture cardiac and breathing activity, respectively, and are commonly used to account for high-frequency fMRI artifacts due to cardiac pulsatility (~1 Hz) and breathing motion (~0.3 Hz) using a technique named RETROICOR proposed by Glover et al. (2000). In addition, these two signals are often used with convolution models to account for SLFOs driven by variations in HR (Chang et al., 2009) and breathing patterns (Birn et al., 2008).

In this study, we propose a refinement for the approach taken in RETROICOR for modelling cardiac pulsatility. Specifically, we propose the cardiac pulsatility model (CPM) that describes pulse-related fluctuations in fMRI as the convolution of a train of impulses located at the time of cardiac contractions and a cardiac pulsatility waveform (CPW). Using a cross-validation framework we compare RETROICOR with CPM in terms of variance explained in fMRI. We hypothesize that these two approaches yield similar performance when HR is relatively stable whereas CPM yields better model fit for subjects with high HR variability (HRV). Moreover, we examine whether the variations observed in the amplitude of the pulses in PPG (PPG-Amp) are also present in fMRI pulse-related fluctuations. Finally, we examine the association between PPG-Amp and the low-frequency fluctuations present in the GS. Previously, we have presented a framework for estimating scan-specific physiological response functions (PRFs) that are used to model the effect of HR and breathing pattern in SLFOs present in the GS (Kassinopoulos and Mitsis, 2019b). Doing so, someone can model SFLOs and use this timeseries as a nuisance regressor in the preprocessing of the fMRI data. Here, we investigate whether there is any benefit of considering PPG-Amp, in addition to HR and breathing activity, to model SLFOs present in the GS. Based on our results, CPM explains more variance than RETROICOR, particularly for subjects with high HRV. The amplitude of the cardiac-related pulses observed in fMRI does not covary with PPG-Amp. Our results suggest, however, that PPG-Amp explains variance in the GS not captured by measurements of cardiac or breathing activity.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2. Methodology

Unless stated otherwise, the preprocessing and analysis described below were performed in Matlab (R2018b; Mathworks, Natick MA).

2.1 Human Connectome Project (HCP) Dataset

We used resting-state scans from the HCP S1200 release (Glasser et al., 2016; Van Essen et al., 2013). The HCP dataset includes, among others, T1-weighted (T1w) images and resting-state (eyes-open and fixation on a cross-hair) data from healthy young (age range: 22-35 years) individuals acquired on two different days. On each day, two 15-minute scans were collected, one with a left-right phase encoding (PE) direction and one with a right-left PE direction. fMRI acquisition was performed with a multiband factor of 8, spatial resolution of 2 mm isotropic voxels, and a repetition time (TR) of 0.72 s (Glasser et al., 2013).

The minimal preprocessing pipeline for the resting-state HCP dataset is described in (Glasser et al., 2013). In brief, the pipeline included gradient-nonlinearity-induced distortion correction, motion correction, EPI image distortion correction and non-linear registration to MNI space. The motion parameters are included in the database for further correction of motion artifacts.

In the present work, we used the minimally-preprocessed data along with T1w images provided in volumetric MNI152 space. We considered 100 subjects which included good quality signals of photoplethysmograph (PPG) and breathing activity (respiratory bellow) in all four scans, as assessed by visual inspection. The 41 subjects considered in (Kassinopoulos and Mitsis, 2019b) were excluded from this study to examine whether some of the previous findings can be replicated in a different subset of HCP subjects.

2.2 Preprocessing and analysis of physiological signals

The detection of all peaks in PPG with good time accuracy was important for the techniques examined in this study. The timings of the peaks in PPG were used to model either the high-frequency (~1 Hz) cardiac pulsatility artifacts in fMRI or low-frequency (~0.1 Hz) physiological artifacts due to variations in HR reflected on the GS. In addition, the amplitudes of the peaks in PPG were used to model low-frequency artifacts that in principle may be unrelated to HR variations. Therefore, to facilitate the detection of peaks, PPG was initially band-pass filtered with a 2nd order Butterworth filter between 0.3 and 10 Hz. The minimum peak distance specified for peak detection varied between 0.5 and 0.9 s depending on the subject's average HR and was chosen based on visual inspection of the HR that results for a specific minimum peak distance value. For a given set of detected peaks, the HR signal was computed in beats-per-minute (bpm) by multiplying the inverse of the time differences between pairs of adjacent peaks with 60, and evenly resampled at 10 Hz. The amplitudes of the peaks, referred to later as photoplethysmographic amplitude (PPG-Amp), were also evenly resampled at 10 Hz. The resampling of HR and PPG-Amp was done using linear interpolation. Note that several subjects in HCP demonstrated a constant HR of 48 bpm which is likely due to erroneous PPG recording. None of these subjects was considered in our study. That said, many scans considered here illustrated HR traces with outliers at different timepoints. The values of HR at the timepoints of outliers were corrected using linear interpolation (for more information see Methods in (Kassinopoulos and Mitsis, 2019b)).

The breathing signal was detrended linearly and corrected for outliers using a median filter in a similar manner with (Kassinopoulos and Mitsis, 2019b). Subsequently, the breathing signal was low-pass filtered at 5 Hz with a 2nd order Butterworth filter and z-scored.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.3 Preprocessing of fMRI data

fMRI data as well as 16 nuisance regressors were first high-pass filtered at 0.008 Hz. The nuisance regressors consisted of the 6 motion parameters along with their derivatives and 4 regressors related to breathing motion (i.e. 2nd order RETROICOR). Subsequently, the nuisance regressors were removed from the fMRI data. To extract the GS of each scan, first we performed tissue segmentation on the T1w images at the MNI152 space using FLIRT in FSL 5.0.9 that generated probabilistic maps for the gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) compartments (Zhang et al., 2001). Afterwards, the GS was calculated by estimating the mean timeseries across all voxels with probability above 0.25 to belong to GM. The choice of the threshold value was done based on visual inspection while overlaying the probabilistic map of GM on the T1w images.

2.4 High-frequency cardiac fluctuations

2.4.1 RETROICOR

RETROICOR stands for RETROspective Image-based CORrection and is a technique proposed by Glover et al. (2000) for removing high-frequency fluctuations due to cardiac pulsatility and breathing motion. While similar steps are performed to remove the aforementioned confounds, different mechanisms underly them. In this study we compared the variance explained in fMRI data with RETROICOR and with respect to CPM which is a physiologically plausible model for pulsatility artifacts. Therefore, if not explicitly stated, by RETROICOR we refer to the steps related to the cardiac fluctuations.

To obtain the pulse-related RETROICOR regressors we followed the following three steps:

1. The cardiac phase was first defined based on the timings of the PPG peaks using the relation:

$$\varphi(t) = 2\pi \frac{t - T_p(t)}{T_p(t) - T_f(t)}, \qquad [1]$$

where $T_p(t)$ and $T_f(t)$ indicate the time of the nearest peaks from past and future timepoints, respectively.

2. A basis set of cosines and sines was created as follows:

$$\cos_m(t) = \cos(m\varphi)$$
$$\sin_m(t) = \sin(m\varphi) \qquad [2]$$

where m is the order of the Fourier basis set. Typically, a 2nd order of RETROICOR is employed which corresponds to four nuisance regressors. It has been suggested that while higher orders improve the fit they carry the risk of overfitting the data (Harvey et al., 2008). However, as the optimal order may depend on parameters of the fMRI pulse sequence, such as the duration of scan and TR, one of the goals of this study was to determine the optimal order for RETROICOR and CPM when applied to the resting-state fMRI data of HCP (see Section 2.6).

3. Finally, the nuisance regressors $\cos_m(t)$ and $\sin_m(t)$ were downsampled to match the fMRI acquisition rate yielding the regressors $\cos_m[n]$ and $\sin_m[n]$. Note that we use parentheses and brackets to distinguish the high and low-sampled timeseries, and not for distinguishing continuous and discrete signals.

The yielded nuisance regressors were used in the design matrix of the general linear model (GLM) in order to model the pulse-related artifacts in each voxel timeseries. Through linear regression in GLM, each nuisance regressor is assigned with a beta parameter β. In the case that only cardiac-related regressors are included in the design matrix, the voxel timeseries can be represented as:

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$y[n] = \sum_{m=1}^{M} \beta_{cos,m} \cos_m[n] + \beta_{sin,m} \, sin_m[n] + \varepsilon[n] \qquad [3]$$

where $\varepsilon[n]$ corresponds to random Gaussian error. Section 2.6 describes how the fit of RETROICOR regressors on the data was assessed and compared with CPM. Note that the beta parameters essentially define the cardiac pulsatility waveform (CPW) during a cardiac cycle. Specifically, CPW is obtained using the relation:

$$CPW(\varphi) = \sum_{m=1}^{M} \beta_{cos,m} \, cos(m\varphi) + \beta_{sin,m} \sin(m\varphi) \qquad \varphi \in [0,2\pi] \qquad [4]$$

2.4.2 Cardiac pulsatility model (CPM)

Here, we propose an alternative model for capturing pulse-related artifacts in fMRI, named cardiac pulsatility model (CPM). CPM assumes that pulse-related artifacts can be modelled as the output of a causal linear time-invariant (LTI) system where the input is a train of pulses corresponding to cardiac contractions. Therefore, to describe these artifacts we employed the convolution representation:

$$y(t) = \int_{\tau=0}^{T} CPW(t-\tau) \, x(\tau) \, d\tau = x(t) * CPW(t) \qquad [5]$$

where $\chi(t)$ is the input of the model reflecting the cardiac contractions and $CPW(t)$ is the cardiac pulsatility waveform and has the role of an impulse response. The timepoints of the cardiac contractions were considered to coincide with the timings of the PPG peaks. Two different inputs were examined, one input referred to as $x_{CA}(t)$ which indicates that all pulses had a constant amplitude of one, and one input referred to as $x_{VA}(t)$ where the amplitudes of the pulses was equal to the amplitudes of the PPG peaks. Both inputs were defined at a sampling rate of 10 Hz. The pulses consisted of one sample duration each with a constant (unit value) or varying amplitude depending on the kind of input examined (i.e. $x_{CA}(t)$ or $x_{VA}(t)$) and all remaining samples were equal to zero. The impulse response $CPW(t)$ was defined as a linear combination of modified Fourier basis functions of one cycle length as shown below:

$$CPW(t) = \sum_{m=1}^{M} \beta_{cos,m} \, L_{cos,m}(t) \; + \beta_{sin,m} \, L_{sin,m}(t)$$

$$\text{where} \quad L_{cos,m}(t) = 1 - cos\left(\frac{2m\pi t}{T}\right) \quad for \; 0 \le t \le T \qquad [6]$$

$$= 0 \qquad elsewhere$$

$$\text{and} \quad L_{sin,m}(t) = \sin\left(\frac{2m\pi t}{T}\right) \quad for \; 0 \le t \le T$$

$$= 0 \qquad elsewhere$$

The fundamental period T varied on a scan basis and was equal to the average cardiac cycle duration within a scan (i.e. average time difference across pairs of adjacent PPG peaks). The cosine terms were subtracted by one so that all basis functions begin and end at zero. This mathematical manipulation ensures the physiological plausibility of the model as it leads to impulse responses that have zero amplitude at time zero and at times larger that T.

Note that based on the properties of convolution, Eq. 5 can also be expressed as:

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$y(t) = \sum_{m=1}^{M} \beta_{cos,m} L_{cos,m}(t) * x(t) + \beta_{sin,m}\, L_{sin,m}(t) * x(t) \quad [7]$$

As in RETROICOR, nuisance regressors were extracted from the CPM that were subsequently downsampled to the fMRI acquisition timeline and used in the design matrix of GLM.

2.4.3 Time alignment of pulse-related regressors and fMRI data

When the heart contracts a pulse pressure wave propagates from the heart through the vasculature to the whole body giving rise to cardiac pulses in the arteries that are captured, among other hemodynamic signals, in the PPG. The vascular path between heart and finger where PPG is typically recorded may differ in terms of travelling distance compared to the path between heart and arteries in the brain vasculature. As a result, a pulse originated from the heart may reach the finger at a different time compared to an artery in the brain. This time difference should be ideally incorporated in the analysis when extracting pulse-related regressors. To examine whether there is a time difference in HCP fMRI data that can improve the performance of the examined models, we repeated the analysis with each of the models considered in this study for lag times varying from -2 s to 2 s in steps of 0.1 s. In practice, to account for a specific lag time, the PPG signal depending on the lag time examined was shifted towards negative or positive times before extracting the pulse-related regressors. Section 2.6 describes how RETROICOR and the two variants of CPM (i.e. using either input $x_{CA}(t)$ or $x_{VA}(t)$) were compared as well as how the optimal lag time was determined for each of the models.

2.5 Systemic low-frequency physiological oscillations (SLFOs)

Apart from removing high-frequency cardiac and breathing fluctuations, physiological recordings can also be used to yield nuisance regressors that account for the effect of SLFOs (Tong et al., 2019). A common approach for removing SLFOs is to use convolution models where the inputs are the HR and a breathing-related variable (e.g. respiration volume per time; RVT) and the impulse responses are the so-called cardiac (CRF) and respiration (RRF) response functions (Birn et al., 2008; Chang et al., 2009). The outputs of these convolutions are subsequently used as nuisance regressors in the GLM. Falahpour et al. (2013) first demonstrated that the GS of a scan can be used as the model output for estimating subject-specific PRFs. Specifically, the authors showed that nuisance regressors extracted using subject-specific PRFs explain significantly more variance in fMRI data compared to regressors extracted with canonical PRFs. In a recent study, we proposed a new framework for estimating PRFs based on the GS and provided evidence that the PRFs vary significantly even across scans within a subject (Kassinopoulos and Mitsis, 2019b).

Apart from fluctuations to HR and breathing pattern, the GS has been recently shown to be linked, during sleep, to fluctuations in PPG-Amp (Özbay et al., 2019, 2018). However, as PPG-Amp is affected by various physiological processes such as breathing, cardiac activity and arterial blood pressure (Reisner et al., 2008), it is unclear whether fluctuations of PPG-Amp explain unique variance in the GS compared to HR and breathing pattern variations.

In this study, we conducted a cross-correlation analysis to better understand the relation between the physiological sources (HR, breathing patterns and PPG-Amp) and the GS during resting conditions. Fluctuations in breathing patterns were quantified using the respiration volume (RV) measure which is defined as the standard deviation of the breathing signal on a sliding window of 6 s (Chang et al., 2009). The GS was upsampled to 10 Hz to match the sampling rate of physiological variables before proceeding with the cross-correlation analysis.

In addition, we employed the framework proposed in (Kassinopoulos and Mitsis, 2019b) to examine whether a convolution model associated to PPG-Amp variations may explain variance in the GS not attributed to HR and

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

breathing variations. Specifically, we examined whether considering a PRF associated to PPG-Amp variations, (referred to later as PPG-Amp response function; PARF) along with the components related to HR and breathing pattern variations can improve the variance explained on the GS compared to the standard approach of considering only HR and breathing pattern. The nuisance regressors related to HR, PPG-Amp and RV were defined as follows:

$$X_{HR}(t) = HR * CRF, \quad [8]$$

$$X_{PA}(t) = PA * PARF, \text{ and} \quad [9]$$

$$X_{RV}(t) = RV * RRF \quad [10]$$

where PA is the PPG-Amp. We considered the double gamma function (i.e., the sum of two gamma functions) as the structure basis of the PRF curves. The gamma function was defined as:

$$\Gamma(\tau, \delta, t) = a(\tau, \delta) \cdot t^{\frac{\sqrt{\tau}}{\delta}} \cdot e^{-\frac{t}{\delta\sqrt{\tau}}} \quad [11]$$

where the parameters τ and δ indicate the (approximate) time of peak and dispersion of the function, and the parameter α is a scaling factor which normalizes the peak value of the gamma function to 1. The PRF curves were defined as follows:

$$CRF(t) = \beta_{1,c} \cdot \Gamma(\tau_{1,c}, \delta_{1,c}, t) + \beta_{2,c} \cdot \Gamma(\tau_{2,c}, \delta_{2,c}, t),$$
$$PARF(t) = \beta_{1,p} \cdot \Gamma(\tau_{1,p}, \delta_{1,p}, t) + \beta_{2,p} \cdot \Gamma(\tau_{2,p}, \delta_{2,p}, t), \text{ and} \quad [12]$$
$$RRF(t) = \beta_{1,r} \cdot \Gamma(\tau_{1,r}, \delta_{1,r}, t) + \beta_{2,r} \cdot \Gamma(\tau_{2,r}, \delta_{2,r}, t)$$

The procedure for estimating scan-specific PRFs is described in detail in (Kassinopoulos and Mitsis, 2019b). In brief, the parameters of the PRFs for a given scan were first estimated using a genetic algorithm (GA) implemented in Matlab R2018b's Global Optimization Toolbox. The parameter vectors $\boldsymbol{\tau}$ $(\tau_{1,c}, \tau_{2,c}, \tau_{1,p}, \tau_{2,p}, \tau_{1,r}, \tau_{2,r})$ and $\boldsymbol{\delta}$ $(\delta_{1,c}, \delta_{2,c}, \delta_{1,p}, \delta_{2,p}, \delta_{1,r}, \delta_{2,r})$ were bounded between 0-20 seconds and 0-3 seconds, respectively. A stopping criterion of 100 generations was set, as it was found to be adequate for convergence. GA searches in the parameter space defined by the boundaries to estimate the parameters that maximize the objective function (i.e. a Pearson correlation coefficient) yielded through the following steps: 1. for a set of given parameters, the three PRF curves are constructed. Subsequently, 2. the HR, PPG-Amp and RV signals (sampled at 10 Hz) are convolved with CRF, $PARF$ and RRF, respectively, to extract the corresponding nuisance regressors and then downsampled to match the fMRI acquisition rate. 3. The beta parameters are estimated through linear regression (GLM) whereby the GS is the dependent variable and the three nuisance regressors are the three explanatory variables. 4. Finally, the Pearson correlation coefficient between the GS and the model prediction is calculated. The estimated parameters of the GA were subsequently refined using the interior-point gradient-based algorithm with a stopping criterion of 100 maximum iterations (implemented in Matlab R2018 as well).

2.6 Model comparison

In the case of high-frequency physiological fluctuations, the goodness-of-fit for a given model was assessed on each individual voxel using the Pearson correlation between the output of the model and the voxel timeseries. Then, the correlation was averaged across all voxels in the brain. In contrast, when examining low-frequency physiological fluctuations, the goodness-of-fit for a given model was assessed on the GS using the Pearson correlation between the output of the model and the GS.

To compare the performance across models a 3-fold cross-validation framework was employed as the models differed in flexibility and, inevitably, risk of overfitting due to a different number of parameters. When modelling low-frequency physiological fluctuations, the GS of each scan was partitioned into three segments of about 5 min each. One segment was used as the validation set for assessing the performance of the model and the remaining two

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

segments were used as the training dataset. This step was repeated three times with each of the three segments used exactly once as the validation data. In each fold, scan-specific PRFs and beta parameters were estimated from the training dataset and, subsequently, used in the validation dataset to model the SLFOs. The goodness-of-fit was assessed based on the correlation between the GS and estimated SLFOs in the validation dataset. Finally, the mean correlation across the three folds was calculated. To compare the standard with the proposed model, the mean correlations of the four scans of each subject were first averaged and, then, a paired t-test was performed based on the mean correlation values of the 100 subjects for the two models. The 3-fold cross-validation and model comparison was also performed in a similar manner for the high-frequency physiological fluctuations.

With regards to high-frequency cardiac pulsatility oscillations, three models were examined, namely RETROICOR, the CPM_{CA} and the CPM_{VA} (see Section 2.4). The two last models only differed on whether the input signal was $x_{CA}(t)$ or $x_{VA}(t)$. Apart from the model type, we also examined the optimal model order of Fourier series and the optimal lag time as these two parameters may have significant impact on the performance. Therefore, the model performance was initially assessed for each of the three model types for model order varying from 1 to 8 and lag time varying from -2 s to 2 s in steps of 0.1 s. The comparison of the three models was subsequently done using for each model the order and lag time that yielded the best performance across all subjects. To illustrate how well the three models capture cardiac pulsatility and how their waveforms may differ between them, we also repeated the comparison considering the PPG raw signal as the output target rather than the voxel timeseries.

To assess regional variability in the performance of the pulse-related models statistical maps were generated. For visualization purposes, maps shown here were overlaid on structural images after being transformed to MNI152 space (1 mm spatial resolution) with FSL's FLIRT registration tool (Jenkinson and Smith, 2001) as incorporated in the MANGO software (Lancaster, Martinez; www.ric.uthscsa.edu/mango).

With respect to the low-frequency noise (Section 2.5), we compared two models for extracting the SLFOs using the GS. The standard model estimates SLFOs presented in the GS using only the variations in HR and breathing patterns whereas the proposed model estimates SLFOs using HR, breathing patterns and PPG-Amp. The two models can be expressed as follows:

Standard model: $$GS = SLFOs + \varepsilon = HR * CRF + RV * RRF + \varepsilon \quad [13]$$

Proposed model: $$GS = SLFOs + \varepsilon = HR * CRF + RV * RRF + PA * PARF + \varepsilon \quad [14]$$

2.7 Cardiac pulsatility waveforms (CPW)

Various MRI techniques have been proposed to measure intracranial CPWs such as phase-contrast and magnetic resonance encephalography as these waveforms are considered a useful biomarker in certain cerebrovascular diseases (Bianciardi et al., 2016; Wagshul et al., 2011) and are often studied in order to understand the role of intracranial cardiac pulsatility in the glymphatic activity (Fultz et al., 2019; Kiviniemi et al., 2016). With this in mind, here we sought to investigate whether the CPM applied in BOLD fMRI data may provide an alternative technique for measuring intracranial pulsatility. Specifically, we sought to investigate the consistency of CPWs across subjects.

To extract CPWs averaged across subjects we followed the following steps for the two variants of CPM separately: For a given scan, we derived the CPW in each voxel based on Eq. 6 using the beta parameters estimated in the GLM. Note that when reconstructing the CPW, a cardiac cycle duration T of 1 s was used for all subjects and a time-step of 0.025 s. Subsequently, the CPW in each voxel was normalized to a maximum absolute value of one and multiplied by the correlation coefficient related to the variance explained with CPM in the associated voxel. Finally, the CPWs were averaged across subjects considering only scans from the first session and the same PE direction. This way, CPWs in regions that were prone to pulse-related fluctuations had larger amplitudes than other regions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3. Results

3.1 High-frequency cardiac fluctuations

All three models examined for pulse-related oscillations were able to model the cardiac pulsatility in the PPG quite well. When considering a 3-fold cross-validation framework, the three models showed improved performance for higher orders of Fourier series reaching a plateau at about the 4th order. Considering the highest examined (8th) order which demonstrated the best performance, we observed a maximum cross-correlation averaged across subjects at time lags between -0.1 s and -0.5 s (Suppl. Fig. 1a). RETROICOR, CPM_{CA} and CPM_{VA} yielded maximum cross-correlation values of 0.900, 0.901 and 0.934, respectively. CPM_{VA} which, in contrast to RETROICOR and CPM_{CA}, takes into account variations in PPG-Amp, demonstrated significantly higher performance compared to the other two models ($p<10^{-31}$), while RETROICOR and CPM_{CA} yielded similar performance. Fig. 1 shows the goodness-of-fit for the three models when applied on the PPG of a subject that illustrated strong variations in HR (the 3-fold cross validation framework was omitted for this figure). As we can see, all three models explained fairly well the variations

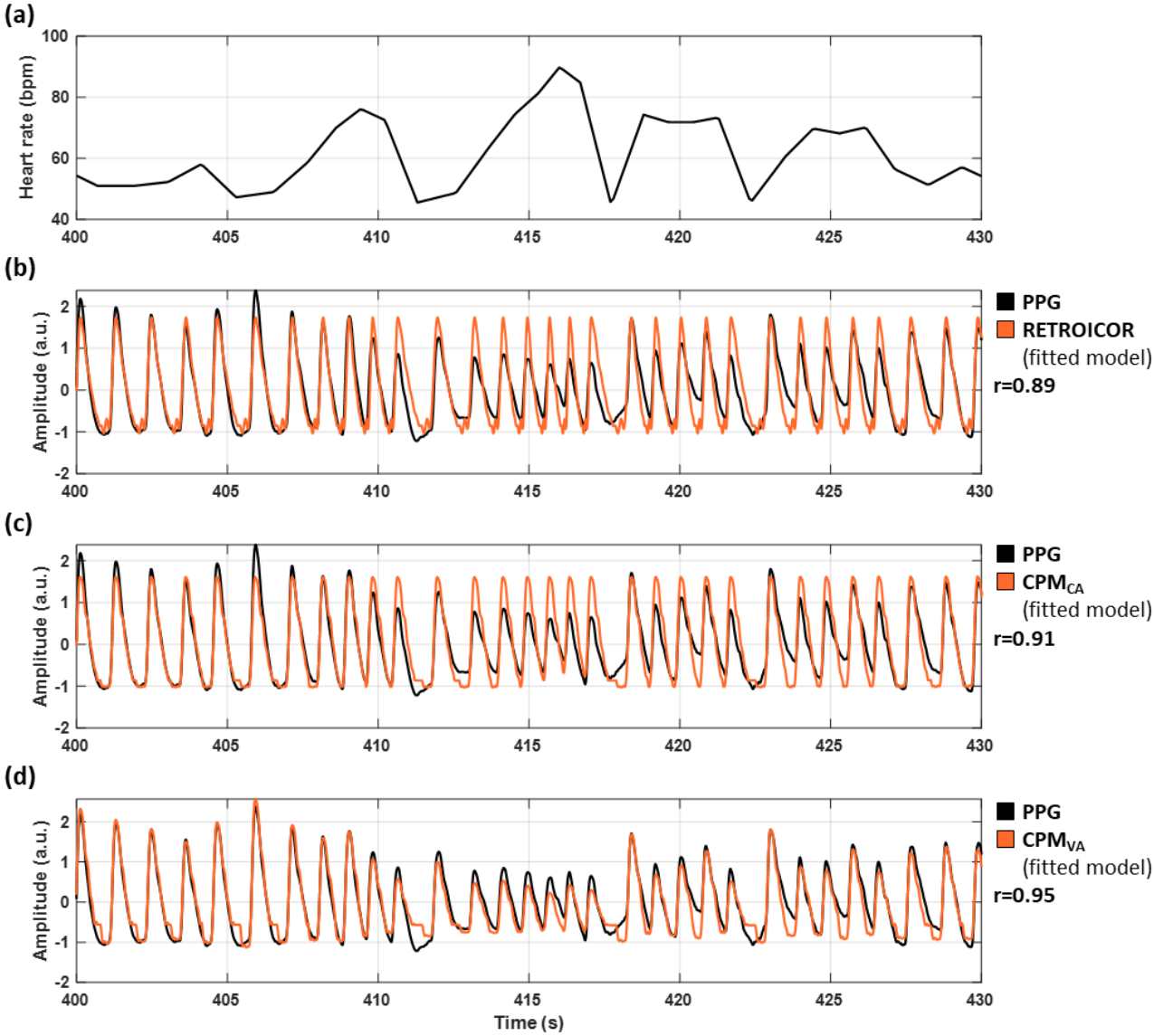


Fig. 1. Model fit of pulse-related models on PPG for a subject with high heart rate variability (HRV; S159138-R1LR). (a) Trace of HR during a 30 s time segment with strong fluctuations. (b)-(d) Model fit of RETROICOR, CPM_{CA} and CPM_{VA} on the PPG timeseries. All three models captured the high-frequency (~1 Hz) fluctuations related to cardiac pulsatility. However, only CPM_{VA} that incorporates in its input the low-frequency (~0.1 Hz) fluctuations in PPG-Amp was able to represent these fluctuations on the model output.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in the PPG signal, although only CPM_{VA} was able to capture the low-frequency fluctuations of the PPG-Amp. Apart from the absent of low-frequency fluctuations in the case of RETROICOR and CPM_{CA}, we were not able to visually observe any other differences in the waveforms across the three models. That said, as shown later, when examining fMRI data, RETROICOR and CPM_{CA} exhibited better performance compared to RETROICOR. Due to this, we

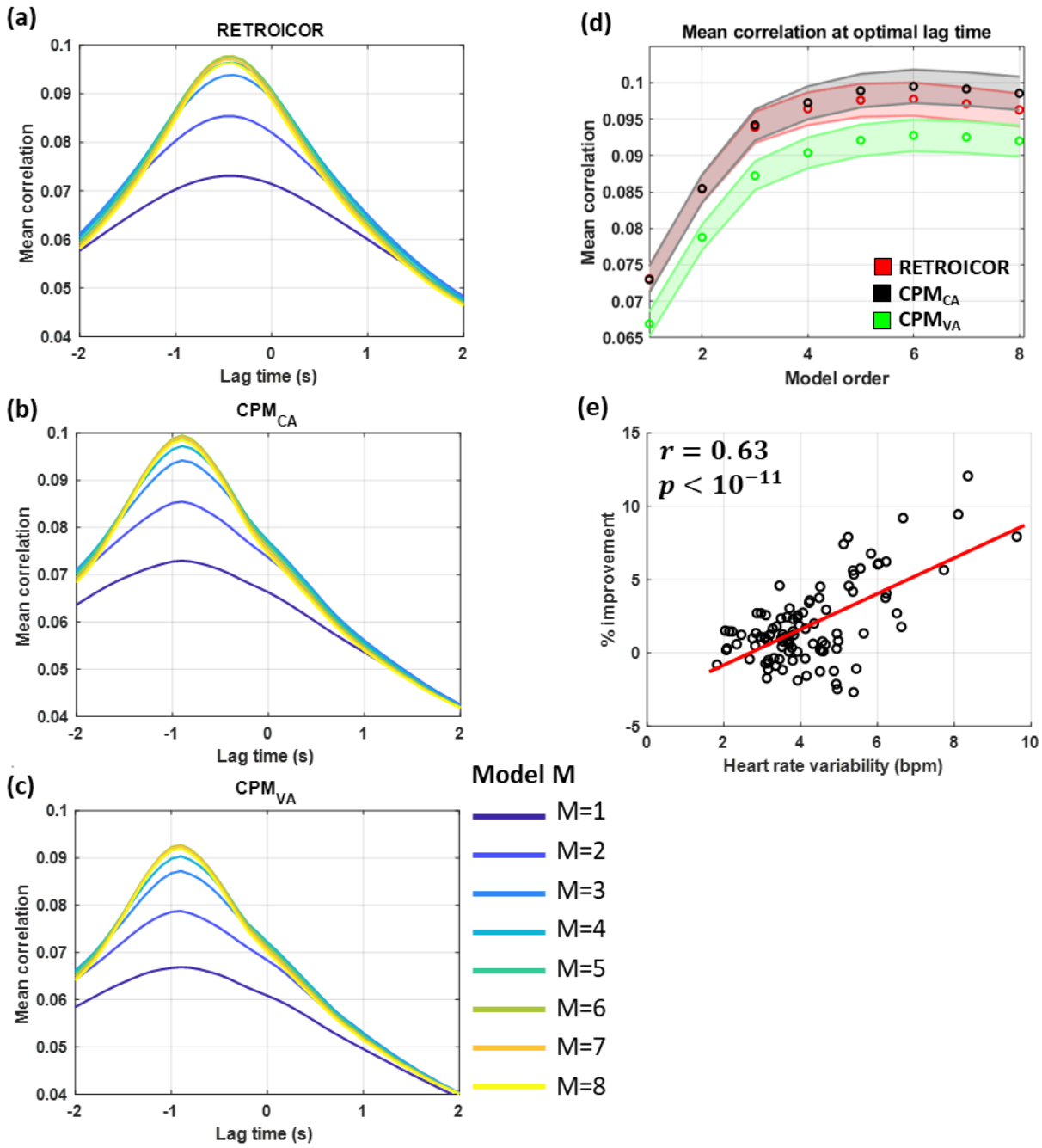


Fig. 2. Performance of pulse-related models on fMRI data. (a)-(c) Cross-correlation averaged across subjects for RETROICOR, CPM_{CA} and CPM_{VA}, respectively. Colors vary from blue to yellow indicating performance for higher model orders of Fourier series. For all three models a unimodal curve was observed with highest mean correlations for lag times between -0.4 and -0.9 s. The optimal lag time in each model was consistent across model orders. (d) Mean correlation with respect to model order for optimal lag time of -0.4 s for RETROICOR and -0.9 s for the two variants of CPM. All models yielded a maximum mean correlation for a 6th model order. In contrast to the performance on the PPG, when examining fMRI data, incorporating the low-frequency fluctuations of PPG-Amp in the input signal was found to harm the model fit (i.e. CPM_{CA} vs CPM_{VA}). (e) Relative percentage (%) improvement with respect to HRV when comparing the proposed model CPM_{CA} with RETROICOR. The stronger were the fluctuations in HR the larger was the improvement achieved with CPM_{CA} compared to RETROICOR ($p<10^{-11}$).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Comparison of pulse-related models in terms of variance explained in fMRI data

A/A	Model type:	Order	Lag time:	Mean correlation (SD):	*P*-value of improvement:
M1	RETROICOR	2nd	0 s	0.082 (0.019)	-
M2	RETROICOR	2nd	-0.4 s	0.085 (0.019)	$P(M2>M1) < 10^{-24}$
M3	RETROICOR	6th	-0.4 s	0.097 (0.023)	$P(M3>M2) < 10^{-29}$
M4	CPM_{CA}	6th	-0.9 s	0.100 (0.023)	$P(M4>M3) < 10^{-8}$

further tested in the analysis with the PPG as output whether scans with high HRV showed relative improvement with CPM_{CA} compared to RETROICOR (for this test 8th order was used with the optimal lag time for each model). HRV was defined as the standard deviation of HR. Indeed, when examining the goodness-of-fit on the PPG, scans with high HRV were characterized by higher relative improvement (Suppl. Fig. 1b; r=0.32, p<0.001).

Fig. 2a-c shows the cross-correlation averaged across subjects (and across all voxels and scans in each subject) for the three models when output target were the fMRI timeseries. As many voxels are not prone to pulse-related fluctuations the correlation values averaged across all voxels were relatively low. In all model orders examined, RETROICOR exhibited a peak in cross-correlation at lag time of -0.4 s whereas the two CPM models exhibited a peak at lag time of -0.9 s (a negative lag time indicates that the PPG signal was shifted backward in time). As such,

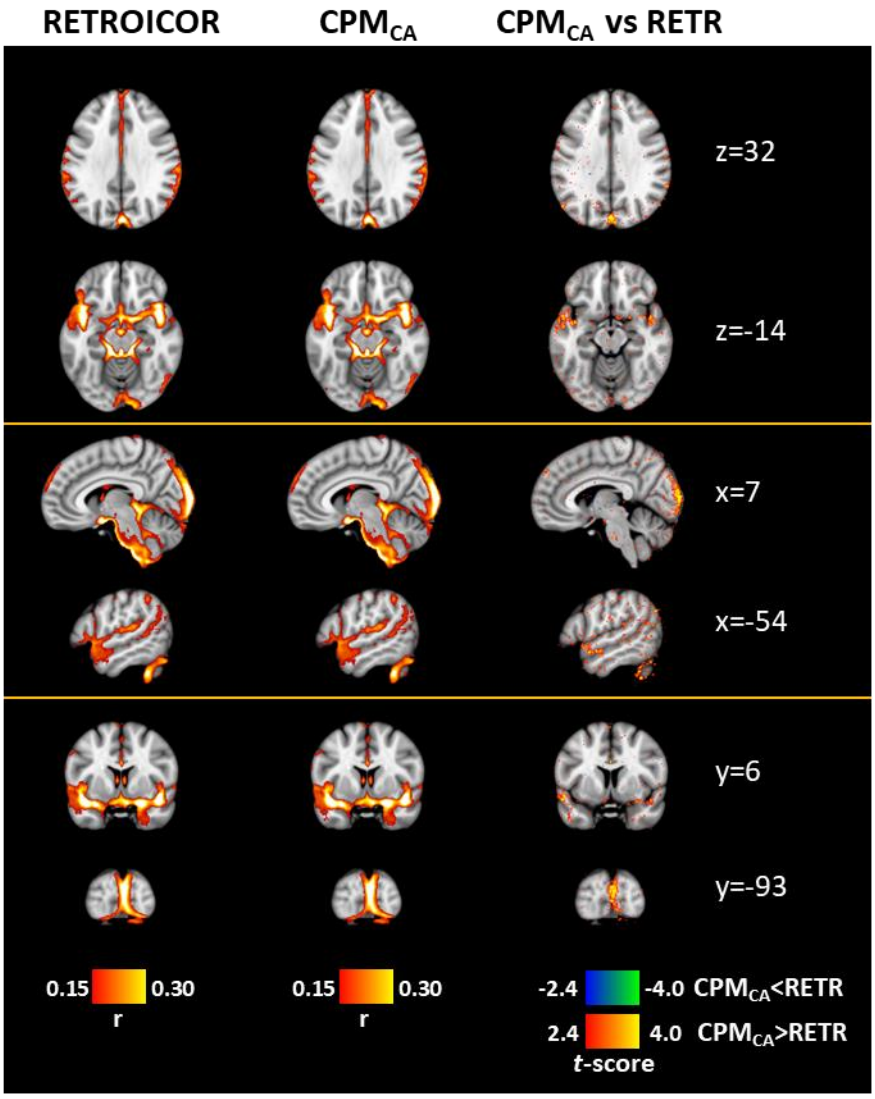


Fig. 3. Correlation maps for pulse-related models averaged across all subjects (N=100; only scans with left-right PE directions from the first sessions were included here). The first and second columns show the variance explained using RETROICOR and the proposed model CPM_{CA}, respectively, while the third column shows t-stat maps indicating the areas with significant differences in correlation values between the two models (p<0.01). The correlation threshold value was chosen arbitrarily so that only the regions more susceptible to cardiac pulsatility are presented. Both RETROICOR and CPM_{CA} found regions close to the basilar and vertebral arteries, in the 4th ventricle, in the superior sagittal sinus, in the lateral sulcus and in the occipital lobe to be more prone to pulse-related oscillations. However, based on the t-stat map, CPM_{CA} explained significantly larger fraction of variance than RETROICOR in the occipital cortex, the lateral sulcus and superior sagittal sinus. The statistical maps shown in this figure are available on: https://neurovault.org/collections/DHFETQTN/.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to ease the comparison between models, Fig. 2d shows the mean correlation for all models along with the associated standard error when considering the optimal lag time of each model type. We observe that, in contrast to the analysis that had the PPG as target output, the CPM_{VA} yielded poorer performance compared to RETROICOR and CPM_{CA}. Moreover, for all three models the highest performance was achieved with a 6th order Fourier series whereas higher orders yielded slightly lower mean correlations.

Table 1 summarizes the performance of four specific models in order to illustrate the main steps that can lead to an improvement in the variance explained in the fMRI data. Model M1 corresponds to the 2nd order RETROICOR that is commonly employed in fMRI studies without considering any lag time whereas model M2 differs in that it considers the optimal lag time of RETROICOR (i.e. -0.4 s). Based on a paired t-test, considering a lag time when extracting the nuisance regressors led to a significant improvement in the variance explained ($p<10^{-24}$). Moreover, employing a 6th instead of a 2nd order Fourier series resulted also in an additional improvement ($p<10^{-29}$). Finally, when optimal model order and lag time were considered, further improvement was achieved when modelling cardiac pulsatility using the CPM_{CA} model proposed here rather than RETROICOR ($p<10^{-8}$). As hypothesized, when comparing RETROICOR (6th order and -0.4 s lag time) and CPM_{CA} (6th order and -0.9 s) the higher was the HRV the larger was the relative improvement achieved (Fig. 2e; r=0.63; $p<10^{-11}$).

Fig. 3 shows correlation maps averaged across subjects (only scans with left-right PE direction from the first session were included) as obtained with RETROICOR and CPM_{CA} when considering 6th order Fourier series and the optimal lag time of each model (i.e. -0.4 s and -0.9 s for RETROICOR and CPM_{CA}, respectively). It also shows t-score maps derived with paired t-test indicating brain regions with significant differences in goodness-of-fit between the two models. As expected, both RETROICOR and CPM_{CA} explained significant variance in areas with CSF (e.g. areas around the brainstem, in the 4th ventricle and superior sagittal sinus) as well as in lateral sulcus and occipital cortex. CPM_{CA} demonstrated better performance in terms of variance explained, particularly in the occipital cortex, lateral sulcus and superior sagittal sinus, while none of the regions illustrated better fit with RETROICOR than with CPM_{CA}.

In an earlier studied we sought to investigate the role of physiological processes in fMRI connectome-based subject discriminability (Xifra-porxas et al., 2020). One of our findings was that the connectome signature driven by cardiac pulsatility differs between left-right and right-left PE direction (see for example Fig. 5 in Xifra-porxas et al. (2020)). To shed light on this phenomenon, here we compared the variance explained with CPM_{CA} for scans with left-right vs right-left PE direction including only scans from the first session of each subject. Interestingly, in the third column of Fig. 4 that corresponds to t-scores of correlations for left-right vs right-left PE direction we observe antisymmetric patterns with respect to the anterior-posterior axis (see specifically the first three rows corresponding to axial slices). In other words, if for example a region in the left hemisphere was more prone to pulse-related oscillations for left-right PE direction compared to right-left PE direction, then the contralateral region in the right hemisphere was more prone to these oscillations for right-left PE direction. Moreover, we observe that the regions that show this PE direction dependence are not necessarily the regions more susceptible to pulse-related oscillations.

Moreover, we investigated the regional variability of the CPWs. As described in Methods (Section 2.4.2), the beta parameters associated to the CPM nuisance regressors define the temporal waveforms of the impulse responses used in the CPM, referred to here as CPWs (cardiac pulsatility waveforms). Based on the beta parameters estimated with the GLM, we inspected the temporal evolution of the CPWs both at the individual (subject) and group (averaged across subjects considering only scans with the same PE direction from the first session) level. For this analysis, a 6th order CPM_{CA} with a lag time of -0.9 s was considered as the cross-validation analysis presented earlier showed that this choice of parameters yields the best performance. At the individual level we were not able to observe any clear pattern apart from discontinuities in adjacent slices that were likely due to time acquisition differences across slices. Note that we did not incorporate any slice-timing correction in our analysis because this is not a trivial task for the multi-band fMRI data examined here. However, at the group level (N=100), we were able to see smooth spatiotemporal patterns without discontinuities in adjacent slices. Videos showing the temporal evolution of CPWs in a mid-sagittal plane for left-right and right-left PE direction are available on repository https://doi.org/10.6084/m9.figshare.c.4946799 (Kassinopoulos and Mitsis, 2020). In addition, Suppl. Fig. 2 shows

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

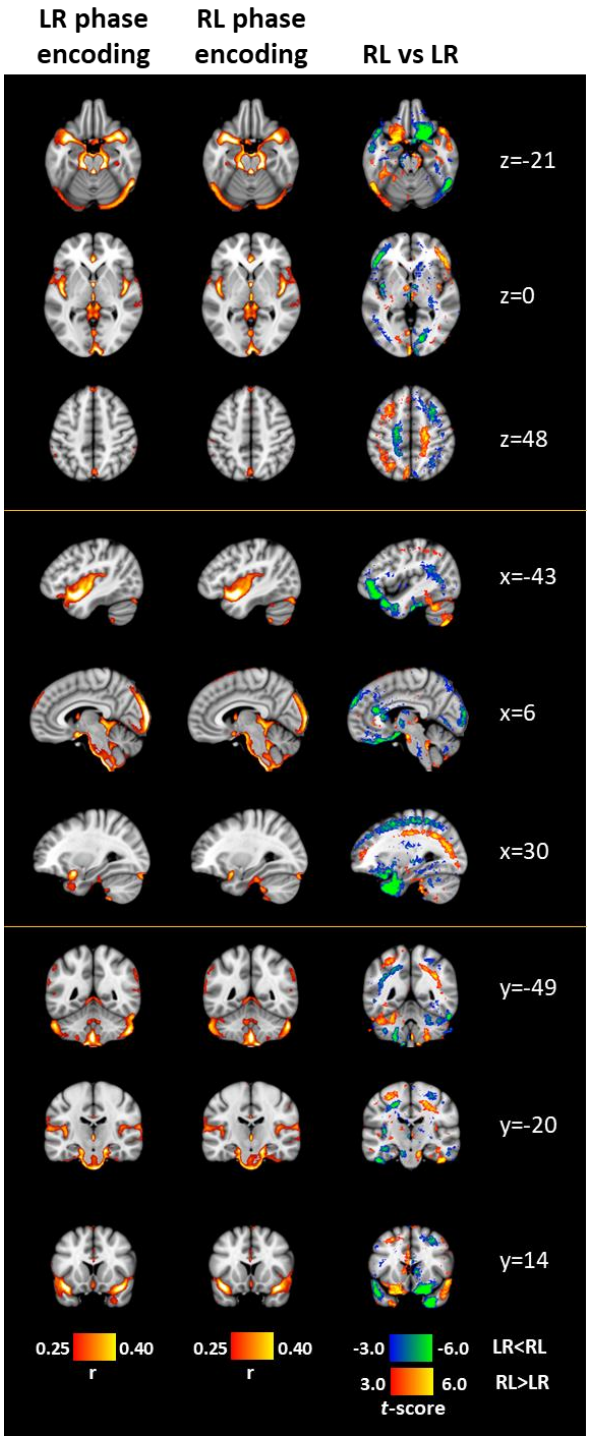


Fig. 4. The effect of phase encoding (PE) direction on pulse-related fluctuations. The first and second columns show the variance explained using CPM_{CA} averaged across all subjects for scans with left-right and right-left PE direction, respectively (N=100; only scans from the first session were included here). The third column shows t-stat maps indicating the areas with significant differences between the two PE directions (p<0.01). The correlation threshold value was chosen arbitrarily so that only the regions more prone to cardiac pulsatility are presented (note that the cross-validation framework was omitted here as we do not compare models rather than examining the effect of PE direction, hence the higher correlation values compared to Fig. 3). As it can be seen from the axial slices, the regions with significant differences between the two PE directions were characterized by antisymmetric patterns with respect to the anterior-posterior axis. For instance, if a brain area in the left hemisphere was more prominent to fluctuations due to cardiac pulsatility for a left-right PE direction compared to a right-left PE direction, then the homologues area in the opposite hemisphere was more prominent to these physiological fluctuations for a right-left PE direction. Furthermore, the areas that showed the aforementioned PE direction dependence for pulse-related fluctuations were not necessarily the areas that were strongly affected by these pulse-related fluctuations. The statistical maps shown in this figure are available on: https://neurovault.org/collections/DHFETQTN/.

the CPWs at five timepoints of a cardiac cycle for the left-right PE direction. Similar dynamics were observed between the two PE directions. While some regions demonstrated an increase in the BOLD fMRI signal after the onset of a cardiac contraction (e.g. regions in the posterior cingulate cortex) other regions showed the opposite trend (e.g. third ventricle and regions in the anterior cingulate cortex). Interestingly, the temporal dynamics shown in the video revealed patterns that may relate to some fluid movements. Particularly, we observed along the cerebral aqueduct that connects the 3rd with the 4th ventricle two spatial layers with opposite temporal responses that peaked at about the middle of the cardiac cycle.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.2 Systemic low-frequency physiological oscillations (SLFOs)

Fig. 5 presents the results of the cross-correlation analysis that was used to investigate the relation between physiological variables (RV, HR and PPG-Amp) and the fMRI GS during resting conditions. Note that cross-correlation is often used to estimate the impulse response for a given input and output. However, this strategy is only suitable for input signals with zero-mean white noise signal. Since in this work we are dealing with physiological variables that are sluggish, we use cross-correlation to get a rough idea of whether timeseries share covariance between them, either with the same or opposite polarity, and whether there is a significant lag time between them.

As can be seen in the diagonal of the diagram, all timeseries exhibited a somewhat monotonic decrease in autocorrelation for increasing (absolute) lag times with the autocorrelation approximating zero for lag times between 10 and 20 s. This trend is expected considering the sluggishness of the examined signals. Looking at the non-diagonal plots we observe that, despite the low correlation values, all three physiological variables demonstrated interactions between them that were consistent across subjects. Increase in breathing activity as quantified with RV was accompanied by concurrent increase in HR and decrease in PPG-Amp. However, PPG-Amp, apart from the decrease, exhibited also a positive peak

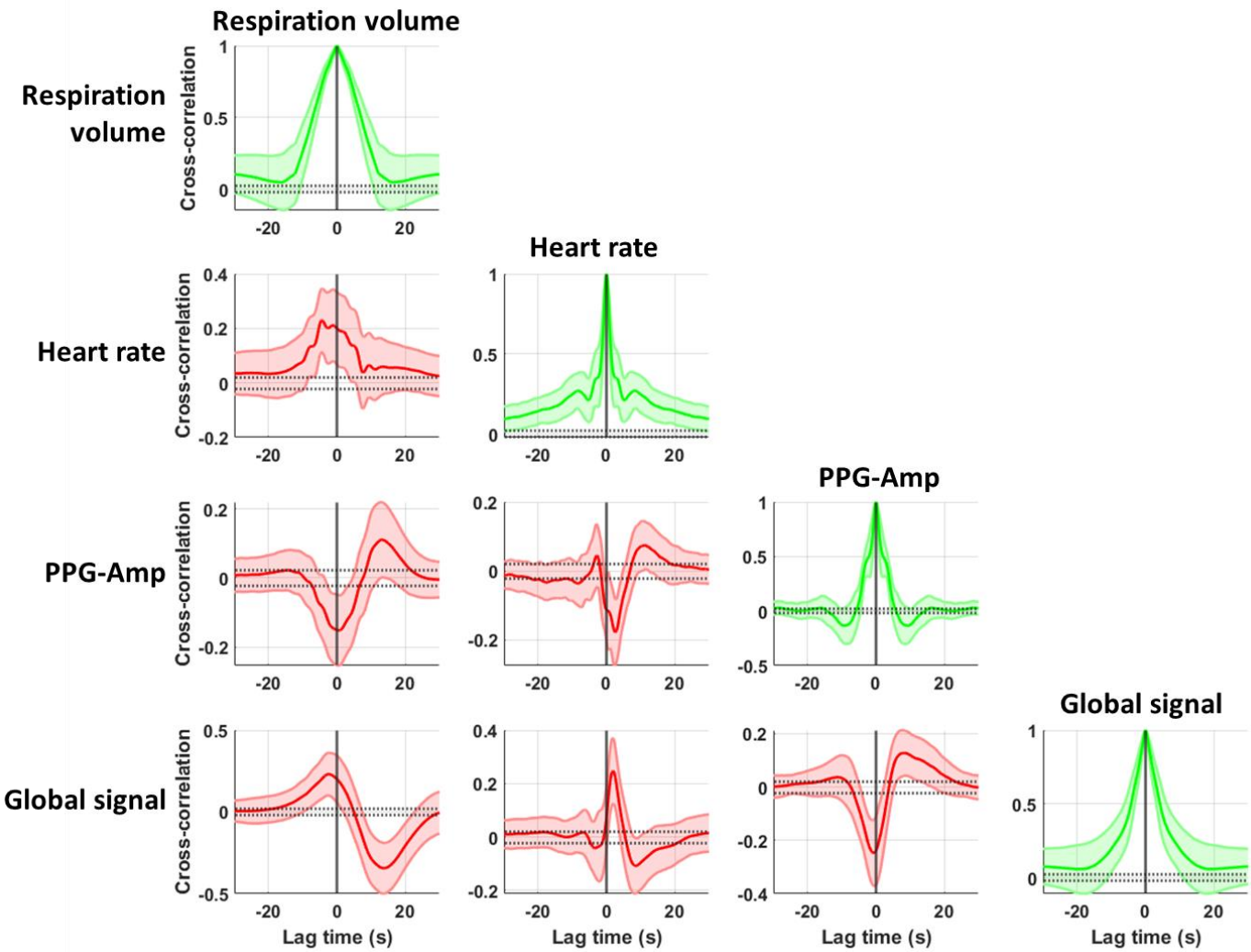


Fig. 5. Cross-correlation between physiological variables and fMRI global signal (GS) averaged across subjects (n=100). The value of cross-correlation between row x and column y at lag time τ indicates the correlation of the corresponding signals when signal of row x is shifted forward in time by τ. A way we can interpret the observed curves is as follows: if we consider for example the case of HR and GS, the fact that the maximum absolute correlation is at +1.9 s lag time and has a positive correlation indicates that HR maximizes its covariance with the GS when is shifted forward in time by 1.9 s (a negative correlation would indicate that HR should be inverted to maximize the covariance). Overall, we observe that all physiological signals share covariance with each other as well as with the GS. Note that diagonal plots correspond to autocorrelations (i.e. cross-correlation of a signal with itself). Shaded areas indicate the standard deviation across subjects.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

about 13 s after the peak in RV. PPG-Amp showed a similar bimodal trend with respect to HR increases event though with different dynamics. Specifically, PPG-Amp demonstrated a negative peak at about 2.5 s after the increase in HR followed by a positive peak, albeit with a smaller amplitude, at a lag time of 10 s.

In addition, GS was found to share covariance with all physiological variables. Perhaps not surprisingly, the cross-correlation of GS with RV and HR somewhat resemble, respectively, the CRF and RRF reported in an earlier study (Kassinopoulos and Mitsis, 2019b). As the main trend of the corresponding cross-correlations suggests, an increase in RV and HR leads to an increase in the GS followed by a negative undershoot. An opposite trend was observed in the cross-correlation between PPG-Amp and GS. Specifically, the GS exhibited a strong negative peak at -1 s lag time as well as a positive peak at 8 s lag time with respect to the PPG-Amp increase.

The strong association observed between PPG-Amp and GS in the cross-correlation analysis (Fig. 5) raised the question whether considering PPG-Amp variations in addition to RV and HR variations may provide additional information when modelling SLFOs in the GS. To address this question, we compared the standard approach for modelling SLFOs (i.e. considering only HR and RV) with the extended model that accounts also PPG-Amp variations, using a 3-fold cross-validation framework (for more information see Section 2.5). The cross-validation framework was necessary for this comparison due to the larger number of parameters in the extended model. Our results revealed a significant improvement in terms of variance explained in the GS when considering PPG-Amp variations. Specifically, the correlation between GS and predicted SLFOs exhibited a statistically significant increase from 0.65 (±0.10) to 0.66 (±0.10) when taking into account PPG-Amp variations ($p < 10^{-4}$).

Suppl. Fig. 3a shows the estimated PRFs averaged across all subjects, namely the cardiac (CRF), PPG-Amp (PARF) and respiration (RRF) response functions. Cross-validation was omitted when estimating the PRFs presented in Suppl. Fig. 3a. To obtain the main trend of PRFs observed across subjects, the estimated PRFs were averaged across scans using

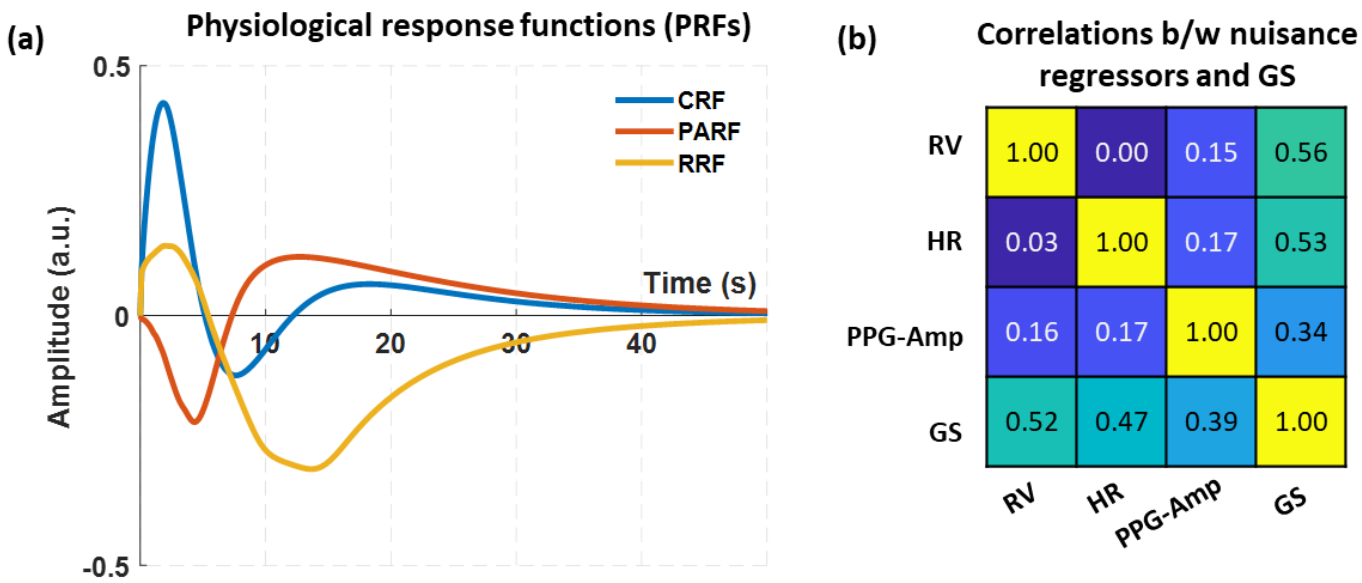


Fig. 6. Estimated physiological response functions (PRFs) when considering the 5 s shifted PPG-Amp (PA_5). (a) PRFs averaged across all subjects and scans using weighted average with the correlation between GS and the predicted output of the model (i.e. SLFOs) for each scan as a weighted coefficient. (b) Correlation between nuisance regressors (i.e. X_{RV}, X_{HR} and X_{PA}) and GS, averaged across all scans. The lower-diagonal elements correspond to correlations whereas the upper-diagonal elements correspond to partial correlations. The partial correlations between pairs of the three nuisance regressors did not control for GS variations as GS is not considered to affect the three associated physiological variables. Note that as in this analysis we do not compare models, the cross-validation framework was omitted. CRF and RRF exhibited a positive peak at around 2 s followed by a negative peak at 8 s for CRF and 13-14 s for RRF. PARF was characterized by a negative peak at 4.3 s followed by a positive peak at 12.5 s. All PRFs demonstrated a slow decay that approximated zero at around 40 s. While the nuisance regressors demonstrated relatively low correlations between them ranging from 0.03 to 0.17, all of them exhibited high correlation with the GS (≥0.39). Similar observations were made for the partial correlations.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

weighted average where the correlation between GS and the associated nuisance regressors (X_{HR}, X_{PA} and X_{RV}) was used as the weighted coefficient. Both CRF and RRF exhibited smooth bimodal curves with a positive peak followed by a negative peak and agreed in terms of dynamics with the PRFs reported in Kassinopoulos and Mitsis (2019a). In contrast, PARF demonstrated a sharp negative peak at 0.2 s followed by a slow positive overshoot.

The sharp negative peak in the PARF (Suppl. Fig. 3a) combined with the minimum peak observed at negative lag time in the cross-correlation between GS and PPG-Amp (Fig. 5) suggested that fluctuations in GS may precede fluctuations in PPG-Amp. As this may be indeed the case, to examine whether incorporating a time difference in the model could lead to further improvement, we also examined the extended model for SLFOs with the timeseries of PPG-Amp shifted back in time by 5 and 10 s. Later we refer to the original and the 5 s and 10 s shifted variants of PPG-Amp as PA_0, PA_5 and PA_{10}, respectively. Using the cross-validation framework, PA_5 was found to yield the highest mean correlation of 0.67 (±0.09) which was statistically higher than the mean correlation (0.66) achieved with the original timeseries (PA_0; $p < 10^{-4}$). Fig. 6a shows the PRFs averaged across subjects when using the PPG-Amp shifted back in time by 5 s (i.e. PA_5). As we can see, PARF exhibited again a bimodal curve with a negative peak at 4.3 s followed by a positive peak at 12.5 s. Due to the time shift of PPG-Amp, the negative peak of the PARF was smoother compared to the negative peak observed with the original PPG-Amp timeseries (Suppl. Fig. 3a). We also notice that the time shift of PPG-Amp did not have any effect on CRF and RRF. Similar to our previous study (Kassinopoulos and Mitsis, 2019b), we observed variability in the estimated curves across scans for all types of PRFs, with the CRF being the most consistent (the estimated PRFs from all scans examined here can be found on https://doi.org/10.6084/m9.figshare.c.4946799 (Kassinopoulos and Mitsis, 2020).

Fig. 6b shows the correlation averaged across all scans between the GS and the nuisance regressors extracted with the PRFs (X_{RV}, X_{HR} and X_{PA}). As we can see, the nuisance regressors did not show high mean correlations between them. However, looking at individual scans, we often observe pairs of nuisance regressors (X_{RV} vs X_{HR}, X_{RV} vs X_{PA} and X_{HR} vs X_{PA}) to be highly correlated. Moreover, in Fig. 6b we see that all three nuisance regressors were strongly associated with the GS, with X_{RV} exhibiting the highest mean correlation (0.52) and X_{PA} the weakest one (0.39). As the original physiological variables exhibited strong interactions between them, we also considered partial correlations to quantify the fraction of variance explained on the GS from each physiological variable when controlling for the variance explained from the other two. Similar observations were made for partial correlation (Fig. 6b). Fig. 7 shows the performance of the extended model for a scan where the nuisance regressors extracted from HR, PPG-Amp and RV were very similar and also explained fairly well the low-frequency oscillations in the GS.

While the contribution of PPG-Amp variations on the GS was on average lower compared to the contribution of HR and RV variations, several subjects exhibited stronger relation of GS with PPG-Amp than with HR or RV. To shed light on this subject variability, we examined whether the partial correlation between PPG-Amp and GS depends on the body type and blood properties of the participants. Specifically, among the measures collected from participants in HCP, we considered the body weight, body mass index, height, systolic and diastolic pressure as well as hematocrit. In addition, we considered the HRV (i.e. standard deviation of HR) and standard deviation of RV as estimated from the physiological recordings. We chose these measures due to their strong relation to hemodynamic properties such as blood viscosity and total peripheral resistance in the blood circulation. Among the eight tests performed, at a significance level of 0.05, only HRV and hematocrit were strongly associated to the partial correlation between GS and PPG-Amp (Fig. 8). Specifically, both HRV and hematocrit exhibited a strong positive relation to the partial correlation between GS and PPG-Amp (p<0.001). Based on this observation, we also examined the relation between hematocrit and HRV. However, we did not find any relation between them (p=0.67).

Finally, consistent to previous studies, variations in HR and breathing pattern had a strong effect across widespread regions in the gray matter (Birn et al., 2006; Kassinopoulos and Mitsis, 2019b; Shmueli et al., 2007). The same regions were found to be also strongly associated with variations in PPG-Amp (Suppl. Fig. 4).

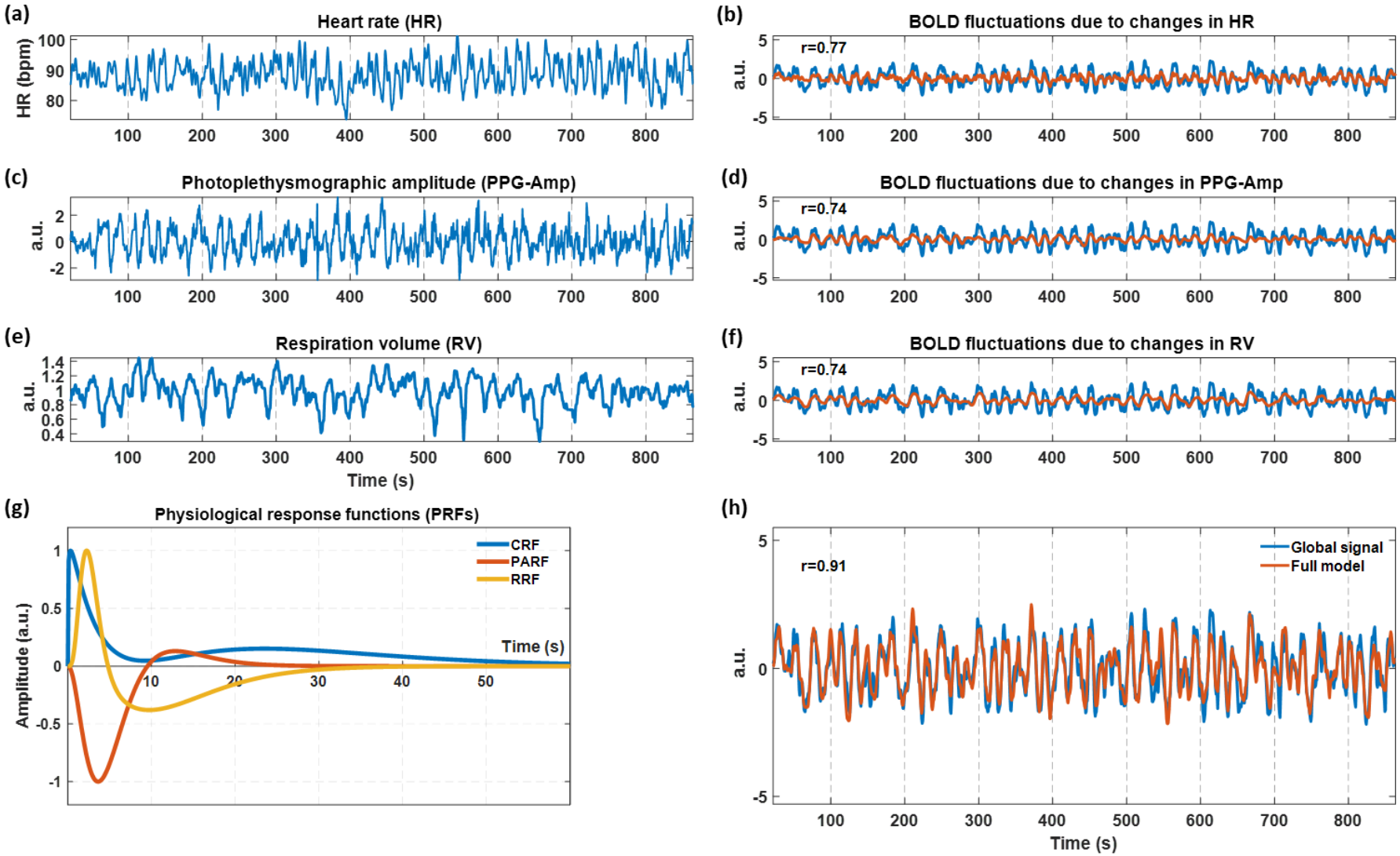


Fig. 7. Demonstration of the estimated SLFOs in the GS for scan S103818-R2RL. (a), (c) and (e) show, respectively, the traces of HR, PPG-Amp and RV during the scan, whereas (b), (d) and (f) show the fit of the nuisance regressors extracted from the physiological variables (orange color) on the GS (blue color). (g) Scan-specific PRFs estimated using the framework proposed in Kassinopoulos and Mitsis (2019a). (h) Model fit of predicted output (i.e. SLFOs) on the fMRI GS. For the purposes of visualization, the HR in (a) was smoothed using a moving average filter of 3 s. For this particular scan, all three nuisance regressors obtained from HR, PPG-Amp and RV explained a large fraction of variance in the GS. The corresponding figures for the remaining of the scans can be found on https://doi.org/10.6084/m9.figshare.c.4946799 (Kassinopoulos and Mitsis, 2020).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

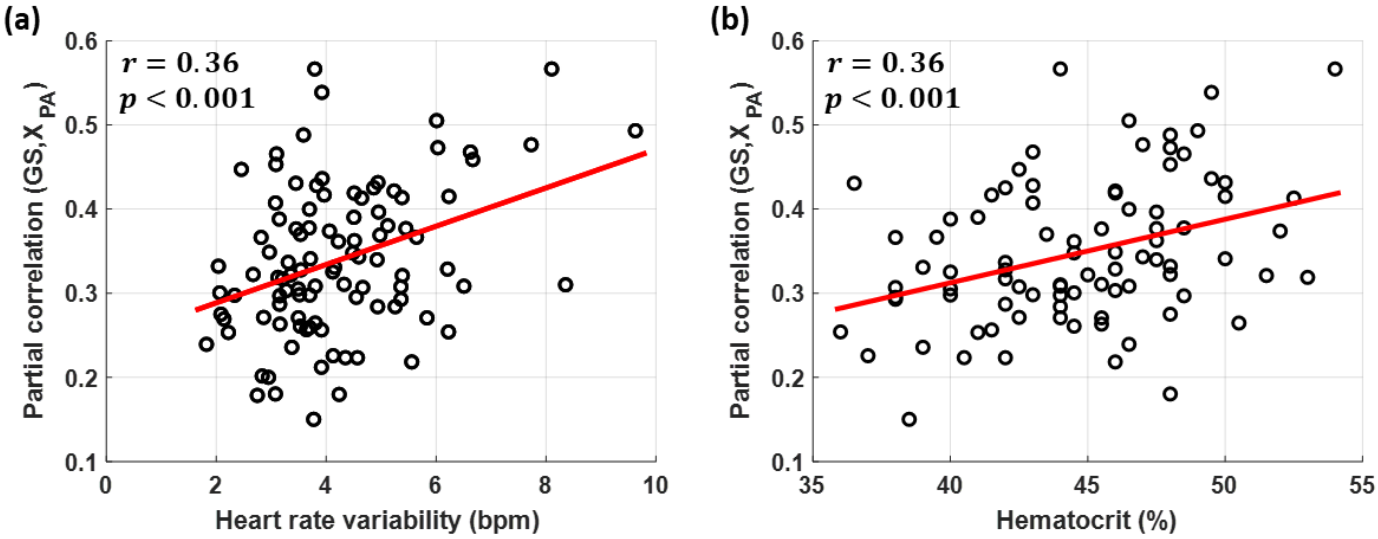


Fig. 8. Scatterplots of the variance explained on the GS with PPG-Amp across subjects with respect to (a) HRV and (b) hematocrit. The hematocrit was measured for 87 out of the 100 subjects. For subjects that had two measures of hematocrit, we used the mean value. The variance explained on the GS from PPG-Amp was based on the partial correlation of GS with the nuisance regressor X_{PA} averaged across the four scans of a subject. Similarly, HRV was averaged across scans within a subject. Interestingly, the higher was the HRV or hematocrit of a subject, the larger was the fraction of variance in the GS explained with PPG-Amp.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4. Discussion

In this study, we have revisited techniques commonly employed in the fMRI literature for modelling low- and high-frequency physiological-related fluctuations and proposed refinements and extensions. Furthermore, we sought to answer whether the low-frequency (~0.1 Hz) oscillations observed in the PPG recordings, referred to here as PPG-Amp, carry important information in the context of physiological noise modelling.

4.1 High-frequency cardiac fluctuations

With regards to high-frequency oscillations induced by cardiac pulsatility, we considered RETROICOR which is a widely used technique proposed by Glover et al. (2000) implemented in various fMRI preprocessing toolboxes such as the physiological noise modelling (PNM) toolbox of FSL (Jenkinson et al., 2012) and the PhysIO SPM toolbox (Kasper et al., 2017). RETROICOR assumes that the pulse-related oscillations are phase-locked to the cardiac cycles. The cardiac cycles are assumed to start and end at time intervals indicated by adjacent peaks in a PPG recording or adjacent R-waves in an electrocardiograph (ECG). RETROICOR essentially uses a Fourier series of order M to define a CPW for each voxel timeseries which is repeated in each cardiac cycle. Depending on how long the period of a cardiac cycle is the CPW is extended or shrinked in time in order to begin and end along with that cycle.

From a systems theory perspective, RETROICOR assumes that the pulse-related oscillations are described by a non-causal system as it requires knowledge of future input values (i.e. the timing of the following peak in PPG) to estimate the output at a specific timepoint. However, as the idea of considering a causal system that depends only on current and past input values may sound more physiologically plausible, we examined the feasibility of a convolution model to capture pulse-related oscillations in fMRI. As in RETROICOR, a Fourier series was used to construct a CPW for each voxel. The CPW in the proposed model plays essentially the role of an impulse response function consistent with the notion of hemodynamic response function that has been adopted in the fMRI literature (Boynton et al., 2012). The input signal is defined as a train of pulses located at the timings of PPG peaks. Two variants of input signals were considered, one with all pulses having equal amplitude (value of one) and a second one where the amplitudes of the pulses matched the amplitudes of the PPG peaks. Note that the peaks observed in PPG are slightly delayed with respect to cardiac contractions due to pulse transit time effects (Allen, 2007). However, this time difference was implicitly accounted for when comparing RETROICOR and CPM by considering for each model the time shift that yielded the best performance.

To compare the performance of the proposed model CPM with RETROICOR, we employed a 3-fold cross-validation framework. In addition, we sought to optimize the model order and determine whether incorporating some time shift in the PPG recordings can improve the performance in terms of variance explained in the voxel timeseries. Indeed, the time shifts of -0.4 s for RETROICOR and -0.9 s for CPM were found to significantly increase the variance explained in fMRI (Fig. 2a-c, Table 1). In the case of RETROICOR, the time shift of -0.4 s is likely related to the fact that we considered the onset of an fMRI volume as the acquisition time for all slices in the volume. In the HCP fMRI data examined here, the TR which corresponds to the time interval between volumes was 0.72 s. However, as the slices of an fMRI volume were being acquired at different times during this time interval, we could consider that the effective time acquisition of a volume (i.e. the time that deviates the least from the acquired times of all slices) was the time indicated by the analog trigger shifted by half TR forward. If we had accounted for the aforementioned effective time acquisition then the optimal time shift needed for RETROICOR would be almost zero. With respect to the half second difference in optimal time shift observed between CPM and RETROICOR, this difference may be related to the fact that the basis functions used in CPM were slighted altered compared to the basis functions used in RETROICOR (specifically, the cosines were subtracted by one) so that the CPWs used in the convolution models, by design, started and ended at zero. Note that the two models showed a similar difference in optimal time shift when used to model the fluctuations on the PPG timeseries itself (Suppl. Fig. 1a).

Significant improvement in the variance explained was also observed when using higher orders of Fourier series than the 2nd order typically used in the literature (Fig. 2d, Table 1). Specifically, for both RETROICOR and CPM, the 6th model order yielded the best goodness-of-fit. Model orders higher than the 6th demonstrated a small decreasing trend suggesting

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that higher orders may be prone to overfitting and, thus, carry the risk of removing signal of interest. While our results demonstrated the superiority of the 6th model order compared to lower order, we acknowledge that the optimal order may vary across datasets, and particularly across datasets with different pulse sequence parameters that determine the degrees of freedom in the data such as the TR and duration of scan. Furthermore, while we did not examine variability of optimal order across brain regions, we speculate that the optimal model order may by higher in regions prone to pulse-related artifacts such as regions close to large arteries, ventricles and venous sinuses.

Among the three models examined in this work for pulse-related fMRI fluctuations, the proposed model CPM_{CA} exhibited the best performance (Fig. 2). As hypothesized, CPM_{CA} performed better than RETROICOR for subjects with high HRV (above ~5 bpm), whereas the performance of the two models was similar for subjects with low HRV. Bear in mind that RETROICOR and CPM_{CA} differ only in that the former assumes a CPW that is phase-locked to the cardiac cycle whereas the latter assumes a CPW of constant duration. Therefore, in the case of a constant HR we would expect the two models to perform equally well.

The rationale for examining the model CPM_{VA} that incorporates the low-frequency fluctuations in the PPG-Amp was to examine whether the varying amplitude of cardiac pulses observed in the PPG (i.e. PPG-Amp) is also present in the BOLD fMRI signal. Compared to RETROICOR and CPM_{CA}, CPM_{VA} explained significantly less variance in the fMRI timeseries suggesting that cardiac pulses measured in the brain with fMRI do not have the same amplitude as the pulses measured on the finger with PPG. We believe that the differences in amplitude variations observed between fMRI and PPG are not due to the different recording sites rather than differences in the physical principles underlying each modality. PPG is based on near-infrared spectroscopy (NIRS) whereby a biological tissue is illuminated with near-infrared light from a laser diode and the light detected by a receiver in a nearby site is analyzed in order to provide information about the compounds present in the illuminated tissue (Delpy and Cope, 1997; Pellicer and Bravo, 2011; Scheeren et al., 2012). Based on the Beer-Lambert law, the attenuation of the light that is measured with PPG depends on the concentrations of oxyhemoglobin (HbO) and deoxyhemoglobin (Hb) as well as on their absorption coefficient for the wavelength of the incident light. Due to that HbO and Hb are characterized by different absorption spectra, it is very likely the a single-wavelength PPG shows different sensitivity to changes of HbO compared to Hb. Therefore, the variations observed in PPG-Amp may be the result of variations in the relative fractions of HbO and Hb without necessarily any changes in the total hemoglobin (i.e. sum of HbO and Hb). On the other hand, while BOLD fMRI is considered to reflect changes in Hb it is also very prone to motion artifacts. Therefore, the pulses in fMRI may originate mainly from vessel expansion and tissue movement due to the propagating blood pressure waves in the arteries rather than fluctuations in Hb. If this is indeed the case then we would expect the pulse waveforms to be independent of the exact composition of blood (in other words, any fluid with similar viscosity to the blood could lead to the same fluctuations) and, therefore, independent of the relative changes in HbO and Hb. Another possible explanation for the differences in pulse amplitude between PPG and fMRI is that while in PPG the signal is linearly proportional to the levels of HbO and Hb, and subsequently to the total hemoglobin and blood volume, in fMRI the signal may present a non-linear relation to blood volume changes. Therefore, variations in PPG-Amp due to changes in blood volume may be reflected differently in fMRI.

In an earlier study we demonstrated that cardiac pulsatility induces systematic biases in FC that to some extent depend on whether the PE direction is left-right or right-left (Xifra-porxas et al., 2020). Here, we provided evidence that the regional sensitivity of fMRI data to cardiac pulsatility depends partly on PE direction (Fig. 4) which may explain why biases in FC due to cardiac pulsatility have this dependence as well. Note that dependence to PE direction was also reported for breathing motion fMRI artifacts by Raj et al. (2001). Based on simulations and experimental data, Raj et al. suggested that magnetic susceptibility variations, caused likely by the expansion of lungs, induce variations in the static magnetic field within the brain being sampled (Raj et al., 2001, 2000). As a result, the spatial encoding during fMRI acquisition is unavoidably affected leading to a shift of the reconstructed image in the PE direction as well as distortion of voxel timeseries with artifact waveforms that depend on both the phase of the breathing cycle and the location of each voxel with respect to the PE direction. In our study, the amplitude of the pulse-related fluctuations varied across voxels depending on their location with respect to the PE direction which may suggest that, similar to breathing motion, the

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mechanism by which pulsatility-induced vessel expansion gives rise to fluctuations in fMRI is partly through local variations in static magnetic field. Note though that as vessel expansion causes also fluid and tissue movement, spin-history effects are also thought to be another source of fluctuations (Caballero-Gaudes and Reynolds, 2017; Murphy et al., 2013).

One can use the proposed model CPM_{CA} to remove fMRI fluctuations due to cardiac pulsatility and facilitate the detection of neural-related activity. However, another potential application of this model is to visualize blood flow pulsatility in cerebral arteries as well as pulsatility-induced CSF movement. There is accumulating evidence that altered cardiac pulsatility in the brain is associated with neurodegenerative diseases such as Alzheimer's disease (Harrison et al., 2018; Iliff et al., 2013; Mestre et al., 2018; Schley et al., 2006). As such, there is a growing interest in developing non-invasive techniques for measuring intracranial pulsatility. When we examined the CPWs (i.e. waveforms of the pulse-related fMRI fluctuations) extracted with CPM_{CA}, we were able, at the group level, to observe patterns that resembled somewhat CSF movement, particularly in areas along the cerebral aqueduct (Suppl. Fig. 2, a video with the temporal dynamics is available on https://doi.org/10.6084/m9.figshare.c.4946799 (Kassinopoulos and Mitsis, 2020)). This finding suggests that CPM_{CA}, combined with a suitably designed fMRI pulse sequence, may be a potential tool for studying the pulsating brain.

4.2 Systemic low-frequency physiological oscillations (SLFOs)

Using a cross-correlation analysis we showed that the low-frequency fluctuations (~0.1 Hz) of GS were preceded by changes in physiological variables, namely the RV, HR and PPG-Amp (Fig. 5). In addition, these three physiological variables were, to some degree, associated to each other. When we considered convolution models to quantify their contributions on the GS, GS was found to share unique variance with each of the three variables with RV being the most influential factor and PPG-Amp the weakest one (Fig. 6b).

The curves of the estimated RRFs and CRFs exhibited significant variability across scans. However, the main trend was similar to the trend reported in our earlier study (Fig. 6a; Kassinopoulos and Mitsis, 2019). Both CRF and RRF illustrated a positive peak at around 2 s followed by a negative peak at 8 s for CRF and 13-14 s for RRF. With regards to the shape of the CRF, as has been suggested in our previous study, the abrupt positive peak may reflect the increase in the blood flow that accompanies the HR increase whereas the negative peak followed a few seconds later may reflect a regulatory feedback mechanism, potentially mediated by a decrease in stroke volume, that aims to bring the blood flow back to its baseline, despite the changes in HR. On the other hand, the early positive peak in RRF may indicate that an increase in breathing activity, either due to increase in breathing rate or breathing volume, leads to an abrupt increase in levels of HbO which eventually results to an increase in the BOLD signal. However, increased breathing activity leads also to a decline in levels of CO_2. And as CO_2 is a strong vasodilator, its decrease leads to a somewhat delayed vasoconstriction which, in turn, reduces the blood flow and also the BOLD signal.

Inspired by the earlier studies of Birn et al. (2008) and Chang et al. (2009), here we introduced an impulse response function, termed photoplethysmographic amplitude response function (PARF), that relates PPG-Amp variations to changes in GS. However, an important note with respect to this convolution model is that the PPG-Amp had to be shifted backward by 5 s to achieve the best fit on the GS. A negative time-shift for RV and HR would not be considered when modelling the SLFOs in the GS as the cardiac and breathing activity are processes thought to drive the fluctuations in the GS and, therefore, cannot be preceded by it. On the other hand, the PPG signal collected from the participant's finger, as with the BOLD signal, can be seen as a hemodynamic signal whose fluctuations are controlled, among others, by the cardiac and breathing activity. Therefore, the -5 s time shift needed for the PPG signal may suggest that the blood pumped by the heart to the aorta first arrives at the brain vasculature and then (through a different branch) at the finger.

The shape of the estimated PARF averaged across all subjects exhibited a somewhat opposite trend compared to CRF and RRF. Specifically, it presented a negative peak at 4.3 s followed by a positive peak at 12.5 s (Fig. 6a). The shape of PARF cannot be easily interpreted as PPG-Amp reflects several processes. Consistent with previous studies (Özbay et al., 2019, 2018), PPG-Amp was found to exhibit an almost instantaneous drop when HR increased (Fig. 5) which may

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

be associated to decreased stroke volume. Moreover, PPG-Amp was found in our data to be reduced during inhalation, a phenomenon well-documented in the literature that has been attributed to changes in intrathoracic pressure resulting also in reduced stroke volume (Meredith et al., 2012). In addition, as PPG is sensitive to changes in HbO and Hb, it also captures slower effects of cardiac and breathing activity related to changes in blood oxygenation and volume. While PARF may be lacking a clear physiological interpretation, the cross-validation analysis conducted in this study revealed that the inclusion of PPG-Amp convolved with PARF in the model of SLFOs substantially improves the goodness-of-fit on the GS compared to considering only HR and RV. Specifically, in the cross-validation analysis the mean correlation increased from 0.65 to 0.67 which was found to be statistically significant ($p<10^{-7}$), while when the model was trained and tested on the same dataset the mean correlation increased from 0.73 to 0.76 ($p<10^{-26}$).

The contribution of PPG-Amp on the GS illustrated variability across subjects with partial correlation between PPG-Amp and GS, when controlling for HR and RV, ranging between 0.15 and 0.57. Subjects with higher HRV presented stronger relationship between GS and PPG-Amp (Fig. 8a; $p<0.001$) supporting the notion that PPG-Amp explains variance on the GS induced partly by HR changes. Furthermore, we observed that the higher was the hematocrit of a subject the larger was the contribution of PPG-Amp on the GS (Fig. 8b; $p<0.001$). Hematocrit which is defined as the proportion of red blood cells in the blood is considered as one of the factors determining the amplitude of the PPG signal even though its exact effect is still not very clear (Fine, 2014; Jubran, 2015; Ochoa and Ohara, 1980). The strong relationship between hematocrit and contribution of PPG-Amp to the GS can be explain as follows: higher levels of hematocrit lead to stronger weighting of HbO and Hb to the PPG signal compared to other compounds. As a result, PPG is more sensitive to changes in oxygenation; hence the increased variance explained in the GS using the amplitude of the PPG pulses for subjects with high hematocrit levels. Note that a positive linear relationship has been previously reported between the amplitude of task-induced BOLD responses and hematocrit (Gustard et al., 2003; Levin et al., 2001) which would suggest that the mean or standard deviation of the GS may differ between subjects with different hematocrit. However, when we examined the mean and standard deviation in the GS, we did not find any strong association with hematocrit (results not shown).

The capability of PPG-Amp to remove SLFOs from fMRI data was first demonstrated by Van Houdt et al. (2010) who showed that removing these fluctuations facilitates the detection of the epileptogenic zone in epileptic patients. Van Houdt et al. reported a low negative correlation between RVT (a measure of breathing activity similar to the measure RV used here) with PPG-Amp (-0.08±0.09) even though the two variables yielded similar spatial maps with regions associated to variations to RVT and PPG-Amp. While we also found similar low values in the cross-correlation analysis, the trends were consistent across subjects (Fig. 5). Importantly, consistent with Van Houdt et al. (2010), we found that the group-level correlation maps with areas associated to HR, RV and PPG-Amp exhibited same spatial patterns (Kassinopoulos and Mitsis, 2019; Suppl. Fig. 4). This may not be surprisingly as we expect physiological processes to affect areas close to the vasculature, and particularly close to draining veins, such as in the occipital cortex, where the concentration of Hb varies the most. That said, the three variables seem to consist of shared but also unique variance, hence the increased variance explained when considering all variables in the analysis.

A main difference between the analysis employed here and the analysis in Van Houdt et al. (2010) is that we accounted for the dynamics for the effects of physiological processes on fMRI data using convolutional models and basis expansion techniques. By doing so, we ensured the plausibility of the estimated PRFs while also keeping the flexibility of the model low. In contrast, Van Houdt et al. (2010) considered 10 lagged versions of the PPG-Amp in the GLM to account for the underlying dynamics. This multi-lagged approach reduces degrees of freedom in the data, particularly in the frequency range where we expect neuronal-related activity (Bright et al., 2017) and, therefore, is prone to removing signal of interest. On a related note, in an earlier study we found that there was no benefit in allowing variability in the shape of the PRFs across voxels (Kassinopoulos and Mitsis, 2019b). This finding suggests that one can reduce the potential of removing signal of interest by modeling the SLFOs in the GS and, subsequently, use the model fit as a nuisance regressor for denoising of the fMRI timeseries as done in (Kassinopoulos and Mitsis, 2019a; Xifra-porxas et al., 2020).

The results of this study were inconsistent with two recent studies from Özbay et al. (2019, 2018) that investigated the role of PPG-Amp in fMRI during sleep. While our data showed a significant negative (cross-)correlation between GS

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and PPG-Amp at -1 s lag time and a positive correlation at 8 s (Fig. 5), Özbay et al. (2019, 2018) found only a positive correlation at about 5-6 s. Moreover, in our data, regions in GM exhibited a negative correlation with PPG-Amp at zero lag and only a few regions in the CSF around the brainstem presented positive correlations, albeit with very low values (Suppl. Fig. 4). In contrast, Özbay et al. (2019, 2018) reported positive correlations in GM and negative correlations in WM. While we find these results puzzling, the inconsistency between the two studies may be somehow explained by the fact that the PPG signal is driven by several physiological processes (e.g. HRV) which in turn exhibit different trends across stages of sleep and wakefulness (Elsenbruch et al., 1999). As the role of PPG-Amp in fMRI has been somewhat neglected in the literature, further research is needed in order to shed light on the mechanisms that may determine the relation between PPG-Amp and GS and how this relation may vary between stages of sleep and wakefulness.

In this study, we have thoroughly studied the pulsatile component of PPG as well as the variations in its amplitude. This component, referred to in the NIRS literature as AC component, is commonly used to obtain measurements of HR and breathing rate (Charlton et al., 2018). However, another important component of the PPG that is not well-documented in the fMRI literature is the variable DC baseline of the PPG signal appeared at frequencies below 0.2 Hz. As a matter of fact, this DC component is typically removed from the MRI pulse oximeters using a high-pass filter to facilitate the visualization of cardiac pulses (this is also likely the case with the PPG signals in HCP as they are lacking low-frequency content). Pulse oximeters considering the full spectrum of the PPG signal and that also illuminate at more than one wavelengths can provide additional physiological variables compared to single-wavelength PPG signals, such as variations in relative changes of HbO and Hb, as well as oxygen saturation (Delpy and Cope, 1997). Due to these properties, the DC baseline in PPG is of great importance in clinical cardiovascular monitoring (Jubran, 2015). Note also that the two-wavelength PPG signal is the fundamental signal exploited in functional NIRS (fNIRS) for the study of brain activity (Tachtsidis and Scholkmann, 2016).

Although fMRI studies do not typically consider the low-frequency (<0.2 Hz) fluctuations of PPG, its potential in physiological noise correction was demonstrated a while ago by Tong and Frederick (2010). Specifically, Tong and Frederick (2010) demonstrated that the low-frequency fluctuations in HbO and Hb, obtained from peripheral NIRS, explained a significant variance in fMRI while in a subsequent study they provided evidence that the variables HbO and Hb explained significantly higher variance compared to nuisance regressors related to HR and breathing patterns (Hocke et al., 2016). As the authors stated, this result may not be surprising as HbO and Hb measured with NIRS are factors directly related to the BOLD signal. While the levels of HbO and Hb measured from the finger are affected by cardiac and breathing activity they are also affected by other processes such as fluctuations in blood pressure and activity of autonomic nervous system that are non-trivial to be measured, especially in the MR environment. Therefore, measurements from peripheral NIRS are in principle able to account for several factors apart from HR and breathing pattern variations. With regards to our results, it is very likely that the unique variance explained by PPG-Amp in the GS could also be captured, if not even better, with the low-frequency fluctuations in NIRS. However, in case that only the PPG is available then the PPG-Amp could be considered in addition to the effects of HR and breathing pattern.

5. Conclusion

Here, we examined noise correction techniques that utilize the PPG signal in order to account for low- and high-frequency physiological fluctuations in fMRI. The CPM_{CA} model was proposed as a physiologically plausible refinement of the commonly used technique RETROICOR. CPM_{CA} employs a convolution framework in a similar manner neural-induced BOLD responses are typically modelled in the fMRI literature (i.e. through convolution with the hemodynamic response function). As initially hypothesized, CPM_{CA} performed equally well with RETROICOR for subjects with relatively stable HR and outperformed the latter for subjects with high variability in HR. The variations in PPG-Amp (i.e. pulse amplitude observed in the PPG) did not seem to covary with the amplitude in the fMRI pulse-related fluctuations. However, PPG-Amp was found to explain a significant amount of variance in SLFOs present in the GS, in addition to variance explained by fluctuations in HR and breathing patterns. Overall, the pulsatile component of the PPG signal explained a large fraction of variance in fMRI related to both low- and high-frequency physiological fluctuations. Scripts for the techniques proposed here are available on git repository https://github.com/mkassinopoulos/Noise_modeling_based_on_PPG.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant 34362 awarded to GDM), the Fonds de la Recherche du Quebec - Nature et Technologies (FRQNT; Team Grant PR191780-2016 awarded to GDM) and the Canada First Research Excellence Fund (awarded to McGill University for the Healthy Brains for Healthy Lives initiative). MK acknowledges funding from Québec Bio-imaging Network (QBIN). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The authors would like to thank Alba Xifra-Porxas for her assistant in the selection of subjects from the HCP database with good quality physiological recordings.

References

Allen, J., 2007. Photoplethysmography and its application in clinical physiological measurement. Physiol. Meas. 28. https://doi.org/10.1088/0967-3334/28/3/R01

Aquino, K.M., Fulcher, B.D., Parkes, L., Sabaroedin, K., Fornito, A., 2019. Identifying and removing widespread signal deflections from fMRI data : Rethinking the global signal regression problem 1–18.

Bandettini, P.A., Wong, E.C., Hinks, R.S., Tikofsky, R.S., Hyde, J.S., 1992. Time course EPI during task activation. Magn Res Med 25, 390–397.

Behzadi, Y., Restom, K., Liau, J., Liu, T.T., 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37, 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Bianciardi, M., Toschi, N., Polimeni, J.R., Evans, K.C., Bhat, H., Keil, B., Rosen, B.R., Boas, D.A., Wald, L.L., 2016. The pulsatility volume index: An indicator of cerebrovascular compliance based on fast magnetic resonance imaging of cardiac and respiratory pulsatility. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374. https://doi.org/10.1098/rsta.2015.0184

Birn, R.M., 2012. The role of physiological noise in resting-state functional connectivity. Neuroimage 62, 864–870. https://doi.org/10.1016/j.neuroimage.2012.01.016

Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31, 1536–48. https://doi.org/10.1016/j.neuroimage.2006.02.048

Birn, R.M., Smith, M. a., Jones, T.B., Bandettini, P. a., 2008. The respiration response function: The temporal dynamics of fMRI signal fluctuations related to changes in respiration. Neuroimage 40, 644–654. https://doi.org/10.1016/j.neuroimage.2007.11.059

Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. 34, 537–541. https://doi.org/10.1002/mrm.1910340409

Boynton, G.M., Engel, S.A., Heeger, D.J., 2012. Linear systems analysis of the fMRI signal. Neuroimage 62, 975–984. https://doi.org/10.1016/j.neuroimage.2012.01.082

Bright, M.G., Tench, C.R., Murphy, K., 2017. Potential pitfalls when denoising resting state fMRI data using nuisance regression. Neuroimage 154, 159–168. https://doi.org/10.1016/j.neuroimage.2016.12.027

Caballero-Gaudes, C., Reynolds, R.C., 2017. Methods for cleaning the BOLD fMRI signal. Neuroimage 154, 128–149. https://doi.org/10.1016/j.neuroimage.2016.12.018

Carbonell, F., Bellec, P., Shmuel, A., 2014. Quantification of the impact of a confounding variable on functional connectivity confirms anti-correlated networks in the resting-state. Neuroimage 86, 343–353.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://doi.org/10.1016/j.neuroimage.2013.10.013

Carbonell, F., Bellec, P., Shmuel, A., 2011. Global and System-Specific Resting-State fMRI Fluctuations Are Uncorrelated: Principal Component Analysis Reveals Anti-Correlated Networks. Brain Connect. 1, 496–510. https://doi.org/10.1089/brain.2011.0065

Chang, C., Cunningham, J.P., Glover, G.H., 2009. Influence of heart rate on the BOLD signal: The cardiac response function. Neuroimage 44, 857–869. https://doi.org/10.1016/j.neuroimage.2008.09.029

Chang, C., Glover, G.H., 2009. Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. Neuroimage 47, 1448–1459. https://doi.org/10.1016/j.neuroimage.2009.05.012

Charlton, P.H., Birrenkott, D.A., Bonnici, T., Pimentel, M.A.F., Johnson, A.E.W., Alastruey, J., Tarassenko, L., Watkinson, P.J., Beale, R., Clifton, D.A., 2018. Breathing Rate Estimation from the Electrocardiogram and Photoplethysmogram: A Review. IEEE Rev. Biomed. Eng. 11, 2–20. https://doi.org/10.1109/RBME.2017.2763681

Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G.L., Ruparel, K., Shinohara, R.T., Elliott, M.A., Eickhoff, S.B., Davatzikos, C., Gur, R.C., Gur, R.E., Bassett, D.S., Satterthwaite, T.D., 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. Neuroimage 154, 174–187. https://doi.org/10.1016/j.neuroimage.2017.03.020

Dagli, M.S., Ingeholm, J.E., Haxby, J. V, 1999. Localization of cardiac-induced signal change in fMRI. Neuroimage 9, 407–415. https://doi.org/10.1006/nimg.1998.0424

Delpy, D.T., Cope, M., 1997. Quantification in tissue near–infrared spectroscopy. Philos. Trans. R. Soc. London. Ser. B Biol. Sci. 352, 649–659. https://doi.org/10.1098/rstb.1997.0046

Elsenbruch, S., Harnish, M.J., Orr, W.C., 1999. Heart rate variability during waking and sleep in healthy males and females. Sleep 22, 1067–1071. https://doi.org/10.1093/sleep/22.8.1067

Falahpour, M., Nalci, A., Liu, T.T., 2018. The Effects of Global Signal Regression on Estimates of Resting-State Blood Oxygen-Level-Dependent Functional Magnetic Resonance Imaging and Electroencephalogram Vigilance Correlations. Brain Connect. 8, 618–627. https://doi.org/10.1089/brain.2018.0645

Falahpour, M., Refai, H., Bodurka, J., 2013. Subject specific BOLD fMRI respiratory and cardiac response functions obtained from global signal. Neuroimage 72, 252–264. https://doi.org/10.1016/j.neuroimage.2013.01.050

Fine, I., 2014. The optical origin of the PPG signal. Saratov Fall Meet. 2013 Opt. Technol. Biophys. Med. XV; Laser Phys. Photonics XV 9031, 903103. https://doi.org/10.1117/12.2051228

Fultz, N.E., Bonmassar, G., Setsompop, K., Stickgold, R.A., Rosen, B.R., Polimeni, J.R., Lewis, L.D., 2019. Coupled electrophysiological, hemodynamic, and cerebrospinal fluid oscillations in human sleep. Science (80-.). 366, 628–631. https://doi.org/10.1126/science.aax5440

Glasser, M.F., Coalson, T.S., Bijsterbosch, J.D., Harrison, S.J., Harms, M.P., Anticevic, A., Van Essen, D.C., Smith, S.M., 2018. Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. Neuroimage 181, 692–717. https://doi.org/10.1016/j.neuroimage.2018.04.076

Glasser, M.F., Smith, S.M., Marcus, D.S., Andersson, J.L.R., Auerbach, E.J., Behrens, T.E.J., Coalson, T.S., Harms, M.P., Jenkinson, M., Moeller, S., Robinson, E.C., Sotiropoulos, S.N., Xu, J., Yacoub, E., Ugurbil, K., Van Essen, D.C., 2016. The Human Connectome Project's neuroimaging approach. Nat. Neurosci. 19, 1175–87. https://doi.org/10.1038/nn.4361

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Glover, G.H., Li, T.-Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn. Reson. Med. 44, 162–167. https://doi.org/10.1002/1522-

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2594(200007)44:1<162::AID-MRM23>3.0.CO;2-E

Gustard, S., Williams, E.J., Hall, L.D., Pickard, J.D., Carpenter, T.A., 2003. Influence of baseline hematocrit on between-subject BOLD signal change using gradient echo and asymmetric spin echo EPI. Magn. Reson. Imaging 21, 599–607. https://doi.org/10.1016/S0730-725X(03)00083-3

Harrison, I.F., Siow, B., Akilo, A.B., Evans, P.G., Ismail, O., Ohene, Y., Nahavandi, P., Thomas, D.L., Lythgoe, M.F., Wells, J.A., 2018. Non-1 invasive imaging of CSF-mediated brain clearance pathways via assessment of perivascular fluid movement with diffusion tensor MRI. Elife 7, 1–14. https://doi.org/10.7554/eLife.34028

Harvey, A.K., Pattinson, K.T.S., Brooks, J.C.W., Mayhew, S.D., Jenkinson, M., Wise, R.G., 2008. Brainstem functional magnetic resonance imaging: Disentangling signal from physiological noise. J. Magn. Reson. Imaging 28, 1337–1344. https://doi.org/10.1002/jmri.21623

Hocke, L.M., Tong, Y., Lindsey, K.P., de B. Frederick, B., 2016. Comparison of peripheral near-infrared spectroscopy low-frequency oscillations to other denoising methods in resting state functional MRI with ultrahigh temporal resolution. Magn. Reson. Med. 1707, 1697–1707. https://doi.org/10.1002/mrm.26038

Iadecola, C., 2017. The Neurovascular Unit Coming of Age: A Journey through Neurovascular Coupling in Health and Disease. Neuron 96, 17–42. https://doi.org/10.1016/j.neuron.2017.07.030

Iliff, J.J., Wang, M., Zeppenfeld, D.M., Venkataraman, A., Plog, B.A., Liao, Y., Deane, R., Nedergaard, M., 2013. Cerebral arterial pulsation drives paravascular CSF-Interstitial fluid exchange in the murine brain. J. Neurosci. 33, 18190–18199. https://doi.org/10.1523/JNEUROSCI.1592-13.2013

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62, 782–90. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156. https://doi.org/10.1016/S1361-8415(01)00036-6

Jubran, A., 2015. Pulse oximetry. Crit. Care 19. https://doi.org/10.1186/s13054-015-0984-8

Kasper, L., Bollmann, S., Diaconescu, A.O., Hutton, C., Heinzle, J., Iglesias, S., Hauser, T.U., Sebold, M., Manjaly, Z.M., Pruessmann, K.P., Stephan, K.E., 2017. The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. J. Neurosci. Methods 276, 56–72. https://doi.org/10.1016/j.jneumeth.2016.10.019

Kassinopoulos, M., Mitsis, G.D., 2020. Figures produced in Kassinopoulos & Mitsis (2020) - Physiological Noise Modeling in fMRI based on the pulsatile component of photoplethysmograph [WWW Document]. https://doi.org/10.6084/m9.figshare.c.4946799

Kassinopoulos, M., Mitsis, G.D., 2019a. White Matter Denoising Improves the Identifiability of Large-Scale Networks and Reduces the Effects of Motion in fMRI Functional Connectivity 1–47.

Kassinopoulos, M., Mitsis, G.D., 2019b. Identification of physiological response functions to correct for fluctuations in resting-state fMRI related to heart rate and respiration. Neuroimage 202, 116150. https://doi.org/10.1016/j.neuroimage.2019.116150

Kisler, K., Nelson, A.R., Montagne, A., Zlokovic, B. V., 2017. Cerebral blood flow regulation and neurovascular dysfunction in Alzheimer disease. Nat. Rev. Neurosci. 18, 419–434. https://doi.org/10.1038/nrn.2017.48

Kiviniemi, V., Wang, X., Korhonen, V., Keinänen, T., Tuovinen, T., Autio, J., Levan, P., Keilholz, S., Zang, Y.F., Hennig, J., Nedergaard, M., 2016. Ultra-fast magnetic resonance encephalography of physiological brain activity-Glymphatic pulsation mechanisms? J. Cereb. Blood Flow Metab. 36, 1033–1045. https://doi.org/10.1177/0271678X15622047

Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E., Weisskoff, R.M., Poncelet, B.P., Kennedy, D.N., Hoppel, B.E., Cohen, M.S., Turner, R., Cheng -, H.M., Brady, T.J., Rosen, B.R., 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc. Natl. Acad. Sci. U. S. A. 89, 5675–5679.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://doi.org/10.1073/pnas.89.12.5675

Levin, J.M., Frederick, B.D.B., Ross, M.H., Fox, J.F., Von Rosenberg, H.L., Kaufman, M.J., Lange, N., Mendelson, J.H., Cohen, B.M., Renshaw, P.F., 2001. Influence of baseline hematocrit and hemodilution on BOLD fMRI activation. Magn. Reson. Imaging 19, 1055–1062. https://doi.org/10.1016/S0730-725X(01)00460-X

Liu, T.T., 2016. Noise contributions to the fMRI signal: An overview. Neuroimage 143, 141–151. https://doi.org/10.1016/j.neuroimage.2016.09.008

Liu, T.T., Nalci, A., Falahpour, M., 2017. The global signal in fMRI: Nuisance or Information? Neuroimage 150, 213–229. https://doi.org/10.1016/j.neuroimage.2017.02.036

Meredith, D.J., Clifton, D., Charlton, P., Brooks, J., Pugh, C.W., Tarassenko, L., 2012. Photoplethysmographic derivation of respiratory rate: A review of relevant physiology. J. Med. Eng. Technol. 36, 1–7. https://doi.org/10.3109/03091902.2011.638965

Mestre, H., Tithof, J., Du, T., Song, W., Peng, W., Sweeney, A.M., Olveda, G., Thomas, J.H., Nedergaard, M., Kelley, D.H., 2018. Flow of cerebrospinal fluid is driven by arterial pulsations and is reduced in hypertension. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-07318-3

Mitsis, G.D., Zhang, R., Levine, B.D., Tzanalaridou, E., Katritsis, D.G., Marmarelis, V.Z., 2009. Autonomic neural control of cerebral hemodynamics. IEEE Eng. Med. Biol. Mag. 28, 54–62. https://doi.org/10.1109/MEMB.2009.934908

Murphy, K., Birn, R.M., Bandettini, P.A., 2013. Resting-state fMRI confounds and cleanup. Neuroimage 80, 349–359. https://doi.org/10.1016/j.neuroimage.2013.04.001

Murphy, K., Fox, M.D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. Neuroimage 154, 169–173. https://doi.org/10.1016/j.neuroimage.2016.11.052

Nalci, A., Luo, W., Liu, T.T., 2019a. Nuisance effects in inter-scan functional connectivity estimates before and after nuisance regression. Neuroimage 202, 116005. https://doi.org/10.1016/j.neuroimage.2019.07.018

Nalci, A., Rao, B.D., Liu, T.T., 2019b. Nuisance effects and the limitations of nuisance regression in dynamic functional connectivity fMRI. Neuroimage 184, 1005–1031. https://doi.org/10.1016/j.neuroimage.2018.09.024

Ochoa, W., Ohara, I., 1980. The effect of hematocrit on photoelectric plethysmogram. Tohoku J. Exp. Med. 132, 413–9. https://doi.org/10.1620/tjem.132.413

Ogawa, S., Lee, T.M., Kay, A.R., Tank, D.W., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc. Natl. Acad. Sci. 87, 9868–9872. https://doi.org/10.1073/pnas.87.24.9868

Özbay, P.S., Chang, C., Picchioni, D., Mandelkow, H., Chappel-Farley, M.G., van Gelderen, P., de Zwart, J.A., Duyn, J., 2019. Sympathetic activity contributes to the fMRI signal. Commun. Biol. 2, 421. https://doi.org/10.1038/s42003-019-0659-0

Özbay, P.S., Chang, C., Picchioni, D., Mandelkow, H., Moehlman, T.M., Chappel-Farley, M.G., van Gelderen, P., de Zwart, J.A., Duyn, J.H., 2018. Contribution of systemic vascular effects to fMRI activity in white matter. Neuroimage 176, 541–549. https://doi.org/10.1016/j.neuroimage.2018.04.045

Parkes, L., Fulcher, B., Yücel, M., Fornito, A., 2018. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. Neuroimage 171, 415–436. https://doi.org/10.1016/j.neuroimage.2017.12.073

Pattinson, K.T.S., Mitsis, G.D., Harvey, A.K., Jbabdi, S., Dirckx, S., Mayhew, S.D., Rogers, R., Tracey, I., Wise, R.G., 2009. Determination of the human brainstem respiratory control network and its cortical connections in vivo using functional and structural imaging. Neuroimage 44, 295–305. https://doi.org/10.1016/j.neuroimage.2008.09.007

Pellicer, A., Bravo, M. del C., 2011. Near-infrared spectroscopy: A methodology-focused review. Semin. Fetal Neonatal

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Med. 16, 42–49. https://doi.org/10.1016/j.siny.2010.05.003

Power, J.D., Plitt, M., Laumann, T.O., Martin, A., 2017. Sources and implications of whole-brain fMRI signals in humans. Neuroimage 146, 609–625. https://doi.org/10.1016/j.neuroimage.2016.09.038

Power, J.D., Schlaggar, B.L., Petersen, S.E., 2015. Recent progress and outstanding issues in motion correction in resting state fMRI. Neuroimage 105, 536–551. https://doi.org/10.1016/j.neuroimage.2014.10.044

Prokopiou, P.C., Pattinson, K.T.S., Wise, R.G., Mitsis, G.D., 2019. Modeling of dynamic cerebrovascular reactivity to spontaneous and externally induced CO2 fluctuations in the human brain using BOLD-fMRI. Neuroimage 186, 533–548. https://doi.org/10.1016/j.neuroimage.2018.10.084

Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F., 2015. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage 112, 267–277. https://doi.org/10.1016/j.neuroimage.2015.02.064

Raj, D., Anderson, A.W., Gore, J.C., 2001. Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. Phys. Med. Biol. 46, 3331. https://doi.org/10.1088/0031-9155/46/12/318

Raj, D., Paley, D.P., Anderson, A.W., Kennan, R.P., Gore, J.C., 2000. A model for susceptibility artefacts from respiration in functional echo-planar magnetic resonance imaging. Phys. Med. Biol. 45, 3809–3820. https://doi.org/10.1088/0031-9155/45/12/321

Reisner, A., Shaltis, P.A., McCombie, D., Asada, H.H., 2008. Utility of the Photoplethysmogram in Circulatory Monitoring. Anesthesiology 108, 950–958. https://doi.org/10.1097/ALN.0b013e31816c89e1

Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90, 449–468. https://doi.org/10.1016/j.neuroimage.2013.11.046

Scheeren, T.W.L., Schober, P., Schwarte, L.A., 2012. Monitoring tissue oxygenation by near infrared spectroscopy (NIRS): Background and current applications. J. Clin. Monit. Comput. 26, 279–287. https://doi.org/10.1007/s10877-012-9348-y

Schley, D., Carare-Nnadi, R., Please, C.P., Perry, V.H., Weller, R.O., 2006. Mechanisms to explain the reverse perivascular transport of solutes out of the brain. J. Theor. Biol. 238, 962–974. https://doi.org/10.1016/j.jtbi.2005.07.005

Shmueli, K., van Gelderen, P., de Zwart, J. a., Horovitz, S.G., Fukunaga, M., Jansma, J.M., Duyn, J.H., 2007. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. Neuroimage 38, 306–320. https://doi.org/10.1016/j.neuroimage.2007.07.037

Tachtsidis, I., Scholkmann, F., 2016. False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. Neurophotonics 3, 031405. https://doi.org/10.1117/1.NPh.3.3.031405

Tong, Y., Frederick, B. deB, 2010. 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, 553–564. https://doi.org/10.1016/j.neuroimage.2010.06.049

Tong, Y., Hocke, L.M., Frederick, B.B., 2019. Low frequency systemic hemodynamic “noise” in resting state BOLD fMRI: Characteristics, causes, implications, mitigation strategies, and applications. Front. Neurosci. 13. https://doi.org/10.3389/fnins.2019.00787

Van Dijk, K.R.A., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W., Buckner, R.L., 2010. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J. Neurophysiol. 103, 297–321. https://doi.org/10.1152/jn.00783.2009

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-Minn Human Connectome Project: An overview. Neuroimage 80, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Van Houdt, P.J., Ossenblok, P.P.W., Boon, P.A.J.M., Leijten, F.S.S., Velis, D.N., Stam, C.J., De Munck, J.C., 2010. Correction for pulse height variability reduces physiological noise in functional MRI when studying spontaneous brain activity. Hum. Brain Mapp. 31, 311–325. https://doi.org/10.1002/hbm.20866

Wagshul, M.E., Eide, P.K., Madsen, J.R., 2011. The pulsating brain: A review of experimental and clinical studies of intracranial pulsatility. Fluids Barriers CNS 8, 5. https://doi.org/10.1186/2045-8118-8-5

Whittaker, J.R., Driver, I.D., Venzi, M., Bright, M.G., Murphy, K., Chen, J., Whittaker, J.R., 2019. Cerebral Autoregulation Evidenced by Synchronized Low Frequency Oscillations in Blood Pressure and Resting-State fMRI 13, 1–12. https://doi.org/10.3389/fnins.2019.00433

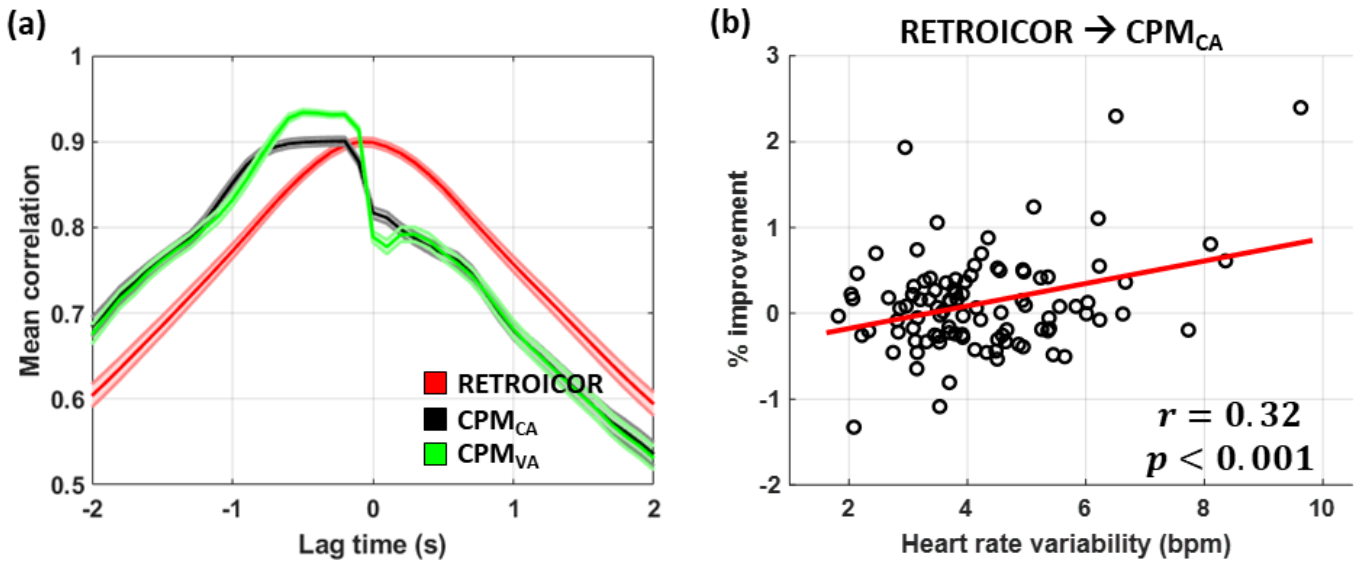
Wise, R.G., Ide, K., Poulin, M.J., Tracey, I., 2004. Resting fluctuations in arterial carbon dioxide induce significant low frequency variations in BOLD signal. Neuroimage 21, 1652–1664. https://doi.org/10.1016/j.neuroimage.2003.11.025

Xifra-porxas, A., Kassinopoulos, M., Mitsis, G.D., 2020. Physiological and head motion signatures in static and time-varying functional connectivity and their subject discriminability. https://doi.org/https://doi.org/10.1101/2020.02.04.934554

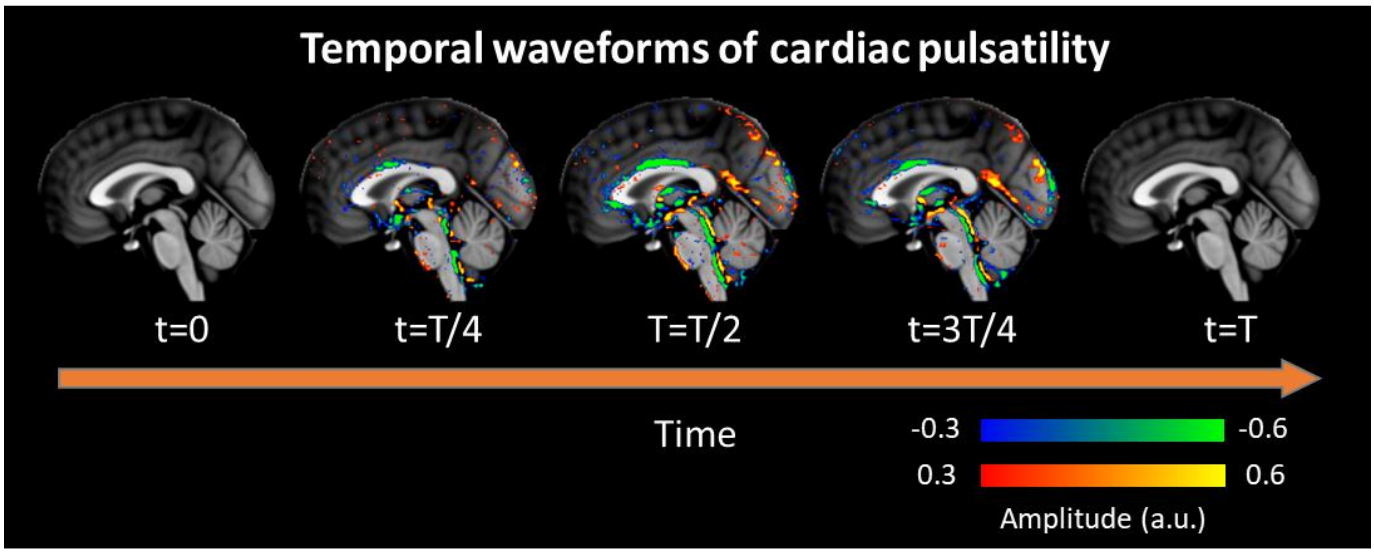
Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57. https://doi.org/10.1109/42.906424

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Material

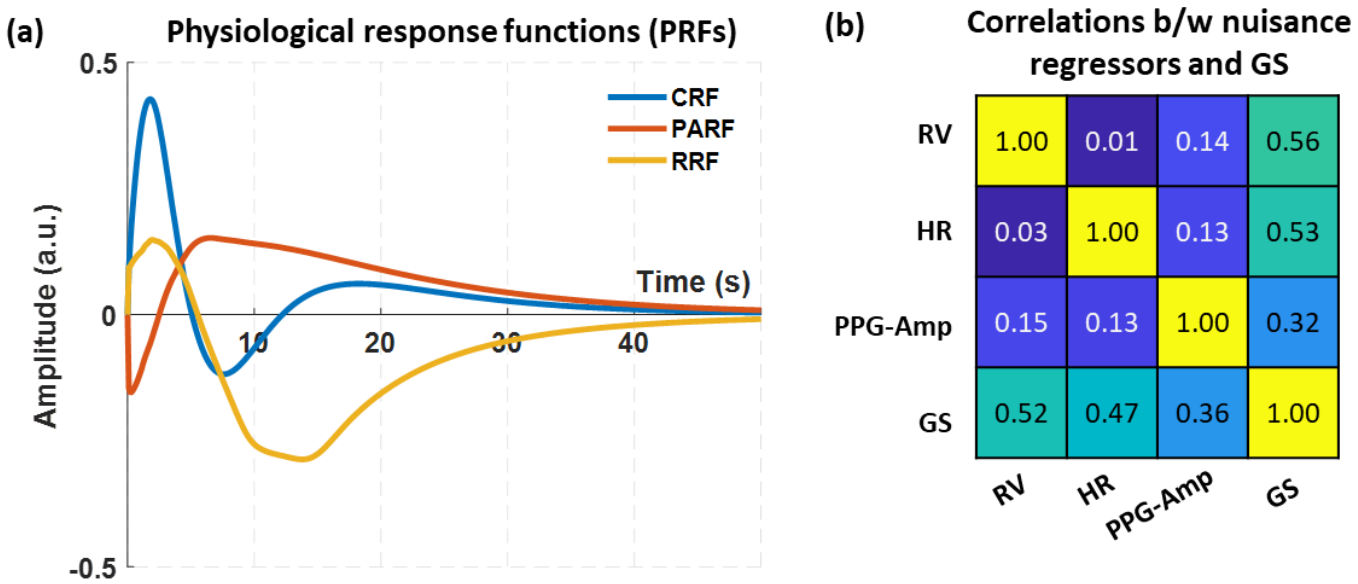


Suppl. Fig. 1. Performance of pulse-related models with respect to variance explained in the PPG. (a) Cross-correlation averaged across subjects for RETROICOR, CPM_{CA} and CPM_{VA}. CPM_{VA} exhibited significantly better performance compared to the other two models as it accounted for the low-frequency fluctuations of PPG. Note that while RETROICOR and CPM_{CA} exhibited similar maximum mean cross-correlation, the peak of the latter was broader (~0.5 s) compared to the peak of the former (~0.1 s). (b) Relative percentage (%) improvement with respect to HRV when comparing the proposed model CPM_{CA} with RETROICOR. The stronger were the HR variations within a scan the larger was the improvement achieved with CPM_{CA} compared to RETROICOR (p<0.001). The comparison was performed between these two models as these models exhibited better performance than RETROICOR when considering the variance explained in the fMRI data (Fig. 3).



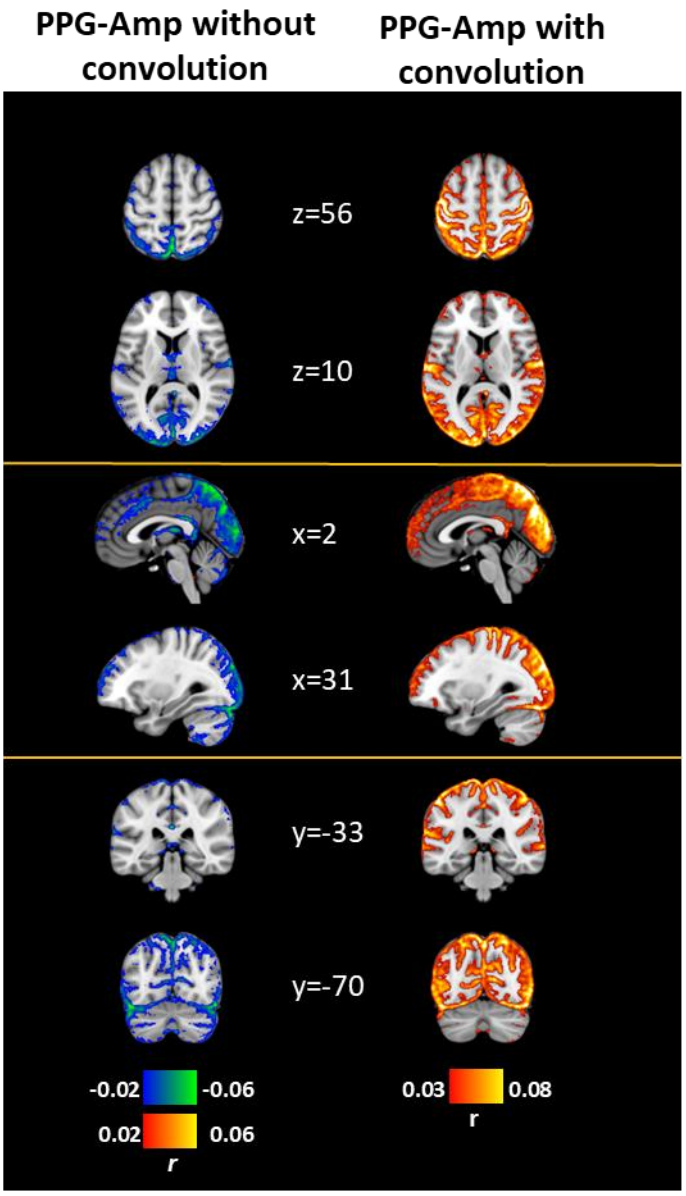
Suppl. Fig. 2. Cardiac pulsatility waveforms (CPWs) averaged across all subjects (N=100; only scans with left-right PE direction from the first session were included). Several areas such as the 3rd ventricle and the cerebral aqueduct as well as areas in the anterior and posterior cortex exhibited similar CPWs across subjects. A video presenting these waveforms induced by cardiac pulsatility in BOLD fMRI can be found on https://doi.org/10.6084/m9.figshare.c.4946799 (Kassinopoulos and Mitsis, 2020). Intriguingly, these CPWs reconstructed with CPM_{CA} revealed different dynamics across regions which may be related in some way to fluid movements.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Suppl. Fig. 3. Estimated physiological response functions (PRFs) when considering the original PPG-Amp (PA_0). (a) PRFs averaged across all subjects and scans using weighted average with the correlation between GS and the predicted output of the model (i.e. SLFOs) for each scan as a weighted coefficient. (b) Correlation between nuisance regressors (i.e. X_{RV}, X_{HR} and X_{PA}) and GS, averaged across all scans. The lower-diagonal elements correspond to correlations whereas the upper-diagonal elements correspond to partial correlations. The partial correlations between pairs of the three nuisance regressors did not control for GS variations as GS is not considered to affect the three associated physiological variables. Note that as in this analysis we do not compare models, the cross-validation framework was omitted. CRF and RRF exhibited a positive peak at around 2 s followed by a negative peak at 8 s for CRF and 13-14 s for RRF. PARF was characterized by a negative peak at 0.2 s followed by a positive peak at 6.5 s. All PRFs demonstrated a slow decay that approximated zero at around 40 s. While the nuisance regressors demonstrated relatively low correlations between them ranging from 0.03 to 0.15, all of them exhibited high correlation with the GS (≥0.36). Similar observations were made for the partial correlations.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.01.128306; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Suppl. Fig. 4. Contribution of PPG-Amp variations in fMRI averaged across all subjects. (1st column) Correlation maps related to the original PPG-Amp variable. (2nd column) Correlation maps related to the nuisance regressor X_{PA} extracted by shifting the PPG-Amp back in time by 5 s and convolving it with the PARF. The correlation maps were estimated on fMRI data corrected for head and breathing motion as well as cardiac pulsatility (6th order of CPM). We observe that both the original PPG-Amp variable and its associated nuisance regressor explained variance in widespread regions across GM. However, as we can see, in the case of the nuisance regressor, the correlation values were significantly higher.