Extraction of cardiac-related signals from a suprasternal pressure sensor during sleep

Document status and date:
Published: 01/03/2023

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

IPEM

**Institute of Physics and
Engineering in Medicine**

PAPER • OPEN ACCESS

Extraction of cardiac-related signals from a suprasternal pressure sensor during sleep

To cite this article: Luca Cerina *et al* 2023 *Physiol. Meas.* **44** 035002

View the article online for updates and enhancements.

Physiological Measurement

IPEM
Institute of Physics and
Engineering in Medicine

PAPER

Extraction of cardiac-related signals from a suprasternal pressure sensor during sleep

Luca Cerina[1,*] , Gabriele B Papini[1,2] , Pedro Fonseca[1,2] , Sebastiaan Overeem[1,3],
Johannes P van Dijk[1,3] and Rik Vullings[1]

1   Electrical Engineering,Technische Universiteit Eindhoven, Eindhoven, Noord Brabant, The Netherlands
2   Philips Research, Eindhoven, Noord Brabant, The Netherlands
3   Center for Sleep Medicine, Kempenhaeghe Foundation, Heeze, Noord Brabant, The Netherlands
*   Author to whom any correspondence should be addressed.

E-mail: l.cerina@tue.nl, gabriele.papini@philips.com, pedro.fonseca@philips.com, overeems@kempenhaeghe.nl, dijkh@
kempenhaeghe.nl and r.vullings@tue.nl

Abstract

Objective. The accurate detection of respiratory effort during polysomnography is a critical element in the diagnosis of sleep-disordered breathing conditions such as sleep apnea. Unfortunately, the sensors currently used to estimate respiratory effort are either indirect and ignore upper airway dynamics or are too obtrusive for patients. One promising alternative is the suprasternal notch pressure (SSP) sensor: a small element placed on the skin in the notch above the sternum within an airtight capsule that detects pressure swings in the trachea. Besides providing information on respiratory effort, the sensor is sensitive to small cardiac oscillations caused by pressure perturbations in the carotid arteries or the trachea. While current clinical research considers these as redundant noise, they may contain physiologically relevant information. *Approach.* We propose a method to separate the signal generated by cardiac activity from the one caused by breathing activity. Using only information available from the SSP sensor, we estimate the heart rate and track its variations, then use a set of tuned filters to process the original signal in the frequency domain and reconstruct the cardiac signal. We also include an overview of the technical and physiological factors that may affect the quality of heart rate estimation. The output of our method is then used as a reference to remove the cardiac signal from the original SSP pressure signal, to also optimize the assessment of respiratory activity. We provide a qualitative comparison against methods based on filters with fixed frequency cutoffs. *Main results.* In comparison with electrocardiography (ECG)-derived heart rate, we achieve an agreement error of 0.06 ± 5.09 bpm, with minimal bias drift across the measurement range, and only 6.36% of the estimates larger than 10 bpm. *Significance.* Together with qualitative improvements in the characterization of respiratory effort, this opens the development of novel portable clinical devices for the detection and assessment of sleep disordered breathing.

1. Introduction

Several disorders are affecting breathing during sleep. Obstructive sleep apnea (OSA) is the most common manifestation, with a complex pathophysiology and diverse outcomes that range from increased cardiovascular risk to excessive daytime sleepiness (Lévy *et al* 2015). There is a growing interest in portable diagnostic devices, or home sleep apnea tests (HSAT) (Abrahamyan *et al* 2017), to reduce costs of OSA screening, but despite recent technical advancements they still provide an incomplete view of OSA mechanisms. To obtain a complete perspective of brain, cardiac, and respiratory activity, and body movements altered by sleep disorders, an overnight polysomnography (PSG) recording conducted in specialized sleep clinics remains the gold standard.

Table 1. Demographics of the participants.

Variable		Mean ± std or median and IQR Median and IQR	(Range)
Sex	[#]	43 (57 male)	
Age	[Years]	47 ± 16	$(18:80)$
BMI[a,e]	[kg m^{-2}]	25.8 ± 3.6	$(19:36)$
TST[b]	[Minutes]	418 ± 69	$(241:622)$
average HR[c]	[bpm]	63 ± 10	$(43:106)$
AHI[d]	[Events h^{-1}]	10.9, 18.8	$(0:109.9)$

[a] Body-Mass Index.

[b] Total Sleep Time.

[c] Heart Rate.

[d] Apnea-Hypopnea Index.

[e] Available only for 34 participants.

Concerning respiratory mechanics, PSG provides two types of signals: oronasal airflow, measured with a thermistor or pressure sensor, and respiratory effort, representing the work of breathing during inspiration and expiration. PSG can rely on two methods to measure respiratory effort: respiratory inductance plethysmography (RIP) belts and, albeit less common, esophageal pressure (Pes) sensors. RIP belts measure thoracoabdominal displacement during breathing and are easy to use but account only for thoracic and abdominal movements. These belts measure the respiratory effort indirectly through continuous phase relationships. Conversely, Pes provides a more direct measure of respiratory effort by assessing intrathoracic pressure swings. Despite being the recommended gold standard (Berry *et al* 2012), Pes sensors are invasive, difficult to place accurately, and poorly tolerated by patients during sleep (Brochard 2014, Glos *et al* 2018).

The suprasternal notch pressure (SSP) sensor represents a potential non-invasive alternative to the esophageal sensor, with comparable reliability in both adults (Glos *et al* 2018) and children (Amaddeo *et al* 2016). The sensor consists of a small pressure-sensitive element placed inside an airtight capsule on the skin notch above the trachea. Due to their specific position, researchers observed that both the Pes and SSP sensors measure more than one physiological signal (Ayappa *et al* 1999, Glos *et al* 2018). Specifically, the SSP signal comprises mainly information about respiratory effort and high-frequency content associated with swallowing and snoring. Upon careful inspection, a component with a smaller amplitude is also visible, and quite evident during central apneas when the respiratory effort is completely absent. This component has been associated with cardiac activity (Suarez-Sipmann *et al* 2012), hence dubbed *cardiogenic* or *cardiac* oscillations. Some examples of the presence of cardiac oscillations during normal respiration and apneic events are illustrated in appendix A.

Typically, the SSP sensor is used to analyse only respiratory activity, and cardiac oscillations are often filtered out (Glos *et al* 2018, Mukhopadhyay *et al* 2020). However, we hypothesize that the content of these oscillations may provide additional insights into cardio-respiratory physiology and is therefore worthy of further analysis.

We propose a method to extract the cardiac signal using only the pressure signal of the SSP sensor. Our contribution is an algorithm capable of estimating heart rate (HR) as the fundamental frequency of cardiac oscillations in the SSP signal and a method to separate these oscillations from the respiratory signal with a tuned filterbank. We tested the accuracy of our estimates against the HR measured by the ECG signal. As a secondary outcome of our research, we perform a qualitative comparison of the respiratory effort signal filtered with our method. Since the extraction of cardiac oscillations from SSP or Pes is not treated extensively in literature, we discuss the shortcomings of the proposed method and challenges encountered, including technical or physiological factors that may influence the quality of the signals.

2. Methods

2.1. Dataset

To test and develop our method we employed full single-night PSG recordings which are part of SOMNIA (van Gilst *et al* 2019), a clinical database designed to facilitate research on sleep disorders and unobtrusive monitoring of sleep. We used a subset of 100 recordings where the PSG included a synchronized recording of the SSP sensor. In this subset each recording is unique participant-wise. Demographics of participants are shown in table 1 together with total sleep time (TST) and apnea-hypopnea index (AHI). The values are expressed as average ± standard deviation and range. AHI is expressed as median and inter-quartile range (IQR) (range) due to its skewed distribution.

Figure 1. Time and frequency domain representation of a signal with a loss of dynamics due to sealing failure. Respiration rate (0.25 Hz), fundamental frequency (1.25 Hz) and 2nd harmonic of heart rate (2.5 Hz) are visible before the breaking point.
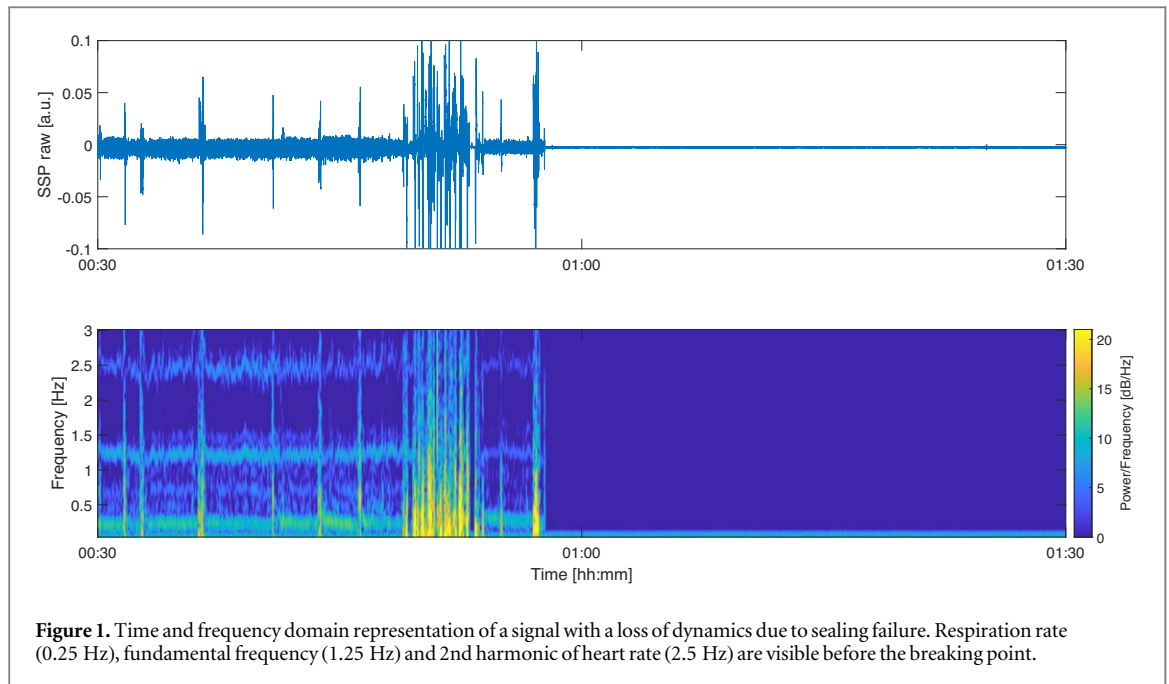
In our dataset, 71 participants out of 100 presented an AHI above 5 events h^{-1} (median 16.4), which is the current clinical threshold for OSA diagnosis. Concerning the manifestation of OSA, 94 participants had more obstructive apneas and hypopneas with respect to central and mixed apneas, with a median ratio of 99.3%. Among the remaining six, four presented more central and mixed apneas than obstructive events but had very low AHI below 1 events h^{-1}. The last two participants had more central or mixed events than obstructive ones, a severe AHI above 40 events h^{-1}, but also respiratory instability in the form of Cheyne-Stokes respiration patterns. The electrocardiography (ECG) Lead II data available in the PSG is used as ground truth to test the estimated heart rate and for the quantification of the average HR in table 1.

2.2. SSP signal characterization

We can represent the raw signal S_{SSP} as the following:

$$S_{\text{SSP}} = S_{\text{Resp}} + S_{\text{Cardio}} + S_{\text{Audio}} + \varepsilon. \tag{1}$$

S_{Resp} represents the pressure swings caused by respiratory activity, while S_{Cardio} contains the pressure waves coming from carotid and pulmonary arteries. S_{Audio} represents noise in the audio frequency range, for example, vocalization or speaking during wakefulness, or snoring during sleep. The ε component is an umbrella term for other noise sources present in the signal: coughs, deglutition, vibrations caused by body movements during sleep, or other unknown transient artifacts.

The SSP signal was recorded with a sampling frequency f_s of 1024 Hz with a low-pass filter DC-285 Hz. The sensor's front-end automatic gain control and filters were disabled to avoid unwanted alterations in the signal's dynamic range. The SSP sensor can suffer from technical limitations. For example, a faulty sealing changes the pressure in the cavity, affecting the amplitude range of the whole signal (figure 1).

When this happens, the vibrations in the trachea are not strong enough and get dispersed before exciting the sensor. In 9 cases, the sensor failed to record any valuable data, and we decided to introduce an exclusion criterion to detect faulty SSP recordings. We opted for the percentage of samples with amplitude smaller than 2% of the sensor range of $0.5A_{\text{pp}}$ (amplitude peak-to-peak) to measure the magnitude of the problem. The median percentage of low amplitude samples in the whole dataset is 50.8% with an IQR of 25.76%. We opted for an exclusion criteria set at 90%.

The SSP sensor can also suffer from saturation and clipping artifacts, likely an effect of sensor misplacement, body movements, or other unknown factors. Figure 1 illustrates an example of spikes observed before the sealing failure. This class of artifacts happened sparsely in the dataset (<0.01% of the samples), thus we did not exclude any SSP recording due to excessive saturation. Instead, we tried to detect local spikes as artifacts and excluded only those segments of the signal from our analysis.

2.3. Methodology overview

Figure 2 provides an overview of the proposed method and the motivation behind the computational steps performed. The overarching goal is to get an accurate estimation of the heart rate at each time step n, $\hat{HR}(n)$, so

Figure 2. High level representation of the proposed methodology. Each color represent a sub-section of section 2, except for the detection of signal artifacts, treated extensively in appendix B. The symbol *n* represents the *n*th time step in the computation.
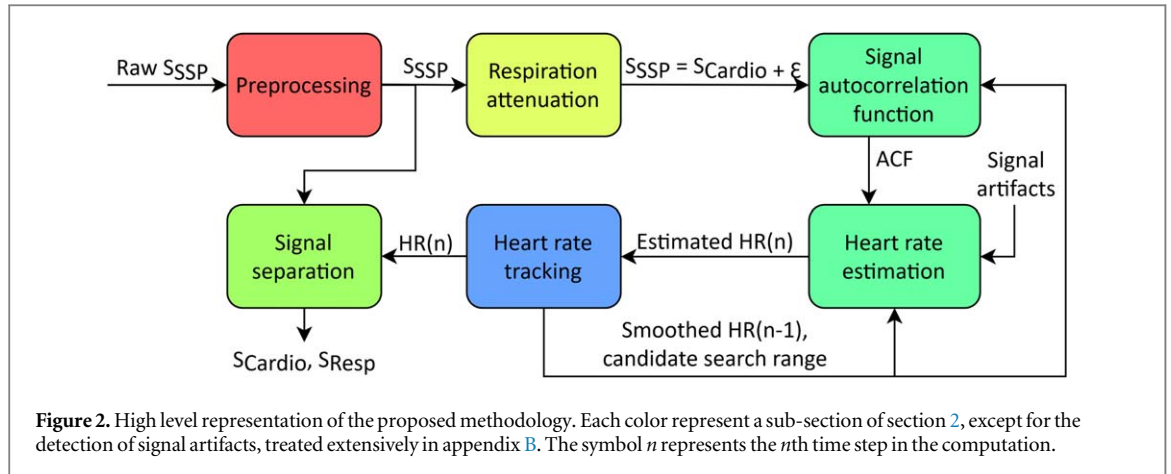
that we can tune a filter to separate the respiratory and cardiac signals S_{Resp} and S_{Cardio}. For all the following processing steps depicted in figure 2 we opted for a time window of 10 s with a 2.5 s step.

We must note that our algorithm includes multiple parameters that concur to its robustness, and that we selected according to technical factors and known physiological ranges during sleep, or that we optimized to improve the performance in our specific population. Nevertheless, none of the parameters was tuned upon known characteristics of the participant, making them entirely participant-agnostic. For reproducibility purposes, all the parameters are listed in appendix C.

Initially, we preprocess the raw SSP signal to remove the S_{Audio} component and powerline interferences that are not related to respiratory and cardiac information using a 4th order Butterworth low-pass filter at 32 Hz. Then the signal is re-sampled from $f_s = 1024$ Hz to $f_s = 256$ Hz using a FIR polyphase filter. This preprocessing optimally removes the S_{Audio} noise component, but ε will still be present in different forms.
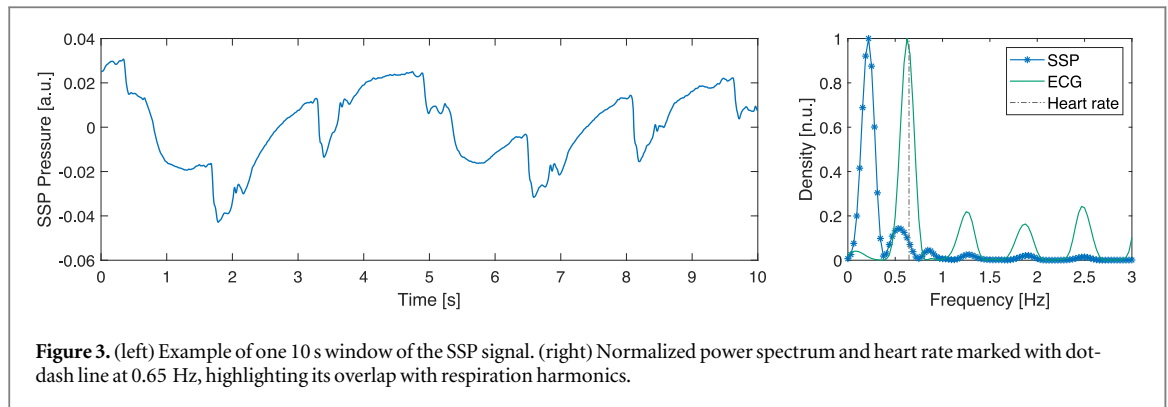
After preprocessing, we would ideally identify a precise frequency separating respiration and heart rate and design a high-pass filter to remove the respiratory signal. However, the process is not trivial because the heart rate may overlap with the higher harmonics of the respiratory rate (RespR), or because broad noise reduces the spectral contrast necessary to distinguish harmonics. In these situations, a filter with the wrong cutoff frequency will either have S_{Resp} leaks or lead to the removal of the main component of S_{Cardio}. Therefore we opted for a different solution that tries to boost the signal in the heart rate frequency range and attenuate the respiration as much as possible.

We now have a signal where S_{cardio} is present along with minimal residuals from S_{Resp}, that we will now consider part of noise ε. We expect that the fundamental frequency $F0$ of this signal will coincide with the heart rate, such that $F0 \approx HR$. However, residual noise may have a magnitude comparable with the cardiac oscillations, and maximum likelihood methods that use discrete Fourier transform (DFT) representations will not work reliably. We chose a time-domain method that identifies $F0$ from its autocorrelation function (ACF), which has already been demonstrated to be more robust to noise in challenging situations, such as speech analysis (Strömbergsson 2016).

The downside of the ACF representation compared to DFT is that there is not a single spectral peak describing a signal with a certain frequency $F0$, but multiple peaks at *lags* (indicative of the signal period) that correspond to $f_s/F0$, its multiples $f_s/(k*F0)$, and dividers $f_s/(F0/k)$ with $k = \mathbb{N} > 1$, even if the signal is a single sinusoid. This ambiguous representation in the ACF implies that we need a plausible guess of the range of lags around the expected heart rate and then update the search space over time as the heart rate fluctuates. Starting with a broad range covering observed physiological frequencies ([0.6: 1.8] Hz) we refine the boundaries according to heuristics and assumptions about heart rate dynamics.

We further divide each time window to get multiple sub-estimates of HR and to minimize the impact of short transient artifacts, assuming that sub-estimates will distribute around its true value. Then we select sub-estimates considering the distribution's skewness as a potential indicator of outliers. We also try to detect evident artifacts in the S_{SSP} signal to exclude them *a priori*. The artifact detection method is not our main contribution, so we will only describe it briefly. For a detailed description please refer to appendix B.

We then design a filter tuned on the estimated heart rate that separates S_{Cardio} and S_{Resp} from the preprocessed S_{SSP} signal. For each estimate, we create a band-pass filter from the HR up to its 3rd harmonic and apply it as a convolution in the frequency domain to extract S_{Cardio}. We derive S_{Resp} directly in the frequency domain by subtracting the now known cardiac components from the S_{SSP} spectrum.

Figure 3. (left) Example of one 10 s window of the SSP signal. (right) Normalized power spectrum and heart rate marked with dot-dash line at 0.65 Hz, highlighting its overlap with respiration harmonics.

2.4. Respiration signal attenuation

The first step of our method attenuates most of the S_{Resp} frequencies so that S_{Cardio} becomes the predominant signal in which we can identify the heart rate.

Let us consider the signal in figure 3 as a common example for our separation process. Often, S_{Resp} and S_{Cardio} are not perfectly separable, so we need to design a filter that considers the respiration rate RespR, but also other spectral characteristics of the S_{SSP} signal. The identification of RespR is relatively straightforward using DFT coefficients since S_{Resp} is predominant in the power spectrum and with a high signal-to-noise ratio (SNR). Nevertheless, we must account for the edge case represented by breathing disturbances, in which the estimated RespR is absent or much lower than the real respiration rate, and how that influences our filter design.

In our example, we extracted the heart rate from the ECG signal, observing a frequency of 0.65 Hz. This frequency is not clearly detectable in the S_{SSP} power spectrum, with the closest peaks around 0.53 Hz and 0.875 Hz. The first peak is likely the average between the heart rate and the respiration rate (0.21 Hz) 2nd harmonic. In this and many other cases, the S_{SSP} signal will not guarantee reliable spectral features for a high-pass filter at the optimal edge of respiration and cardiac oscillations.
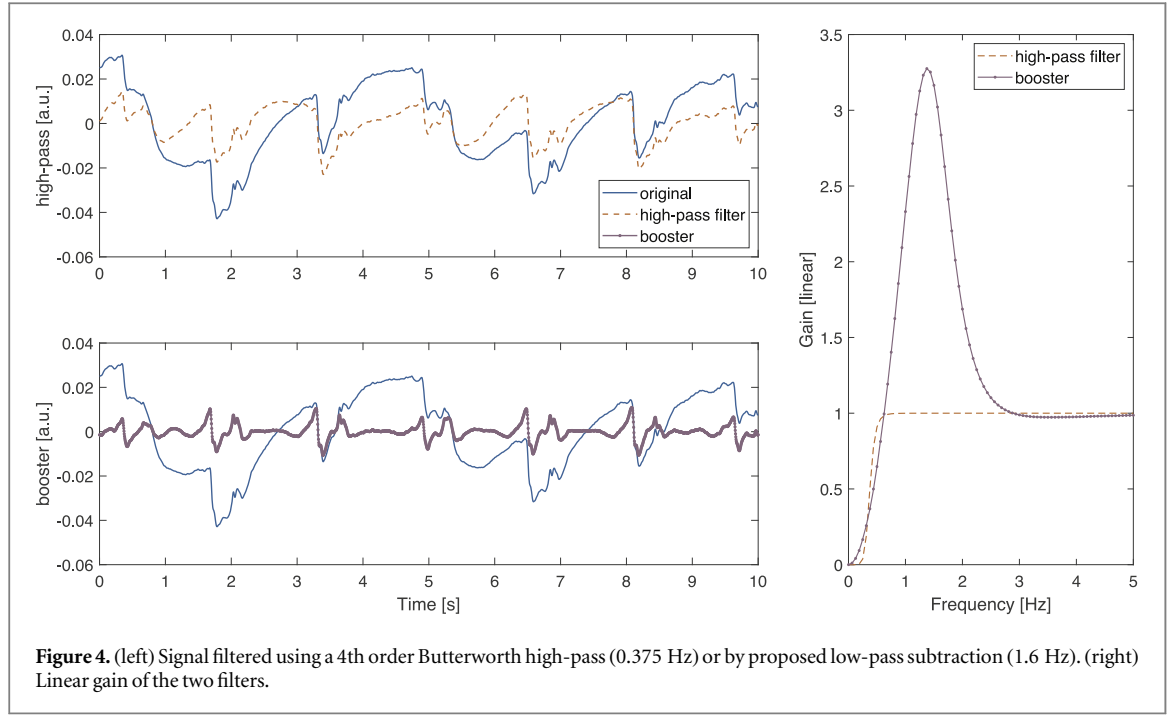
We then opted for a different approach. Instead of a finely tuned high-pass filter, we use a low-pass filter and subtract its output from the original signal. In this way, lower frequencies (where low-pass gain $\text{Gain}_{\text{LP}} = 1$) disappear entirely while frequencies around the low-pass cutoff are boosted ($\text{Gain}_{(1-\text{LP})} > 1$). While a boosting gain would be undesirable in filtering operations, it helps us detect frequencies that resonate together. This method does not guarantee the absence of noise under 1 Hz or respiration harmonics, but it gives us confidence that S_{cardio} frequencies will stand out for the heart rate detection step and it has the advantage of working well also with noisy spectra.

To determine the low-pass filter's cutoff frequency, we search for the first valley point $f_{\min}$ in the range [0.40 : 0.95] Hz. If none exists, we use 0.66 Hz as a fallback candidate (or the second harmonic of a maximum RespR around 20 breaths-per-minute). This frequency is assigned to a new range [$f_{\min}$: 2.4] Hz, where we detect all local maxima of the power spectrum. A pairwise comparison of these maxima flags those with potential higher harmonics. If a harmonic exists, we use its frequency as the low-pass cutoff, as it may be related to the heart rate, and they will resonate together. Other noise sources spread across a wide band of frequencies and should not have definite harmonics. If we cannot find harmonics, we set a fallback value of 1.6 Hz, boosting frequencies from 0.62 to 2.8 Hz (considering a 4th order Butterworth filter).

In figure 4, we see our example signal filtered using a 4th order Butterworth high-pass (0.375 Hz) or by low-pass subtraction (1.6 Hz). The first has some of the frequency content of respiration leaking into the signal. Conversely, our filter obtains unit gain near the real heart rate and a boost peak at its 2nd harmonic. We have chosen a 4th order Butterworth filter taking in consideration the desired boosting effect, its bandwidth and how much it attenuated frequencies in the respiratory band in our specific population and with this specific SSP sensor, and tested the performance of the filter in the whole estimation pipeline, in comparison with other filter typologies, both FIR and IIR.

2.5. Heart rate estimation

We proceed to estimate HR from the S_{Cardio} approximating signal. Assuming its fundamental frequency $F0 \approx HR$, we can reduce the heart rate estimation problem to the identification of the signal's main oscillatory component. In a noisy scenario, DFT-based techniques are influenced by residual noise, so time-domain methods such as the ACF are preferred. They were demonstrated in earlier works on speech analysis (Strömbergsson 2016) or digital transmissions (Chaudhari *et al* 2009) to be more robust to noise and have looser constraints on the signal's stationarity. In its general formulation the ACF is:

Figure 4. (left) Signal filtered using a 4th order Butterworth high-pass (0.375 Hz) or by proposed low-pass subtraction (1.6 Hz). (right) Linear gain of the two filters.

$$r'_x(\tau) = \frac{1}{r_x(0)} \sum_{k=1}^{M} x(k) * x(k+\tau),$$
(2)

$$r_x(0) = \sum_{k=1}^{M} x(k) * x(k),$$
(3)

where x is our signal of interest, and τ represents the time lag between samples over a window of M samples. The r'_x formulation is already normalized compared to the original $r_x(\tau)$ ACF coefficients, dividing them by the global maxima $r_x(0)$. In a noise-free situation, the autocorrelation function (ACF) decreases to zero as τ increases and has local maxima at $\tau = f_s/F0$, its multiples $\tau = f_s/(k*F0)$, and dividers $\tau = f_s/(F0/k)$ with $k = \mathbb{N} > 1$. This representation is the principal distinction between DFT where the fundamental frequency and its harmonics are distinct (given a frequency resolution high enough), and ACF where they conflate. Consequently, we can only estimate the correct frequency by searching the local maximum in a known range of interest $[f_{bottom} : f_{top}]$ in Hz, to be determined. The number of samples M is dependent on the largest period of oscillation (and consequently minimum frequency) that we want to observe correctly as a local maximum. In our case the minimum samples will be ideally $M > 2 * f_s/F0$. Since the frequency to lag conversion $lag(f) = \lfloor f_s/f \rfloor$ introduces a discretization error, we refine the lags' search range by selecting those that would be closest to the original boundaries when the inverse function $f(lag) = lag(f)^{-1}$ is applied.

Additional steps are needed to improve the likelihood that the $F0$ detected in the ACF is truly representative of HR because residual noise may influence the shape of ACF at certain lags, masking the peak of $F0$. Our solution is to subdivide our 10 s analysis window into smaller frames and shift the ACF function, to obtain multiple sub-estimates $\hat{F}0(i)$. If the heart rate remains relatively stable and noise is limited, we would expect that $\hat{F}0$ distributes around the real $F0$ as:

$$\hat{HR} \sim \mathcal{N}\left(\frac{1}{N}\sum_{i=0}^{N} \hat{F}0(i), \sigma^2\right) + \gamma_1.$$
(4)

We already account for the fact that the distribution will have a variance σ^2 and a non-zero skewness γ_1 caused by noisier frames or short heart rate spikes. We will not treat the problem of skewed distributions analytically, but each step of our algorithm applies some heuristics trying to keep $\hat{HR}$ close to a non-skewed normal distribution.

We identified $N = 20$ as the optimal number of frames using a parameter grid-search (range $N = [10 : 30]$, step size 2) on the entire dataset (the same applies to other hyperparameters introduced here). The size M of each frame and subsequently the step size between frames (step $= (L - M)/(N - 1)$ with $L = f_s^* 10$) is dependent on the lower boundary of the search range as $M = 2.0 * lag(f_{bottom})$. If the HR is low, the ACF needs more samples to characterize it, but the corresponding peak in the ACF will change slowly over time. Instead, higher frequencies need fewer samples, but frames will have a a higher overlap to better observe rapid variations.

Sometimes the search range may be misplaced, with local maxima falling on sub-harmonics $1/2 * F0$. Then, for each candidate $\hat{F}0(i)$, we extend the search range up to $2 * \hat{F}0(i)$ (or the first zero-crossing of ACF) to verify

the presence of a larger local maximum. If a new maximum exists and it exceeds $2 * \mathrm{ACF}(\hat{F}0(i))$, it becomes the new sub-estimate. Both the M multiplier and the amplitude threshold for harmonics were selected after a parametric grid-search (range [1.5: 2.5] step 0.1 for both the parameters). Nevertheless, there is a small but non-negligible chance that the extended range and undetected noise in the specific time frame introduces some outlier estimates. If these outliers cause a shift in the $\hat{F}0$ skewness we remove them using the following criteria based on the quantiles of $\hat{F}0$:

$$
\text{outliers} = \begin{cases} \hat{F}0(i) < Q1 - 1.5 * IQR, & \text{if } \gamma_1 < 0 \\ \hat{F}0(i) > Q3 + 1.5 * IQR, & \text{if } \gamma_1 > 0 \end{cases}. \tag{5}
$$

With $Q1$ and $Q3$ being the specific quartiles and IQR the interquartile range. We identified a threshold in the skewness value of ± 1.5 as the best performing in our dataset, although it does not remove all unbalanced estimates.

2.6. Heart rate estimate rejection

In some circumstances, the method presented fails to estimate the heart rate, and some corrective techniques are needed to manage unreliable segments of the signal.

In segments where the sensor disconnects from the PSG recorder the value of ACF is $r_x(0) \simeq 0$, the autocorrelation function does not cross zero or it does not exhibit any local maxima. In all three cases, we skip the estimation process.

In other situations, short, large transient artifact negatively affects the methods used to estimate the heart rate, so we opted for two levels of information rejection when they are present. We detected artifacts using symbolic aggregate approximation (SAX) technique (Keogh *et al* 2005, Senin *et al* 2014). SAX transforms a time series into a sequence of characters using quantized levels, then artifacts (here called discords) are selected as sub-sequences that are maximally different from all the others. A detailed description is available in appendix B. For each analysis window, if artifacts account for more than 12.5% of its samples, the entire window is rejected. The exclusion threshold is a compromise between performance loss caused by artifacts, and coverage. If the artifact samples are few but scattered across frames, we reject only the sub-estimates that contain those samples and calculate the HR averaging the others. If these frames are more than a threshold arbitrarily set at $3/4 * N$, we reject the entire window. For both artifacts and disconnected sensors, we store the estimate as invalid (NaN, not-a-number) and reset the search range for the next window at its default value.

2.7. Search range tracking

The selection of the proper search range is crucial for our method. If the search range is too wide, ACF noise will lead to bi-modal distributions of $\hat{F}0(i)$ and erroneous estimates. Conversely, a too narrow range may not respond well to sudden variations in the heart rate caused by several phenomena during sleep, such as apneas and arousals (Azarbarzin *et al* 2014, 2021). We then distinguish two operational phases: an initial *bootstrap* phase, where we do not know the heart rate nor can make proper assumptions, and a *tracking* phase, in which we aim to keep the search range in a soft spot centered around the average estimates.

The bootstrap phase comprises the first ten estimates ($\hat{HR}_{\text{booth}}$) with a default range of [0.6: 1.8] Hz, identified by looking at the heart rate distribution extracted from the ECG in the same initial interval. After the bootstrap phase we define an initial floor estimate f_{bottom} equal to $\text{median}(\hat{HR}_{\text{booth}}) - 0.5 * IQR(\hat{HR}_{\text{booth}})$. This early approximation is a reasonable guess of the bottom range of heart rate, which we will update with subsequent estimates. In the tracking phase, we reduce the f_{bottom} by a small margin of 0.05 Hz, clipping at a minimum value of 0.6 Hz since frequencies lower than that are unlikely in our sample population. If the bootstrap estimates are too high, the margin allows for corrections of f_{bottom}. The upper search limit f_{top} is instead dependent on a smoothed average of the last 10 HR estimates, as:

$$
\hat{HR}_{\text{smooth}}(n) = \alpha \hat{HR}(n) + (1 - \alpha) \frac{1}{10} \sum_{i=n-10}^{n-1} \hat{HR}(i), \tag{6}
$$

$$
f_{\text{top}}(n + 1) = \hat{HR}_{\text{smooth}}(n) + 0.25 \text{ Hz}. \tag{7}
$$

With n the nth estimate of the heart rate, α a smoothing factor equal to 0.1. We chose both α and the 0.25 Hz margin assuming that generally HR remains stable over time, but it can also increase suddenly and then return to the average value. This two factors offer a compromise between having enough margin to observe spikes while minimizing the risk of the tracker deadlocks in higher search ranges. If the estimate $\hat{HR}(n)$ is invalid we keep the last smoothed value. We opted to focus the HR tracking algorithm on the control f_{bottom} as it links to how many samples we use to calculate the ACF in the sub-estimates. The value of f_{bottom} is moved up or down to track slower changes in the average heart rate according to these rules:

$$\hat{HR}_{\text{smooth}} > f_{\text{bottom}} + 0.64 * (f_{\text{top}} - f_{\text{bottom}}) \rightarrow f_{\text{bottom}} + 0.01 \text{ Hz}, \tag{8}$$

$$\hat{HR}_{\text{smooth}} < f_{\text{bottom}} + 0.36 * (f_{\text{top}} - f_{\text{bottom}}) \rightarrow f_{\text{bottom}} - 0.01 \text{ Hz}. \tag{9}$$

We chose the two thresholds assuming that $\hat{HR}_{\text{smooth}}$ should stay close to the center of the search range. A shift towards f_{bottom} or f_{top} invalidates this assumption and updates the search range. We selected both thresholds empirically, optimizing for the minimum variance of the estimation error.

Furthermore, participants are most likely awake in the initial phase of the night (before sleep onset), and in this phase, heart rate remains relatively constant. After the person falls asleep, the heart rate slightly decelerates around 0.05 Hz–3 bpm (Shinar *et al* 2006), then decreases further during the first half of the night. We introduced a coarse correction mechanism trying to follow this deceleration. After the bootstrap phase, for each window, we decrease f_{bottom} by $1e - 5$ Hz, roughly equivalent to a reduction of 1 bpm every 10 min of recording.

2.8. Signal separation

After heart rate estimation, we separate respiratory and cardiac signals to retain their informative content. The separation happens on the original preprocessed signal before attenuating the respiration.

We opted for a filtering method that employs the short-time Fourier transform (STFT) of the S_{SSP} signal. Each time segment of the STFT corresponds to each window in which we estimated the heart rate. The HR estimation tunes a filter applied to every segment to obtain two distinct representations, one for S_{Cardio} and one for S_{Resp}. The reconstruction of neighboring windows in the STFT domain will have a smooth change in the cardiac frequency, similarly to time-variant adaptive filters.

At first, we calculate a frequency representation of the signal using the STFT as $X_{\text{SSP}} = \text{STFT}\{S_{\text{ssp}}\}$. Each time segment of the STFT is defined with the same length and overlap (10 s, 75%) used so far for the analysis. Then, for each estimated $\hat{HR}(n)$, with n the iterator of all estimates in time, we design a FIR bandpass filter with 2048 taps tuned from $\hat{HR}(n)$ to the 3rd harmonic and estimate its response $H_{\text{filter}}(n)$ in the frequency domain. If $\hat{HR}(n)$ was discarded we use the last valid value of $\hat{HR}(0: n)$. A single large bandpass filter guarantees *gain* $= 1$ in the S_{cardio} bandwidth with a better separation result. If other non-Gaussian noise sources (e.g. very high harmonics of S_{resp}) are present in-between the harmonics, they will leak into the final signal.

The signal S_{cardio} is then filtered using the next formulation:

$$X_{\text{cardio}}(n) = X_{\text{SSP}}(n) * |H_{\text{filter}}(n)|, \tag{10}$$

$$X_{\text{cardio}}(n) = |X_{\text{cardio}}(n)| * e^{(i*\angle X_{\text{cardio}}(n) + i*\theta(n-1))}. \tag{11}$$

Here we opted for a FIR filter with a high number of taps to obtain the highest resolution possible in $H_{\text{filter}}(n)$ and to multiply it directly with $X_{\text{SSP}}(n)$. With $\theta(n)$ in equation (11) accounting for small phase shifts that exist between the phase of $H_{\text{filter}}(n - 1)$ and $H_{\text{filter}}(n)$ at the frequency of $\hat{HR}(n)$. Now we consider $X_{\text{cardio}}(n)$ as additive *noise* in X_{SSP} and therefore obtain S_{resp} as:

$$X_{\text{resp}}(n) = X_{\text{SSP}}(n) * (1 - |X_{\text{cardio}}(n)|/|X_{\text{SSP}}(n)|). \tag{12}$$
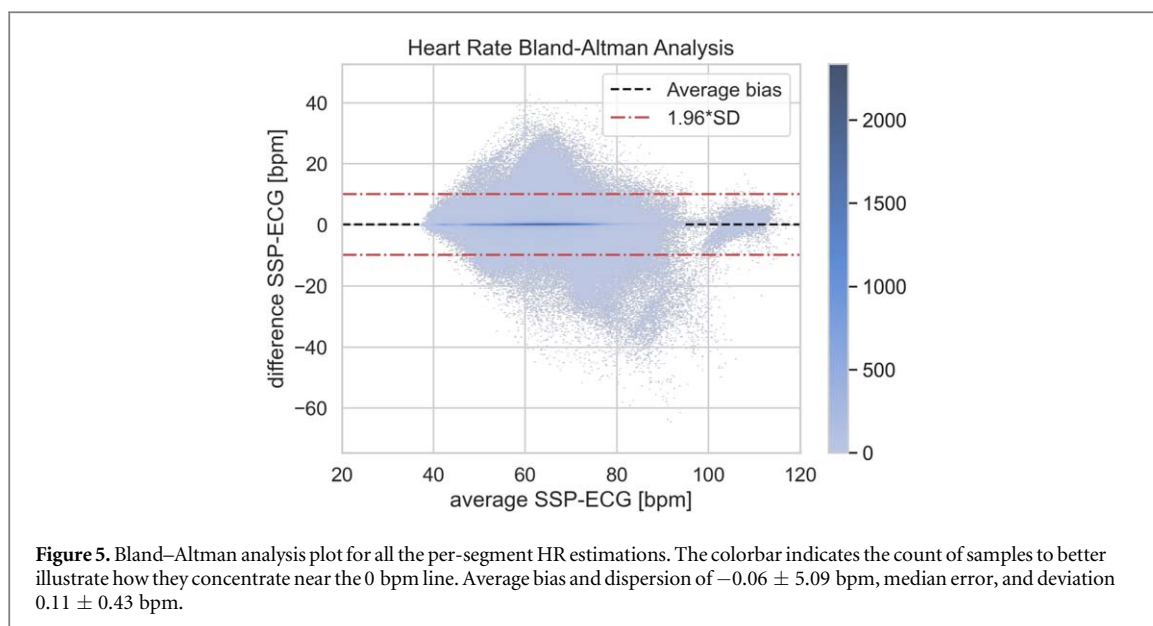
Additionally, X_{resp} is low-pass filtered (FIR 1024 taps) using the 3rd harmonic of the respiratory rate $RR(n)$ as the cutoff point to remove high-frequency noise higher than that. We finally reconstruct the signals from the frequency domain to the time domain using the inverse STFT and store them.

2.9. Performance evaluation

To evaluate how well our method can estimate the heart rate from the SSP sensor, we extracted the heart rate from the synchronous ECG as a ground truth. QRS complexes are detected and localized with the *ecg_peaks* method available in the NeuroKit2 toolbox (Makowski *et al* 2021), checked for detection quality using the *ecg_quality* method (based on Zhao and Zhang 2018), and then visually inspected to correct misclassified ECG peaks with the R-DECO software (Moeyersons *et al* 2019). The HR_{ecg} is the average of the interbeat intervals on the same 10 s time windows used by our proposed method. If the HR is outside a plausible physiological range (set broadly at [24: 220] bpm) or the signal quality is deemed too low, we discarded the window.

We pooled the estimations on the entire population to give a complete comparison between HR_{ecg} and HR_{ssp} using Bland–Altman agreement analysis. Besides the average estimation error (bias) and its deviation in beats-per-minute (bpm), we also considered the percentage of estimates above 10 bpm and below 5 bpm, and the coverage, or the amount of valid HR_{ssp} estimates over the valid HR_{ecg} measurements. Since some estimation errors may be large but otherwise sparse in the signal, we also considered their median and median absolute deviation (MAD) since these metrics are more robust to outliers. The same metrics were calculated separately per recording to detect problematic SSP signals.

We calculated the skewness and kurtosis of the sub-estimates for each estimation window and used a Shapiro–Wilk test to verify if they are normally distributed. We compared the effect of having absolute skewness

Figure 5. Bland–Altman analysis plot for all the per-segment HR estimations. The colorbar indicates the count of samples to better illustrate how they concentrate near the 0 bpm line. Average bias and dispersion of -0.06 ± 5.09 bpm, median error, and deviation 0.11 ± 0.43 bpm.

greater than one, excess absolute kurtosis greater than 0.5, or non-normal distributions using the one-sided Wilcoxon Rank-Sum test on the absolute estimation error.

To quantify how respiratory events dynamics impact our performance, we employed the scored PSG annotations to separate the estimates obtained during normal breathing and during sleep disordered breathing events for each recording. Then, we subdivided breathing events along with their onset and termination, where the SSP signal may be turbulent, and during the event where it is relatively stable. We compared the results using the Wilcoxon Rank-Sum test against normal breathing (null hypothesis: the median difference is 0) and between event classes to evaluate if they have a larger impact on error. To compensate for inter-recording variability, the effect is calculated as the difference between the median estimate error during events and during normal respiration. Lastly, we assessed linear effects of age, body-mass index (BMI), and AHI on our performance indexes using Kendall τ correlation coefficient to reduct the risk of bias caused by model residuals. Potential effects of sex were assessed with two-sided Wilcoxon Rank-Sum test.

3. Results

3.1. Heart rate estimation

Figure 5 shows the Bland–Altman agreement analysis on the entire population. The error bias is -0.06 bpm with a dispersion (standard deviation SD) of 5.09 bpm. Most of the measured bias is near 0 bpm, with 85.49% of errors less than 5 bpm and only 6.36% over 10 bpm. Specifically, the median error is 0.11 bpm with a median absolute deviation of 0.43 bpm. With respect to ECG coverage, the average per-recording is 94.4%, with the first and third quartiles at 92.5% and 97.1%, respectively. 86 out of 100 recordings have a coverage higher than 90%. The average error of all the under-estimated segments is -2.84 bpm, slightly larger than the over-estimated segments, with an average of 2.03 bpm. In the same way, the maximum under-estimation error exceeds over-estimation with a value of -64.65 bpm and 42.70 bpm, respectively.

Figure 6 illustrates the Bland–Altman analysis for each participant, with a median bias of -0.12 bpm and a median deviation of 3.42 bpm. 94% of the recordings have an average bias smaller than 2 bpm, while few recordings have a greater contribution to the error seen in figure 5, with high bias, large error dispersion, or both. Our method only severely under-performs in four recordings, with more than 50% of estimates having an error larger than 5 bpm. In one of these recordings the participants presented Cheyne-Stokes breathing patterns for most of the night.

Of the 100 recordings analyzed, 8 fit the exclusion criteria presented in section 2.1 and 2 are excluded because of persistent Cheyne-Stokes breathing patterns. After excluding these recordings, the bias becomes -0.26 bpm with a standard deviation of 3.88 bpm, with the median error reduced to 0.09 ± 0.36 bpm. The percentage of errors under 5 bpm increases to 90.07% with only 3.24% above 10 bpm. All of the following results do not exclude any recording to give a perspective on the whole experimental population.
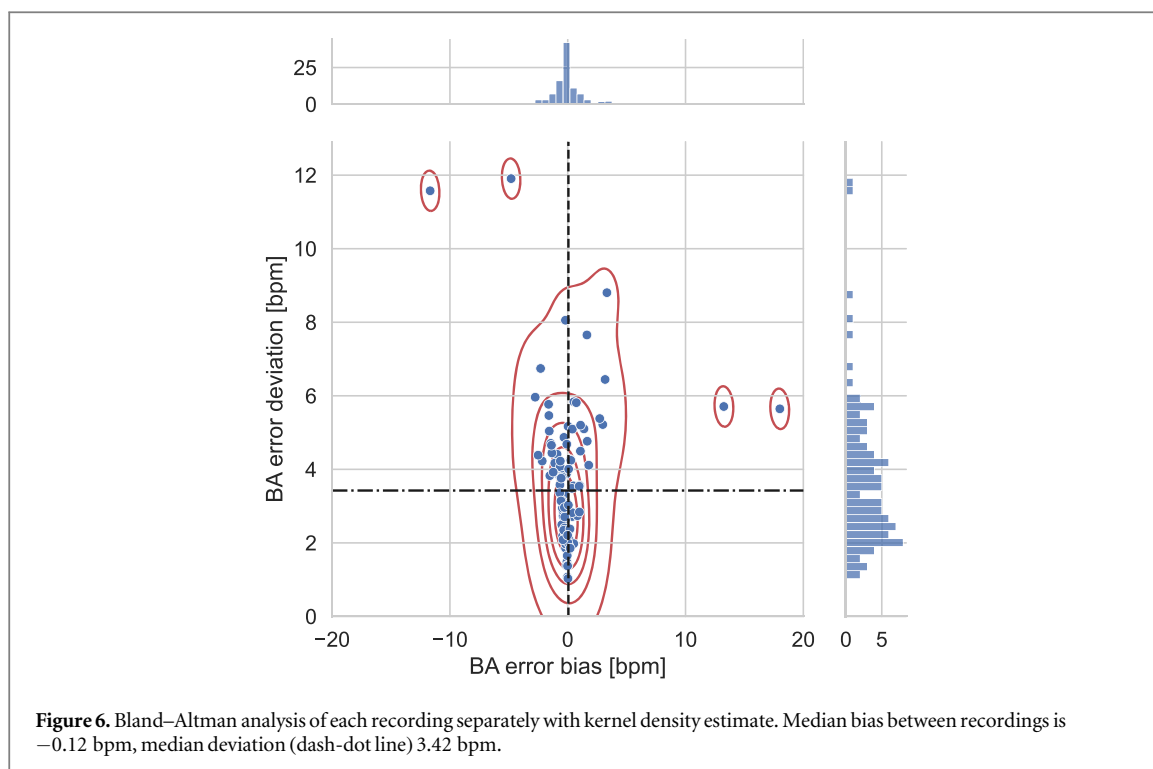
Figure 6. Bland–Altman analysis of each recording separately with kernel density estimate. Median bias between recordings is −0.12 bpm, median deviation (dash-dot line) 3.42 bpm.

Table 2. Effect of respiratory events on method performance as difference with average normal respiration error. Number of observations: 98.

Type	Median Δ error [bpm]	Range	IQR[a]	p
Event end	−0.19	−11.97 : 4.04	0.83	<1e-05
Event onset	0.03	−8.18 : 10.66	0.20	0.054
During the event	−0.04	−8.96 : 5.51	0.25	<5e-02

[a] Inter-Quartile Range.

3.2. Assumptions on normality of estimations

We assumed the sub-estimates of each window to be minimally skewed and normally distributed, particularly after our correction. Concerning sub-estimates' skewness, 84.93% of the windows considered have an absolute skewness lower than 1, with a median value and standard deviation of 0.08 ± 0.67. We contained distributions with absolute skewness above 1.5 using the method presented in section 2.5, but we observed that a minimum amount of estimates (1.14%) remains unaffected by the correction. When the absolute skewness is above 1, the absolute estimation error is significantly higher ($p < 1$e-04) with a median of 1.30 bpm versus 0.39 bpm.

The sub-estimates' kurtosis (Pearson's notation) is distributed around a median of 2.31 with an interquartile range from 1.88 to 2.99. It implies a slightly platykurtic tendency of estimates, which may be an outcome of the outliers rejection presented in section 2.5. The sub-estimates with excess kurtosis are 74.2% of the total, but we did not observe significant differences in the error.

Using the Shapiro–Wilk test with a rejection $\alpha = 0.05$, 42.35% of the valid windows reject the null hypothesis, and they cannot be assumed to be normally distributed. When the sub-estimates distribution is non-normal, the absolute error is significantly higher ($p < 1$e-04) with a median of 0.87 bpm versus 0.34 bpm.

3.3. Effect of respiratory events

We can observe in table 2 that respiratory events introduce an estimation error smaller 1 bpm on average. However, some recordings report non-negligible differences compared to normal respiration. The onset of an event contained the HR estimates from 5 before to 2.5 s after its start. Conversely, the end of an event is considered from 2.5 s before its termination to 5 s after. Out of 100 participants, 2 did not have respiratory events during the night, hence they are not considered here.

Although small, the median difference is significantly different from zero for event transitions and during the events. A paired comparison between the different portions of events suggests that our method performs significantly worse at the end and during the respiratory events than at their onset (table 2). Upon closer

Table 3. Effect of apneic events (obstructive or not) and hypopneas on absolute error as difference with average normal respiration error. Number of observations: 65.

Type	Apneas $\vert\Delta$ error$\vert$ [bpm]		Hypopneas $\vert\Delta$ error$\vert$ [bpm]		p
	Median	IQR[a]	Median	IQR	
Event end	0.27	1.21	0.46	1.14	0.662
Event onset	0.29	0.93	0.10	0.33	<1e-03
During the event	0.41	0.93	0.11	0.26	<1e-05

[a] Inter-Quartile Range.

inspection, we found that the largest positive and largest negative outliers come from subjects that can be excluded according to the criteria in section 2.1.

If we distinguish the absolute effect of apneic events, both obstructive and central, and obstructive hypopnea events (table 3) the effect of apneas is significantly larger at event onset and during the event, with an increased error of 0.24 bpm on average, but not at event end. It must be noted that out of 98 participants considered, 33 had only hypopnea events and no apneas, hence they are excluded from the comparison.

Although respiratory events lead to local errors in the estimation accuracy, we did not observe strong correlations between the AHI of each participant and our performance indexes. We converted the average and median bias to their absolute value, while other variables are unchanged. The correlation coefficients are: average bias (τ of 0.16, $p < 0.05$), error variance (0.17, $p < 0.05$), median bias (0.15, $p < 0.05$), median absolute deviation (0.25, $p < 0.001$), percentage of errors above 10 bpm (0.19, $p < 0.01$), and percentage of errors below 5 bpm (-0.22, $p < 0.01$).

3.4. Effects of demographic factors

The manifestation of OSA in patients may differ significantly with age, sex or gender, and BMI and consequently influence our method's performance. After the exclusion of faulty recordings, our dataset includes 40 females and 50 males. None of our performance indexes exhibited significant differences ($p < 0.05$) with respect to gender. We observed minor correlations between performance and the age of the participant: average bias (τ of 0.18, $p < 0.05$), error variance (0.16, $p < 0.05$), median bias (0.18, $p < 0.05$), median absolute deviation (0.27, $p < 0.001$), percentage of errors above 10 bpm (0.21, $p < 0.01$), and percentage of errors below 5 bpm (-0.27, $p < 0.001$). We did not observe significant relationships between the BMI of participants and our performance indexes. Other details of the participants may influence the quality of the signal, but are currently not accounted, such as skin elasticity, and other anatomical (e.g. fat depositions in that area) or non-anatomical factors (e.g. mucus accumulations caused by chronic obstructive pulmonary disease (COPD), physiological status of the lungs, body position etc).
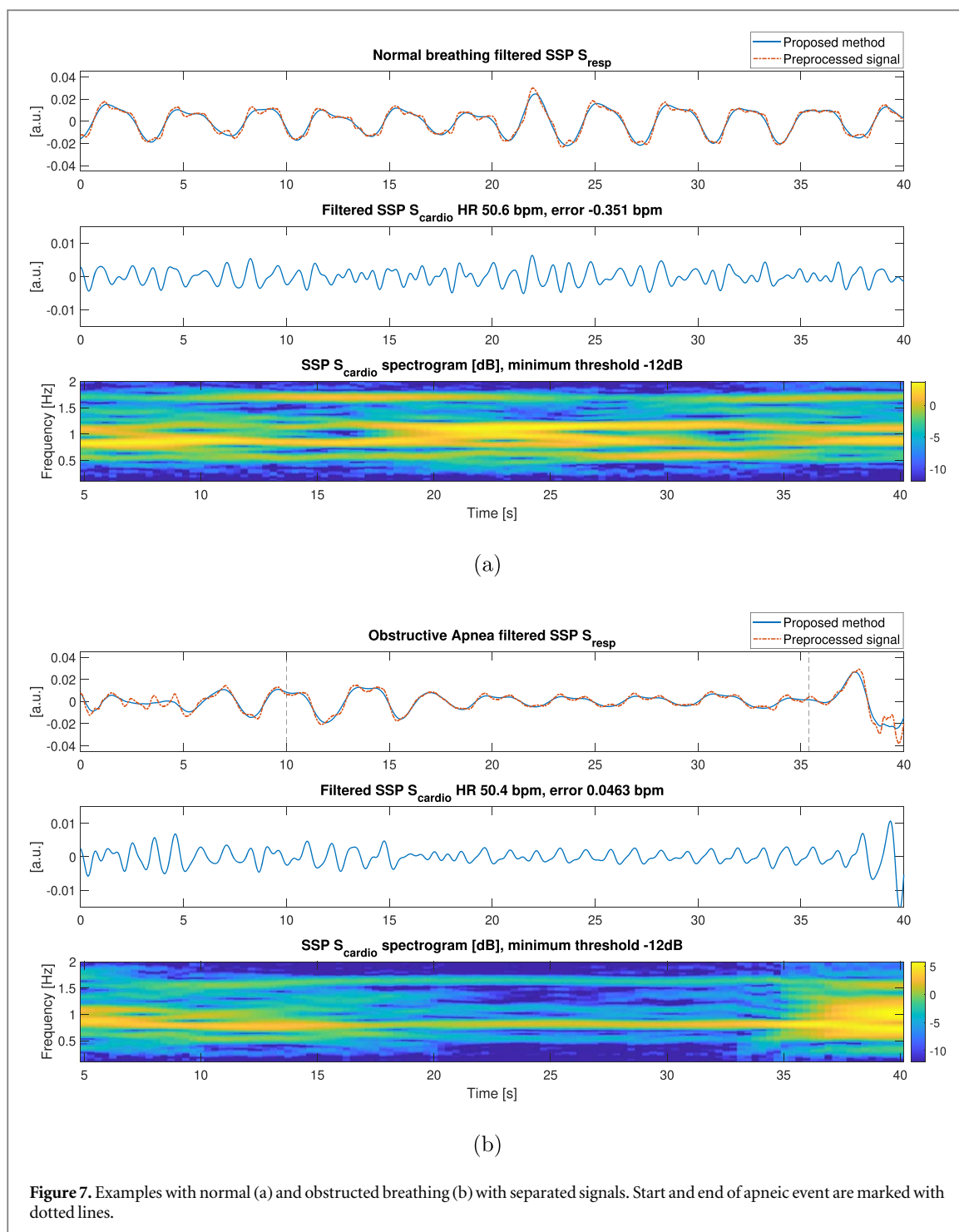
3.5. Qualitative improvement of respiratory effort signal

Figures 7(a), (b) and 8 illustrate examples chosen to highlight how the signal changes after separating S_{resp} and S_{cardio}. The aim of these examples is to show how a well tuned filter can lead to better visual evaluations of the signal.

Figure 7(a) shows that our method smoothens out the small cardiac oscillations in S_{resp}, but the quality of S_{cardio} is low despite the low estimation error. The spectrogram shows that all three harmonics of the HR (0.84 Hz) are present, but another strong signal at 1.12 Hz is the most likely cause of noise here.

In the obstructive apnea example in figure 7(b) our method filters smaller cardiac oscillations at the beginning and during the apnea, so that we can better understand if there is a pattern of increased respiratory effort caused by the obstruction. Also, the cardiac signal shows a better resemblance with cardiac pressure waves (Dillon and Hertzman 1941), qualitatively speaking. We can also see that respiratory overshooting at the end of the event may introduce noise in the signal spread across cardiac harmonics.

The central apnea example in figure 8 shows very interesting results. We can remove the cardiac oscillations from the central apnea, which results in a more accurate representation of the full cessation of breathing effort. Although the cardiac information during the event is fully retained in S_{cardio}, the first half of this example (from 0 to 20 s) presents the same issues observed in the first example, with a higher spectral noise between the fundamental frequency and the first harmonic, which disappears in the second half.

Figure 7. Examples with normal (a) and obstructed breathing (b) with separated signals. Start and end of apneic event are marked with dotted lines.

4. Discussion

We describe a method to extract the cardiac signal present in SSP recordings and separate it from the respiratory signal, retaining the information contained in both. The goal was twofold: first, we aimed to improve the filtering of cardiac oscillations so that respiratory events are better characterized in the resulting signal; second, we aimed to retain the potentially informative content of the cardiac signal. To achieve these goals we designed a processing system that estimates the heart rate fluctuations in the signal, and exploits this information to tune a bandpass filter. This filter is employed to separate the two signals of interest in the frequency domain. Heart rate was estimated using the combination of time-domain representation of the signal in the form of autocorrelation function (ACF), and a set of domain-driven heuristics to track and improve estimates during sleep.

We measured the quality of the method on the agreement between the SSP-estimated heart rate and the equivalent measured from a synchronized ECG signal. Generally, the heart rate estimation error with the proposed method is low, with an average bias of -0.06 bpm and a standard deviation of the error of 5.09 bpm
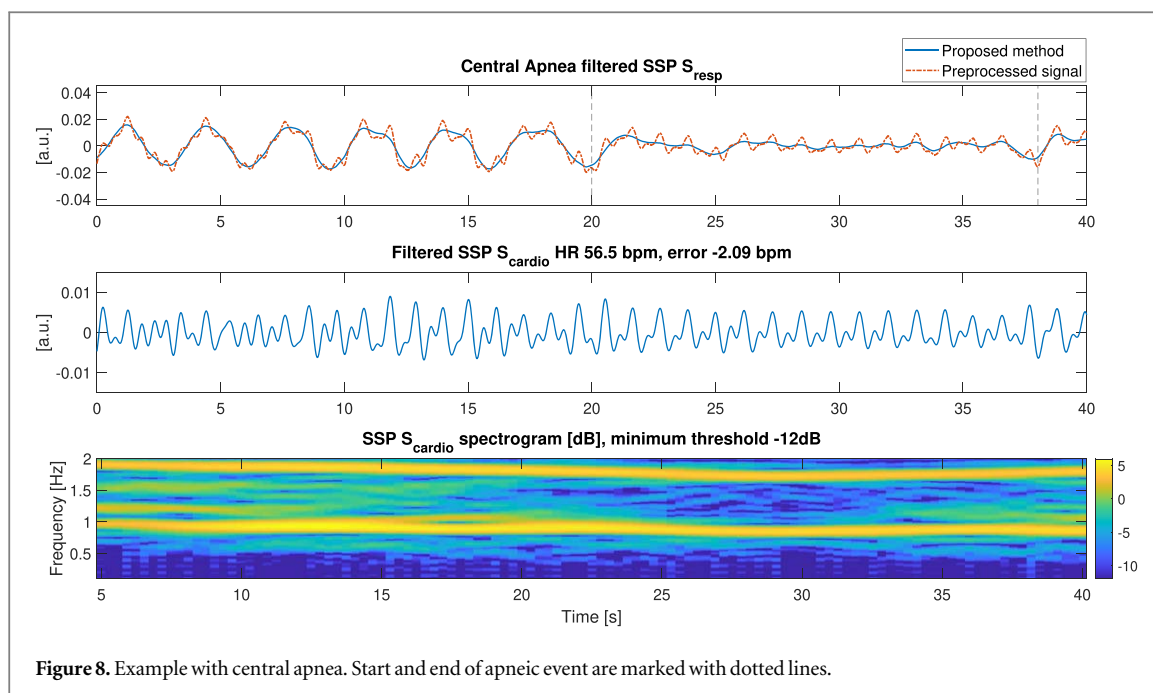
Figure 8. Example with central apnea. Start and end of apneic event are marked with dotted lines.

over 10 s, relatively constant over the entire range of measurements, and well contained, with more than 90% of the recordings having an average bias smaller than 2 bpm. The combination of artifacts detection and outlier removals allowed us to obtain a very high coverage of 94.4% with respect to ECG-derived heart rate.

The quality of the estimation, or at which point the error can be considered large or small, depends on the application of the heart rate signal, the sampling rate of the source, and the intrinsic characteristic of the source signal itself. The authors in Peláez-Coca *et al* (2022) have shown that signals like ours or photoplethysmography (PPG) signals do not have well defined fiducial points like the ECG does, and consequently these signals have some discrepancies with the ECG even at high sampling rates. Furthermore, for certain applications the difference between pulse rate variability (e.g. from the PPG) and heart rate variability (from ECG) may be non-negligible (Mejía-Mejía *et al* 2020). Since this is the first time, to the best of our knowledge, that cardiogenic oscillations are considered as a separate signal, there are no studies in literature examining their physiological significance and their usage in beat-by-beat analysis applications, such as pulse rate variability or the detection of abnormal beats. On the other hand, we believe that our level of accuracy may be sufficient in applications where the heart rate dynamics are smoothed over time, such as monitoring trends in heart rate or sleep staging (Bakker *et al* 2021).
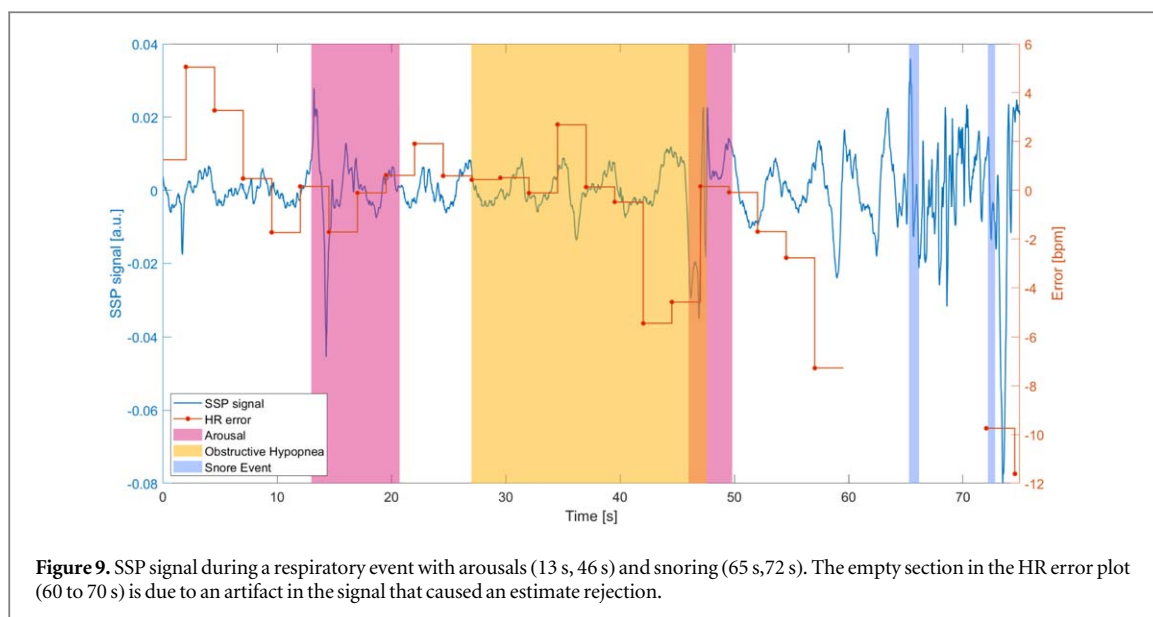
In a broader multi-modal scenario, we believe that the information carried by the SSP signal could be used to improve other signals. For example, the S_{Resp} respiratory signal may be used to reduce fluctuations visible in ECG and PPG signals to improve the detection of heart-beats. Additionally, the peaks in the S_{Cardio} signal could be used as candidates of pulse locations in majority voting fusion systems (after accounting for pulse transit time effects) (Rankawat and Dubey 2017).

We also believe that our proposal opens new research developments in respiratory analysis and that a better comprehension of respiratory effort measurements can lead to their widespread usage in clinical settings. Other than enriching the information available in PSG recordings, we foresee the usage of the SSP sensor in novel integrated home sleep apnea tests devices, with an improved characterization of respiratory effort and upper airways dynamics.

4.1. Factors influencing the accuracy of heart rate estimation

In 10 out of 100 recordings the results were unsatisfactory, due to a combination of high signal noise and sensor failures, large transients in the signal caused by turbulence associated with respiratory events, and technical shortcomings of our method.

The physical properties of the SSP sensor are one of the factors that influences the quality of the raw signal. Requiring an airtight capsule makes it more prone to disruptive effects caused by detachments or faulty sealing. While in some cases the signal loss is evident (figure 1), we also observed drift, baseline wander, or baseline jumps effects potentially caused by smaller leaks, and that are not correctly filtered by our preprocessing. In both cases the system would require additional sensors that measure the capsule status, body movements, or environmental effects, paired with corrective algorithms, such as adaptive filters, source separation techniques

Figure 9. SSP signal during a respiratory event with arousals (13 s, 46 s) and snoring (65 s, 72 s). The empty section in the HR error plot (60 to 70 s) is due to an artifact in the signal that caused an estimate rejection.

or gain correction methods at the sensor frontend. If a correction is not possible, it would be necessary to quantify the reliability of each heart rate estimate and flag or remove them if the quality is too low. Furthermore, analysing the residuals between the original SSP signal and the separated S_{Resp} and S_{cardio} signals after the estimates may help us understanding how well the information content of the signal was captured.

In addition to currently uncontrolled artifacts, multiple phenomena can drastically and rapidly alter the temporal evolution of the S_{SSP} signal. For example, the signal's dynamics can change, both with natural physiological mechanisms of sleep (e.g. heart rate changes with sleep stages and arousals) and pathological respiratory events (e.g. changes in breathing rate during apneas), or anatomical changes in the lungs, trachea and heart that may depend on the age of the participant (see section 3.4). This situation imposes a constraint on the signal processing techniques we can apply, as signal stationarity is quite limited in time. During the analysis, long time windows provide more reliable estimates of low frequencies, but average heart rate may change too rapidly. On the contrary, shorter windows would guarantee limited heart rate fluctuations, but the characterization of slow-varying S_{Resp} before its subtraction would be more challenging. Our selection of a 10 s time window is a compromise between these two contrasting situations. As a future development of our algorithm, we foresee it would be possible to detect apneic events on the S_{Resp} signal, and manipulate the time window as necessary. For example, windows with higher overlap after apneic events, or longer and less overlapped windows when the respiration and heart rate are changing slowly.

Secondarily, we observed how respiratory events and abnormal respiratory patterns such as Cheyne-Stokes breathing increased the estimation error. The error measured at the end of respiratory events may be caused by typical post-event phenomena, such as hypercapnia, arousals, coughs, or gasps that introduce large artifacts in the SSP signal and hide the effect of cardiac oscillations. Figure 9 shows an example of such events. An obstructive hypopnea event accompanies two arousals and snoring events that impact the signal quality, to the point that the resulting artifacts lead to the rejection of the estimates. Although both hypopneas and apneic events had a significant effect on the estimation error, we observed that hypopneas are less disruptive. We hypothesize that the reduction in the airflow without complete obstruction does not introduce the same extreme pressure swings visible during apneas.

Some of the assumptions we made in the design phase of the system are invalidated by the transient non-linearities of respiratory events and corrective measures are necessary. For example the time window should update dynamically if respiratory turbulence is detected, potentially leading to a better attenuation of the respiratory signal. Furthermore, the boosting filter may be the culprit of these errors because of its design. The boosting effect is relatively broad in frequency. On the one hand, a broadband boost helps us as we do not need to identify the cutoff frequency perfectly to catch the cardiac oscillations. On the other hand, the filter may also boost some noise, leading to uncertainties in heart rate detection later. The boosting filter needs also a phase-aware design, as the current implementation introduces undesired phase shifts in the range of boosted frequencies. Phase flips between harmonics are detrimental for ACF representation, leading to uncertainties in the heart rate estimation.

There is also space to improve the heart rate estimation algorithm to address some weak points of ACF representation. For example, using multiple representations of the signal and ensembling the results (Brüser *et al* 2013), employing estimates of sleep stages or respiratory events visible in the signal itself (Fonseca *et al* 2015).

While we did not observe significant differences in the performance with respect to sex and BMI, we observed small correlations with age and AHI. Given our specific population and the known complex interplay of age, gender and BMI in the severity of OSA, further tests on healthy elderly participants are required to understand the nature of these relationships. Currently we can only assume that the strength or nature of cardiogenic oscillations in the SSP signal may be partially influenced by the effects of aging heart and vasculature.

Another important consideration is that comorbid cardiac conditions (quite common in OSA) may produce a discrepancy between the electric activity of the heart and the manifestation of heart pulsations visible in arteries (Gil *et al* 2013, Sološenko *et al* 2017). Further research is necessary to discriminate how irregular pulses of the heart affect the SSP signal. The average night-time heart rate of our participants is in line with other OSA and disturbed sleep cohorts (Choi and Kim 2011, Huang *et al* 2018). Nevertheless, one participant maintained a HR >100 bpm and remained awake for most of the night. In this specific case, the correction for sleep-induced heart rate deceleration likely introduced an unwanted drift. As we already know that the heart rate does not decrease linearly for the entire night, a better mechanism could employ a more precise prior of the deceleration curve, sleep staging information, or data from other sensors (e.g. body movements) to know if the person is asleep or not.

Lastly, we observed how skewed distributions of $\hat{F}0$ lead to larger errors. Other than spurious elements in the signal, for example undetected artifacts, it may happen that our algorithm introduces artificial errors at upper harmonics. When they are not corrected, wrong sub-estimates introduce error creeps, both in the single estimate and by drifting the heart rate tracking. Better selection algorithms are required to properly compensate for skewed or bimodal distributions.

4.2. Considerations on harmonic reconstruction error

The fact that the proposed method is resilient to small estimation errors (as seen in figure 8) opens the way to some considerations on what is the largest $\hat{HR}$ error tolerable if we want to separate respiratory and cardiac signals. If we slightly overestimate HR, the current transition bandwidth of the S_{cardio} filter guarantees that, up to 5 bpm (or 0.1 Hz), the signal will stay over the -3 dB threshold. The amplitude of the signal peaks will not be usable as a feature, but it should be good enough to improve the estimation of HR through the detection of inter-beat intervals. On the contrary, $\hat{HR}$ underestimations have a larger operative margin, thanks to full bandpass bandwidth from $\hat{HR}$ to $3 * \hat{HR}$. If the real $HR < 2 * \hat{HR}$, the filtered S_{cardio} will contain at least the fundamental frequency and 2nd harmonic, as mistaken for the 2nd and 3rd harmonic. However, two limitations remain: every signal or noise with frequency $\hat{HR} < f < HR$ will leak into S_{cardio}. Similarly, if higher harmonics of S_{resp} are in present in that bandwidth, they will be filtered out. We will need more sophisticated filters to guarantee that signals' separation does not disrupt valuable information when we reconstruct S_{Cardio} and S_{Resp}.

4.3. Physiological interpretation of the phenomenon

We can make some hypotheses on the physiological mechanism driving the cardiac oscillations. While earlier studies connected cardiogenic oscillations to the vibration of the heart through the lungs (West and Hugh-Jones 1961), researchers later demonstrated that they derive from the pressure wave of the pulmonary artery near the trachea's bifurcation (Suarez-Sipmann *et al* 2012). If this is the case, they would be more visible when the airflow is low, such as at the end of the expiration or in the presence of central apneas, as illustrated in our examples. It must be noted that the experiments conducted by Suarez-Sipmann and colleagues were performed on pigs, which may present slight anatomical differences compared with human subjects. A second potential source in humans may be the aortic arch that runs left of the trachea, but we are unaware of experimental studies directly supporting this. At the same time, since oscillations seem dampened during normal breathing and amplified during apneas, they could also be related to blood pressure variations. Different studies (Alex *et al* 2017) showed that blood flow velocity and pressure both increase during apneas. In this case, the vibration of the carotid artery may be transmitted transversely through the still air in the trachea. It is also possible that the temporal shift between the two is so small that they both concur in the signal. Unfortunately, we cannot verify these hypotheses with the data available at the moment. Future experiments with the same pressure sensor used to detect the SSP signal or with a PPG sensor placed on the carotid artery could provide the means to determine how the two signals are related.

Concerning the physiological role of cardiogenic oscillations in sleep-disordered breathing conditions, some studies on rodents (Sullivan and Szewczak 1998, Dubsky *et al* 2018) showed a positive oxygen intake and lung air mixing as the result of increased cardiogenic activity during induced apneas or hibernation. We cannot be sure that indirect air mixing represents an active protective mechanism against lowering oxygen levels. Yet, we can hypothesize that strong cardiogenic activity may be a phenotypical trait of OSA. For example, cardiogenic activity may correlate to higher pressure in the pulmonary arteries. If this is the case, it may be a part of the bi-directional link between OSA and pulmonary hypertension, as proposed by different authors (Mesarwi and

Malhotra 2020, Sharma *et al* 2021). Similarly, the interaction between intrathoracic pressure and cardiogenic activity may represent another indicator of the physiological mechanisms that govern cardioventilatory coupling (Sin *et al* 2012). That would add insights to the techniques (such as the loop gain) currently employed to characterize patients with elevated cardiovascular risk (Edwards *et al* 2019).

From a clinical perspective, multiple controlled experiments are necessary to unravel the exact physiological mechanisms that manifest as cardiac oscillations in the SSP signal. A better comprehension of the phenomenon could then lead to improved algorithms to analyze this information-rich signal.

5. Conclusions

This paper presents a method to estimate heart rate and extract the correlated cardiac signal from a Suprasternal Notch pressure sensor during polysomnographic recordings of people with suspected sleep disorders. Performance of heart rate estimation is promising, and we observed definite qualitative visual improvements in the characterization of respiratory events with correct filtering of cardiac oscillations. Future developments will cover both technical improvements necessary to get more reliable estimates and a cleaner cardiac signal, but also to better understand the potential clinical applications of the system. If the information carried by the SSP signal correlates with other cardiac sensors like the PPG, it means that we could design new home sleep apnea tests systems that need fewer sensors but provide a level of detail comparable with obtrusive and expensive PSG systems at a fraction of the cost. Potential coupling mechanisms between the respiratory effort signal and cardiac oscillations may be complementary to all other signals as a descriptor of the complex OSA dynamics. These can also represent an almost direct probe of the activity of the pulmonary artery due to their cardiogenic nature, and could be used, for example, to possibly quantify the risk of pulmonary hypertension in OSA patients (Sharma *et al* 2021).

Acknowledgments

Appendix A. Cardiac oscillations under various respiratory conditions

Figure A1 shows the SSP signal under different conditions: normal respiration (ideally Eupnea), partial or fully obstructed breathing, and central apneas, where the airway is still partially open, but the activity of respiratory
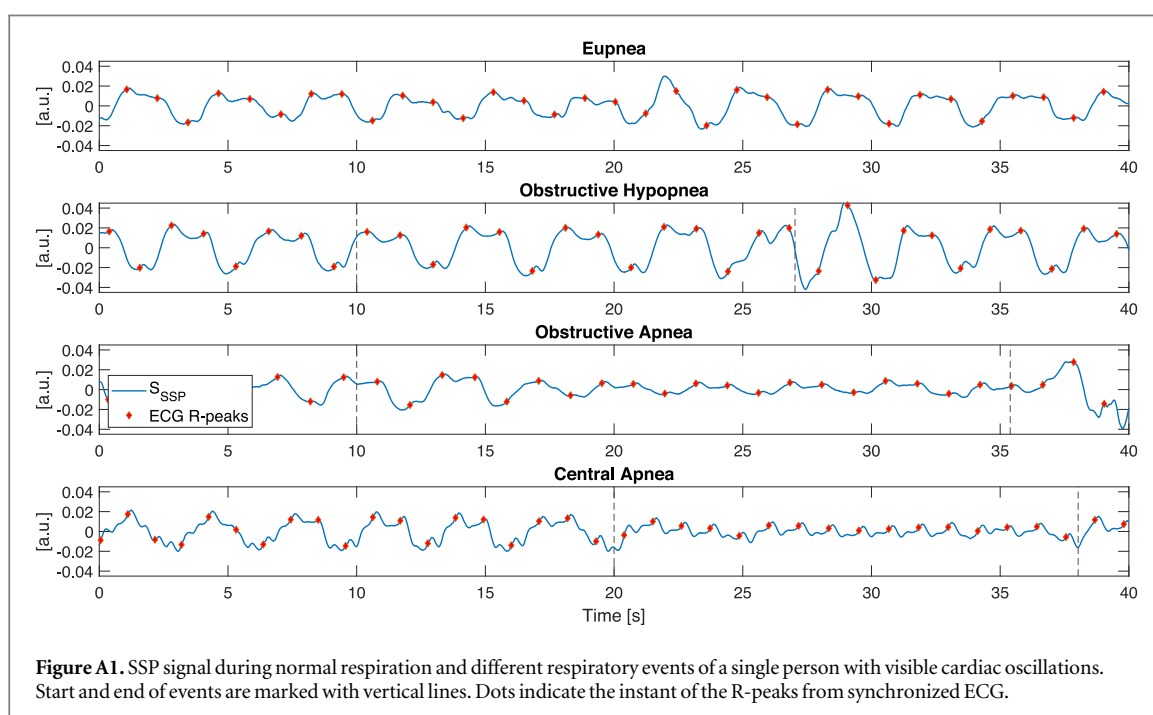

Figure A1. SSP signal during normal respiration and different respiratory events of a single person with visible cardiac oscillations. Start and end of events are marked with vertical lines. Dots indicate the instant of the R-peaks from synchronized ECG.

muscles is absent. In all cases, cardiac oscillations are visible at different levels. Normal respiration masks them entirely, except for the expiration phase with small visible inflections. They become more evident during hypopneas, with distinctive inflections in the signal, and follow the same pattern during obstructive apneas, although with a dampened amplitude profile. Lastly, cardiac oscillations are dominant during central apneas, with a temporal profile quite similar to the known PPG shape (Dillon and Hertzman 1941). The respiratory profile seems to shift towards cardiac oscillations in this example, even before the onset of the apnea.

Appendix B. Artifact rejection with HOT-SAX method

We describe how we can estimate the presence of an artifact (here called discords) using the HOT-SAX (Heuristically Ordered Time series using Symbolic Aggregate Approximation) method (Keogh *et al* 2005, Senin *et al* 2014).
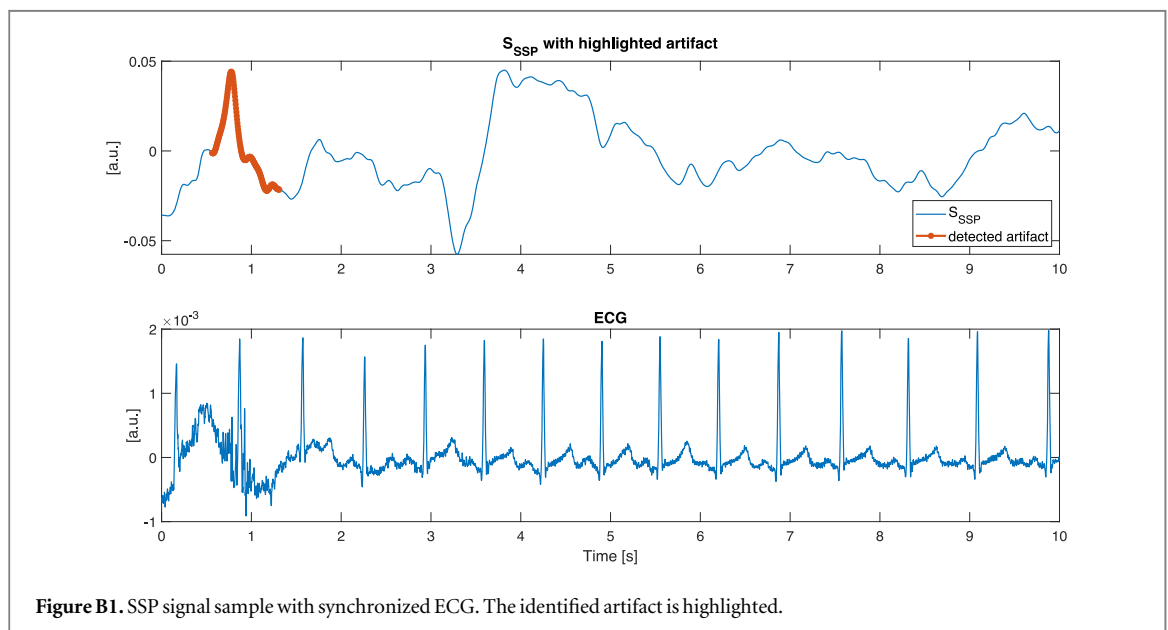
We start with the sample visible in figure B1 with a small artifact present. The method may be prone to false positives (given our choice in the parameters), but the presence of some noise also in the ECG signal increases the chance it is a real artifact.
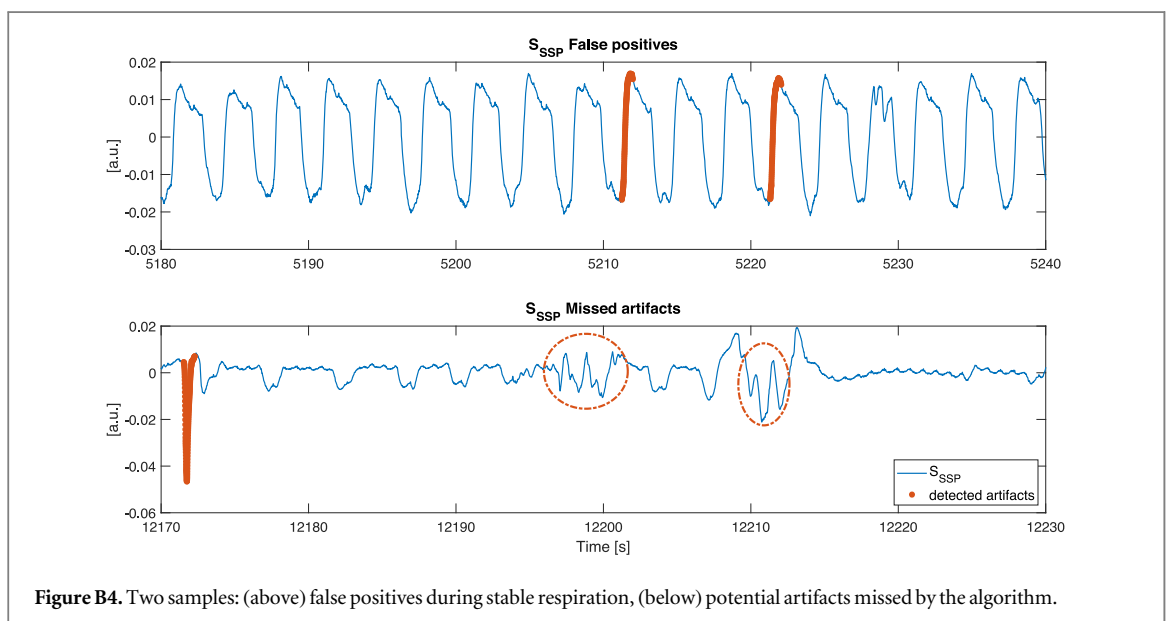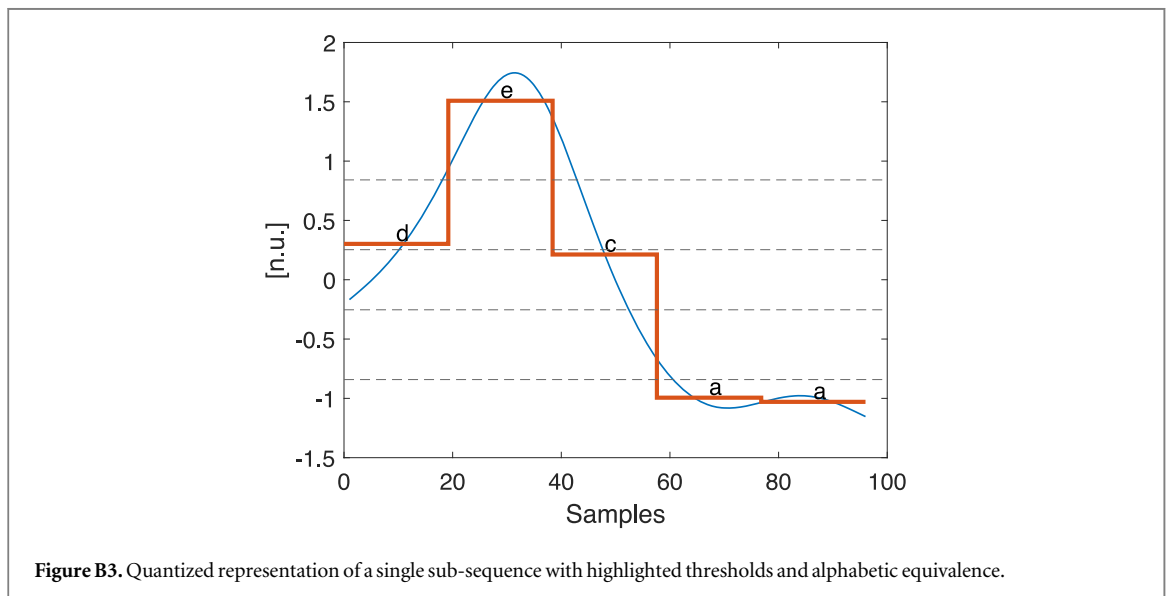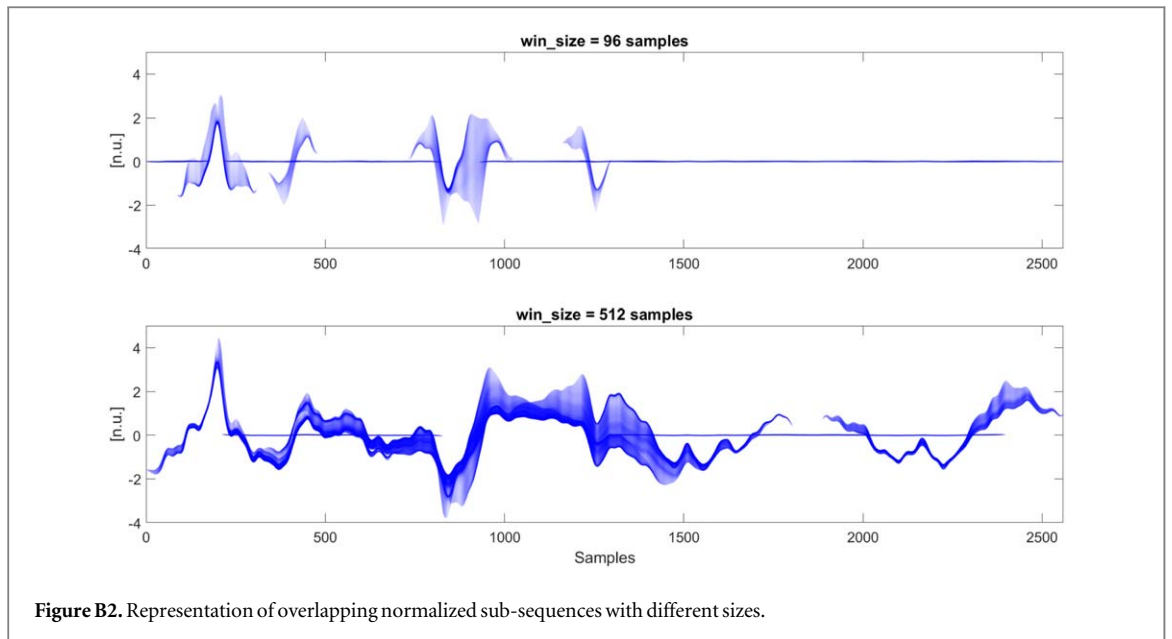
The algorithm starts by dividing the signal into N overlapping sub-sequences (with 1 sample step) of size win_{size}. This value determines the dimension of the discords we will find. The optimal value depends on the sampling frequency, the length of the analysis window, the expected length of discords, etc. In our case $win_{size} = 96$ is a reasonable compromise between granularity and coverage, as a value too large may force us to reject many estimates and at the same time miss smaller artifacts. Each sub-sequence is normalized subtracting the mean and dividing it by the standard deviation if larger than a threshold equal to 0.01. An example with $win_{size} = 96$ and $win_{size} = 512$ is included in figure B2 and it hints already where large unexpected variations are.

Then each sub-sequence is quantized using Piece-wise Aggregate Approximation and thresholds based on the normal probability distribution. The first step is to divide the sub-sequence into equally spaced frames (in our case $paa_{size} = 5$) and calculate the mean of each frame. Then the dimensionality is reduced by assigning a letter to each frame according to the desired resolution. We opted for an alphabet with size $a_{size} = 5$. The choice of paa_{size} and a_{size} is not as critical as win_{size}. The original authors observed that a_{size} between 3 and 5 is good for most applications, while paa_{size} should be high if the signal has fast-changing dynamics. Figure B3 shows an example on a single sub-sequence.

Once all sub-sequences are converted into words, the heuristic of HOT-SAX starts to look for discords sorting words from the least frequent, which is more likely to be a unique discord. To identify if the sub-sequence represents a discord or not, it calculates the euclidean distance between the original sub-sequence, normalized but not quantized, and all other sub-sequences in random order. The sub-sequence which has the largest distance with the others is the discord. If the parameters are tuned correctly, the ordering heuristic provides a higher chance of finding the sub-sequence with the largest distance early on.

Even with this heuristic, the research process can be time-consuming if we are looking for small artifacts in a signal that is hours long. Therefore we sub-divided the signal again in 10 s windows. The shift is 10 s if no discord

Figure B1. SSP signal sample with synchronized ECG. The identified artifact is highlighted.

Figure B2. Representation of overlapping normalized sub-sequences with different sizes.

Figure B3. Quantized representation of a single sub-sequence with highlighted thresholds and alphabetic equivalence.

Figure B4. Two samples: (above) false positives during stable respiration, (below) potential artifacts missed by the algorithm.

is found, or the final instant of the discord otherwise. We apply a distance threshold to separate large discords from false positives. We are still exploring the optimal combination of parameters, as it sparsely happens that normal breathing waves are classified as discord, or that windows with a lot of noise may be *self-similar* internally and therefore missed by this method. Figure B4 shows two examples of the problem. Future optimizations will look for parameters that balance HR estimation performance and coverage.

Appendix C. Algorithm parameters

We describe here in table C1 and figure C1 how different parameters in the algorithm operate and the motivation behind their value. Variations of these parameters may be necessary if the method is applied on other populations, or during different physiological states that are not related to sleep. All the parameters are fixed and do not change during the signal processing, while *tuned* parameters do in accordance to certain rules and are noted in the table with the symbol †. None of the parameters is tuned *a priori* upon known characteristics of the participant.
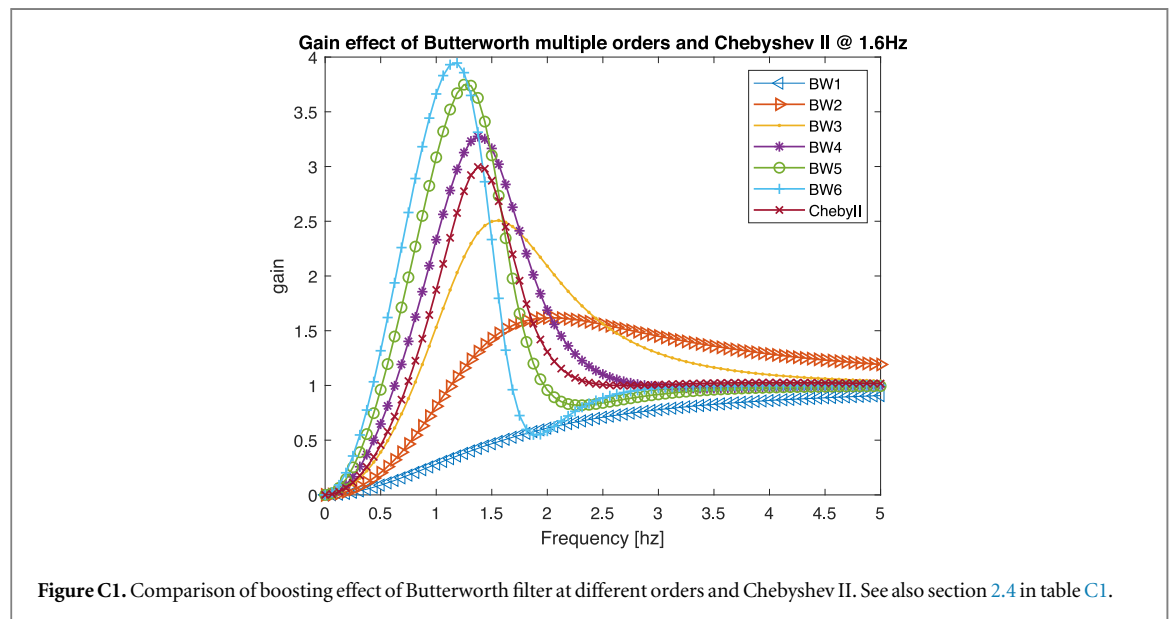

Figure C1. Comparison of boosting effect of Butterworth filter at different orders and Chebyshev II. See also section 2.4 in table C1.

Table C1. Parameters employed in our algorithm.

Parameter	Initial value	Motivation
		Section 2.2
Small signal selection threshold	90%	Elbow cutoff point in performance of our dataset
		Section 2.3
Analysis window (L)	10 s	Compromise between respiratory and heart rate stationarity (see also section 4.1)
Window step	2.5 s	Compromise between potential maximum rate of change in heart rate during sleep, time resolution of estimates, and algorithm execution speed
S_{Audio} lowpass cutoff	32 Hz	Reasonable margin for lower boundary of audible audio frequencies (18 Hz : 20 kHz)
Signal subsampling	256 Hz	algorithm execution speed
		Section 2.4
Respiration f_{min} range	[0.4, 0.95] Hz	Hypothesis of respiration rate harmonics during sleep between approximately between 12 and 30 breaths-per-minute
f_{min} fallback candidate	0.66 Hz	Hypothesis of respiration rate close to 20 breaths-per-minute

Table C1. (Continued.)

Parameter	Initial value	Motivation
Harmonics search range[†]	$[f_{min}, 2.4]$ Hz	Hypothesis on average heart rate observed. See section 2.1.
Harmonic fallback candidate	1.6 Hz	Hypothesis on average heart rate observed.
Respiration filter cutoff[†]	None	See section 2.4, function of the detected harmonics.
Order of respiration filter	Butterworth 4th	Optimization of booster bandwidth and signal gain. See also figure C1

<center>Section 2.5</center>

Parameter	Initial value	Motivation
Number of frames N	20	Empirical optimization of algorithm performance on ours dataset
Size of window (L)	$fs * 10$	Compromise between respiratory and heart rate stationarity (see also section 4.1)
Size of frame (M)[†]	$2.0 * lag(f_{bottom})$	Sampling criterion for ACF f_{bottom} frequency resolution
f_{bottom}, f_{top} HR range[†]	$[0.6, 1.8]$ Hz	Initial value based on heart rate during sleep 36 to 108 bpm, then updated according to section 2.5 rules
Outlier threshold	$Q_{3,1} \pm 1.5 * IQR$	Common choice for outliers removal
Skewness threshold	1.5	Empirical optimization of algorithm performance on our dataset

<center>Section 2.6</center>

Parameter	Initial value	Motivation
Rejection threshold	12.5% or 1.25 s	Correspondent to half window step, arbitrary choice
Sub-estimates rejection threshold	$3/4^*N$	Arbitrary choice on good performance in out dataset in detecting sensor disconnections

<center>Section 2.7</center>

Parameter	Initial value	Motivation
Bootstrap f_{bottom}[†]	$median(\hat{HR}_{booth}) - 0.5 * IQR(\hat{HR}_{booth})$	Assumption on reasonable lower boundary close to 1st quartile of $\hat{HR}_{booth}$
f_{bottom} margin[†]	$f_{bottom} - 0.05$ Hz	Arbitrary margin for potential corrections of erroneous Bootstrap f_{bottom}
f_{top} margin[†]	$\hat{HR}_{smooth}(n) + 0.25$ Hz	See relevant text. Based on assumptions of heart rate dynamics during sleep and OSA
f_{bottom}, f_{top} update thresholds[†]	$[0.36, 0.64]$	Empirical optimization of algorithm performance on our dataset
f_{bottom} deceleration[†]	1e-5 Hz/window	Assumption on heart rate deceleration during sleep. See section 4 for relevant issues

<center>Section 2.8</center>

Parameter	Initial value	Motivation
FIR and STFT window	2048	Balance between optimal nfft for a 10 s window (2560 samples) and resolution of the filter response.

ORCID iDs

Luca Cerina ⓘ https://orcid.org/0000-0001-8571-3971
Gabriele B Papini ⓘ https://orcid.org/0000-0002-5752-9226
Pedro Fonseca ⓘ https://orcid.org/0000-0003-2932-6402

References

Abrahamyan L, Sahakyan Y, Chung S, Pechlivanoglou P, Bielecki J M, Carcone S M, Rac V E, Fitzpatrick M and Krahn M 2017 Diagnostic accuracy of level iv portable sleep monitors versus polysomnography for obstructive sleep apnea: a systematic review and meta-analysis *Sleep Breathing* **22** 593–611
Alex R M, Mousavi N D, Zhang R, Gatchel R J and Behbehani K 2017 Obstructive sleep apnea: brain hemodynamics, structure, and function *J. Appl. Biobehav. Res.* **22** e12101
Amaddeo A, Fernandez-Bolanos M, Olmo Arroyo J, Khirani S, Baffet G and Fauroux B 2016 Validation of a suprasternal pressure sensor for sleep apnea classification in children *J. Clin. Sleep Med.* **12** 1641–7
Ayappa I A, Norman R G and Rapoport D M 1999 Cardiogenic oscillations on the airflow signal during continuous positive airway pressure as a marker of central apnea *Chest* **116** 660–6
Azarbarzin A, Ostrowski M, Hanly P and Younes M 2014 Relationship between arousal intensity and heart rate response to arousal *Sleep* **37** 645–53

Azarbarzin A *et al* 2021 The sleep apnea-specific pulse-rate response predicts cardiovascular morbidity and mortality *Am. J. Respiratory Crit. Care Med.* **203** 1546–55

Bakker J P, Ross M, Vasko R, Cerny A, Fonseca P, Jasko J, Shaw E, White D P and Anderer P 2021 Estimating sleep stages using cardiorespiratory signals: validation of a novel algorithm across a wide range of sleep-disordered breathing severity *J. Clin. Sleep Med.* **17** 1343–54

Berry R B *et al* 2012 Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events: deliberations of the sleep apnea definitions task force of the american Academy of Sleep Medicine *J. Clin. Sleep Med.* **8** 597–619

Brochard L J 2014 Measurement of esophageal pressure at bedside: pros and cons *Curr. Opin. Crit. Care* **20** 39–46

Brüser C, Winter S and Leonhardt S 2013 Robust inter-beat interval estimation in cardiac vibration signals *Physiol. Meas.* **34** 123–38

Chaudhari S, Koivunen V and Poor H V 2009 Autocorrelation-based decentralized sequential detection of ofdm signals in cognitive radios *IEEE Trans. Signal Process.* **57** 2690–700

Choi S J and Kim J S 2011 Clinically different phenotypes of obstructive sleep apnea according to 24 hours heart rate tachogram pattern *Sleep Med. Res.* **2** 21–6

Dillon J B and Hertzman A B 1941 The form of the volume pulse in the finger pad in health, arteriosclerosis, and hypertension *Am. Heart J.* **21** 172–90

Dubsky S, Thurgood J, Fouras A, R Thompson B and Sheard G J 2018 Cardiogenic airflow in the lung revealed using synchrotron-based dynamic lung imaging *Sci. Rep.* **8** 1–9

Edwards B A, Redline S, Sands S A and Owens R L 2019 More than the sum of the respiratory events: personalized medicine approaches for obstructive sleep apnea *Am. J. Respiratory Crit. Care Med.* **200** 691–703

Fonseca P, Long X, Radha M, Haakma R, Aarts R M and Rolink J 2015 Sleep stage classification with ecg and respiratory effort *Physiol. Meas.* **36** 20–27

Gil E, Laguna P, Martínez J P, Barquero-Pérez Ó, García-Alberola A and Sörnmo L 2013 Heart rate turbulence analysis based on photoplethysmography *IEEE Trans. Biomed. Eng.* **60** 3149–55

Glos M, Sabil A, Jelavic K S, Schöbel C, Fietze I and Penzel T 2018 Characterization of respiratory events in obstructive sleep apnea using suprasternal pressure monitoring *J. Clin. Sleep Med.* **14** 359–69

Huang Z, Goparaju B, Chen H and Bianchi M T 2018 Heart rate phenotypes and clinical correlates in a large cohort of adults without sleep apnea *Nat. Sci. Sleep* **10** 111–25

Keogh E, Lin J and Fu A 2005 Hot sax: efficiently finding the most unusual time series subsequence *Fifth IEEE Int. Conf. on Data Mining (ICDM'05)'* (IEEE) 1–8

Lévy P, Kohler M, McNicholas W T, Barbé F, McEvoy R D, Somers V K, Lavie L and Pépin J L 2015 Obstructive sleep apnoea syndrome *Nat. Rev. Dis. Primers* **1** 1–21

Makowski D, Pham T, Lau Z J, Brammer J C, Lespinasse F, Pham H, Schölzel C and Chen S H A 2021 NeuroKit2: a python toolbox for neurophysiological signal processing *Behav. Res. Methods* **53** 1689–96

Mejía-Mejía E, May J M, Torres R and Kyriacou P A 2020 Pulse rate variability in cardiovascular health: a review on its applications and relationship with heart rate variability *Physiol. Meas.* **41** 07TR01

Mesarwi O and Malhotra A 2020 Obstructive sleep apnea and pulmonary hypertension: a bidirectional relationship *J. Clin. Sleep Med.* **16** 1223–4

Moeyersons J, Amoni M, Huffel S V, Willems R and Varon C 2019 R-deco: an open-source matlab based graphical user interface for the detection and correction of r-peaks *PeerJ Comput. Sci.* **5** 1–20

Mukhopadhyay S K, Zara M, Telias I, Chen L, Coudroy R, Yoshida T, Brochard L and Krishnan S 2020 A singular spectrum analysis-based data-driven technique for the removal of cardiogenic oscillations in esophageal pressure signals *IEEE J. Transl. Eng. Health Med.* **8** 1–11

Peláez-Coca M D, Hernando A, Lázaro J and Gil E 2022 Impact of the ppg sampling rate in the pulse rate variability indices evaluating several fiducial points in different pulse waveforms *IEEE J. Biomed. Health Inform.* **26** 539–49

Rankawat S A and Dubey R 2017 Robust heart rate estimation from multimodal physiological signals using beat signal quality index based majority voting fusion method *Biomed. Signal Process. Control* **33** 201–12

Senin P, Lin J, Wang X, Oates T, Gandhi S, Boedihardjo A P, Chen C, Frankenstein S and Lerner M 2014 Grammarviz 2.0: a tool for grammar-based pattern discovery in time series *Joint European Conf. on Machine Learning and Knowledge Discovery in Databases* (Springer) pp 468–72

Sharma S, Stansbury R, Hackett B and Fox H 2021 Sleep apnea and pulmonary hypertension: a riddle waiting to be solved *Pharmacol. Therapeutics* **227** 107935

Shinar Z, Akselrod S, Dagan Y and Baharav A 2006 Autonomic changes during wake-sleep transition: a heart rate variability based approach *Autonomic Neurosci.* **130** 17–27

Sin P Y W, Webber M R, Galletly D C and Tzeng Y C 2012 Relationship between cardioventilatory coupling and pulmonary gas exchange *Clin. Physiol. Funct. Imaging* **32** 476–480

Sološenko A, Petrėnas A, Marozas V and Sörnmo L 2017 Modeling of the photoplethysmogram during atrial fibrillation *Comput. Biol. Med.* **81** 130–8

Strömbergsson S 2016 Today's most frequently used F0 estimation methods, and their accuracy in estimating male and female pitch in clean speech *INTERSPEECH Dresden* pp 525–9

Suarez-Sipmann F, Santos A, Peces-Barba G, Bohm S H, Gracia J L, Calderón P and Tusman G 2012 Pulmonary artery pulsatility is the main cause of cardiogenic oscillations *J. Clin. Monitoring Comput.* **27** 47–53

Sullivan S G and Szewczak J M 1998 Apneic oxygen uptake in the torpid pocket mouse Perognathus parvus *Physiol. Zool.* **71** 624–32

van Gilst M M *et al* 2019 Protocol of the SOMNIA project: an observational study to create a neurophysiological database for advanced clinical sleep monitoring *BMJ Open* **9** e030996

West J B and Hugh-Jones P 1961 Pulsatile gas flow in bronchi caused by the heart beat *J. Appl. Physiol.* **16** 697–702

Zhao Z and Zhang Y 2018 Sqi quality evaluation mechanism of single-lead ECG signal based on simple heuristic fusion and fuzzy comprehensive evaluation *Front. Physiol.* **9** 7–27