Fetal heart rate monitoring by fusion of estimations from two modalities: A modified Viterbi's algorithm

Rémi Souriau, Julie Fontecave-Jallon, Bertrand Rivet

Fetal heart rate monitoring by fusion of estimations from two modalities: a Modified Viterbi's algorithm

Rémi Souriau[a,b], Julie Fontecave-Jallon[a] and Bertrand Rivet[b]

[a]Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, Grenoble, 38000, France
[b]Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, 38000, France

ARTICLE INFO

Keywords:
Hidden Markov Chain
Fetal Heart Rate
Fusion estimation algorithm
Electrocardiogram
Phonocardiogram

ABSTRACT

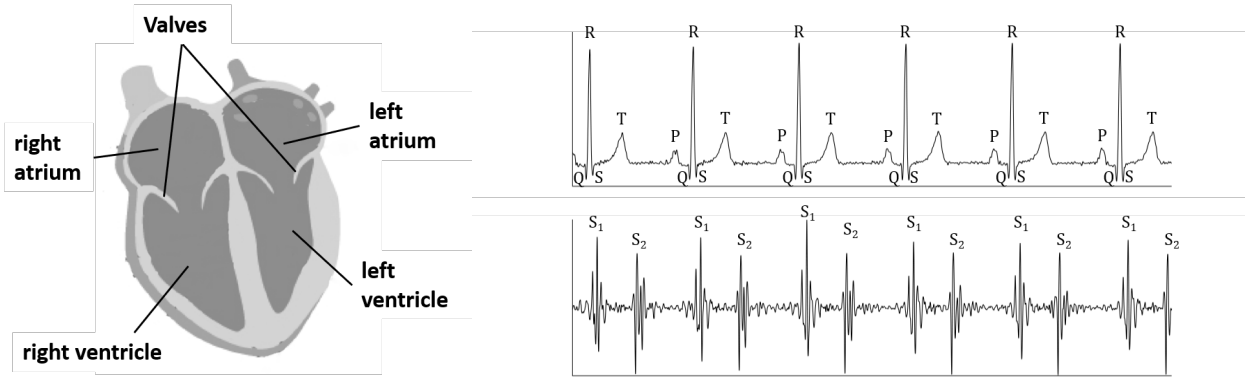
.*Objective*. In the context of fetal heart rate (fHR) estimation, the article addresses the fusion of two monomodal estimations into a multimodal one. Electrical and mechanical modalities are considered through the use of abdominal electrocardiogram (ECG) and phonocardiogram (PCG). The aim of the fusion is to provide a fHR estimation, robust against noise sources, and especially against the risk of confusion with the mother heart rate (mHR). *Approach*. A hidden Markov chain is considered to model variations of the two monomodal fHR estimations and the true fHR. Thanks to the Viterbi's algorithm (VA), the two fHR estimations are then merged. However, the classical VA does not ensure that the merged estimation always follows the true fHR: it may alternate between the fHR and the mHR, especially when the mother component is not correctly removed from the abdominal ECG. Therefore, a modified VA is proposed to efficiently avoid confusion with mHR. The aim is to discourage large variations between successive states. *Results*. Comparisons between classical and modified VA are performed on real pregnant women data. The modified VA reduces the confusion between fHR and mHR in major cases compared to the classical VA. For most recordings, the confusion with mHR decreases under 1%, and for the best case, it is reduced from 59% to 0%. *Conclusion*. The modified VA succeeds to improve the fHR estimation by reducing confusion with mHR. *Significance*. Fusion of estimations from two modalities is a promising approach for more robust fHR monitoring.

1. Introduction

The fetal heart rate (fHR) estimation is a key element in monitoring the health status of the fetus during pregnancy. Study of fHR and its variability may highlight fetal distress that can lead to the need of an emergency delivery. Nowadays, the cardiotocography (CTG) is used as the clinical reference for the fHR monitoring, as recommended by the International Federation of Gynecology and Obstetrics [1]. The CTG is based on Doppler ultrasounds and monitors the fHR as well as the uterine contractions during labor. However, the use of CTG presents some limitations like signal loss or confusion between the fHR and the mother heart rate (mHR) [2]. Some studies show that the use of CTG increases the cesareans rate [3, 4]. These limitations justify the interest of research for alternative techniques for fetal monitoring. One can cite methods like the fetal scalp electrodes [5, 6] or the fetal magnetocardiography [7, 8]. These methods can provide higher quality measures but present also different disadvantages (highly invasive, expensive, etc). Non-invasive methods, much more ergonomic for clinical uses, are of high interest for fHR monitoring. Among them, *electrocardiography* [9] allows to measure the electrical activity of the fetal heart by placing electrodes on the mother abdomen. The obtained abdominal electrocardiogram (ECG) signals are corrupted by maternal heart activity and require a preprocessing step to extract the fetal ECG (ECGf). ECGf is widely studied in the literature: many methods have been proposed to remove noises [10], and there are many databases of synthetic and real ECG recordings available to test algorithms (see [11, 12] for examples). Another non-invasive approach is the *phonocardiography* [13] which consists in measuring the mechanical activity of the heart using a microphone. The use of the phonocardiogram (PCG) for fetal monitoring regains in interest the past recent years [14]. It gives access to fetal PCG (PCGf) and fetal heart sounds, from which fHR can be estimated. In the present study, the ECG and the PCG are used simultaneously to estimate the fHR.

Fig. 1a gives a heart scheme and Fig. 1b is an illustration of ECG and PCG signals. At each cardiac cycle, the passing electric current through the heart (or myocardium) generates P,Q,R,S,T waves during contractions and relaxations of

remi.souriau@univ-grenoble-alpes.fr (R. Souriau); julie.fontecave@univ-grenoble-alpes.fr (J. Fontecave-Jallon); bertrand.rivet@grenoble-inp.fr (B. Rivet)
ORCID(s): 0000-0003-0849-9515 (R. Souriau); 0000-0001-5764-3144 (J. Fontecave-Jallon); 0000-0003-4901-5302 (B. Rivet)



(a) Scheme of the heart. (b) Synchronized ECG and PCG signals.

Figure 1: Heart and biosignals. **(a)** is a scheme which summarizes the different part of the heart and **(b)** gives an illustration of the theoretical ECG (top) and PCG (bottom) signals. The location of P,Q,R,S,T waves are given for the ECG signal. S_1 and S_2 in the PCG signal give the location of, respectively, systoles and diastoles.

the atria and ventricles. The opening and closing of valves during the systole and the diastole product sounds which can be observed in the PCG signal and are noted S_1 and S_2.

From ECG and PCG signals, there are different methods based on R peaks detection [15, 16] or S_1 and S_2 sounds detection [17] to estimate the heart rate of an adult. However, for fHR estimation, a first step consists in extracting the ECGf and the PCGf from, respectively, the abdominal ECG (ECGa) and the abdominal PCG (PCGa). This step is not easy due to noise sources in abdominal signals. This is especially true for the ECGa because the mother heart activity is also present and the signal power of the mother is significantly greater than the fetal heart activity in most cases. Different methods to extract the ECGf have been proposed, like template subtraction [18], source separation [19, 20] or neural network [21]. In order to be robust against noise sources, most of these methods require to add redundancy of information by using more than a single sensor on the woman abdomen (three to four ECGa in [19], five in [20] for example). However, such a solution is not ergonomic for the clinician and the mother in real conditions where emergency surgery can be needed.

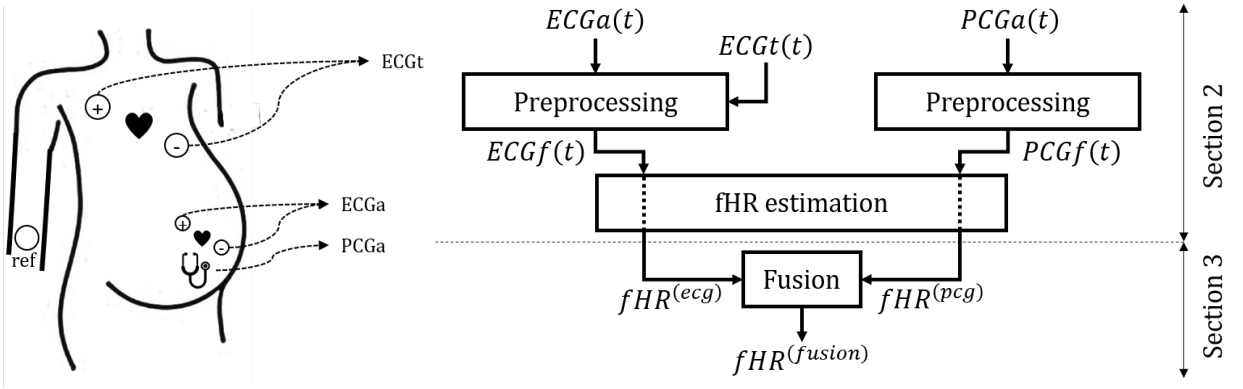
Therefore, an alternative solution to estimate the fHR from a few numbers of sensors is to consider multimodality. Its interest lies on the complementary and redundant information from modalities to gain in robustness against noise sources. Multimodality has been identified as a promising approach for fetal monitoring [22]. For example, the study of [23] used an ECGa signal coupled with a PCGa signal to extract the ECGf. However, only a very few studies focused on fHR estimation from multiple modalities. In [24], four ECG signals are combined with one seismocardiographic recording and one gyrocardiographic recording for fHR estimation. The use of PCG for fHR estimation is considered in [25] and the authors used the ECG modality to validate their fHR.

In a previous study [26], we investigated a preliminary approach of fusion of fHR data from ECG and PCG for fHR estimation. By considering the reference fHR as a latent dynamic process and the two estimations as observations processes, a bimodal *hidden Markov model (HMM)* is used to model these processes. In the HMM, the underlying laws between the variables are probability laws. The reference fHR is assumed to be provided by the CTG (recorded in parallel to the ECG and the PCG). And the probability laws are modelled according to some recording sessions. The *Viterbi's algorithm (VA)* allows then to decode the true fHR from the two fHR estimations and probability laws.

However, because of the possible presence of the mHR in the fHR estimation inputs, the VA can fail to estimate the fHR. In this current paper, a modified version of the VA is then proposed to improve robustness against this maternal noise source.

2. Monomodal estimation of the fetal heart rate

The steps of the complete process of fHR estimation are summarized on Fig. 2b. After description of the clinical protocol (Sec. 2.1), this section summarizes the different steps preceding the fusion for multimodal estimation, which will be described in Sec. 3. Before this fusion, the monomodal fHR estimation (Sec. 2.2) using each single modality (ECG or PCG) is necessary and separated into two steps: (1) the preprocessing (Sec. 2.2.1) to extract, on one hand,



(a) Signals recording protocol.

(b) Methodology of fHR estimation.

Figure 2: Recording and fHR estimation. **(a)** is a scheme of the recording protocol of the studied database. The CTG (not represented here) has also been recorded for evaluation. **(b)** gives the methodology of fHR estimation from the ECG and the PCG modality. After pre-processing, the monomodal fHR estimation step is performed on the ECG and the PCG separately. This step similar for the two modalities is detailed in [9, 13] and summarized in Sec. 2.2. The fusion step is detailed in Sec. 3.

the ECGf from the ECGa and the thoracic ECG (ECGt) and, on the other hand, to extract the PCGf from the PCGa, and (2) the fHR estimation method (Sec. 2.2.2) which is similar for both modalities. It is worth noting that all the steps from abdominal recordings (ECGa and PCGa) to monomodal fHR estimations ($fHR^{(ecg)}$ and $fHR^{(pcg)}$) are recalled for clarity of the paper and do not contain new contribution.

2.1. Clinical protocol

Real data considered in this paper comes from six volunteer pregnant women older than 18 years old, between 38(+0) and 38(+4) weeks(+days) of gestation with no maternal or fetal complication, as part of a clinical protocol established at the University Hospital of Grenoble (study No RCB: 2018-A03182-53) [27]. After signature of informed consent, the volunteer was laying on her back, in a comfortable position, in order to minimize movements and electrical interference. As illustrated on Fig. 2a, recording sessions consisted in the acquisition of synchronous signals (Powerlab acquisition system, ADInstruments, sampling frequency 1KHz):

- one ECGt with two thoracic bipolar electrodes on the maternal heart axis and a reference electrode on the maternal wrist (BioAmp, ADInstruments).
- one ECGa with two abdominal bipolar electrodes across the axis of the fetal heart (defined after clinical auscultation) and the same reference electrode as for the thoracic ECG (BioAmp, ADInstruments).
- one PCGa with a cardio-microphone (MLT201, ADInstruments) put on the skin of the maternal abdomen as close as possible to the fetal heart.

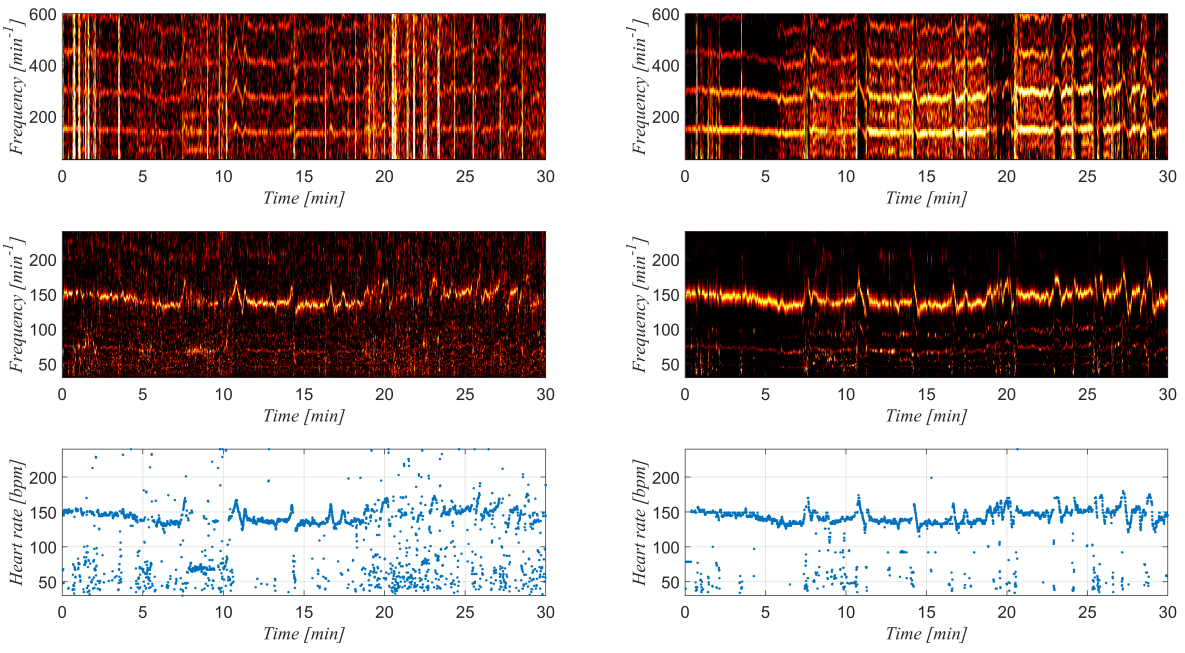
In addition, the fHR was also directly estimated using the reference CTG (Avalon F20/F30, Philips) [1]. The sampling frequency of fHR measure from CTG has been set to 4Hz. The duration of each recording session was about 30 minutes in rest condition for the mother.

2.2. Methodology of monomodal fHR estimation

This section describes the two steps of the methodology for monomodal fHR estimation using a single modality (ECG or PCG).

2.2.1. Preprocessing

The preprocessing step consists in attenuating the noise sources of the ECGa and the PCGa in order to obtain an approximation of the ECGf and the PCGf [9, 13]. An approximation of these signals is sufficient in the context of our proposed fHR estimation, since this latter is not based on events detection. For the ECGa, the baseline is filtered with a high-pass filter (finite impulse response(FIR), order 1024, cut-off frequency at 10 Hz) to enhance the fetal



(a) Estimation of $\text{fHR}^{(\text{ecg})}$ (denoted $Y^{(1)}_{[0:T]}$ in Sec. 3.1).

(b) Estimation of $\text{fHR}^{(\text{pcg})}$ (denoted $Y^{(2)}_{[0:T]}$ in Sec. 3.1).

Figure 3: Monomodal fHR estimation steps. **(a)** and **(b)** give, respectively, fHR estimation for the ECG modality and the PCG modality. From top to bottom, spectrograms of the ECGf and the PCGf (Fs = 1kHz, window = 4s, shift = 250ms, zero padding ratio = 4) are first given. Then, excitation matrices $H^{(e)}$ obtained from spectrograms using the NMF algorithm [13, 9] are provided. fHR estimations deduced from excitation matrices are shown in the last row.

R peaks. Then, the power-line inference is removed using a band-stop filter (Butterworth filter, order 2, stop band [49-51] Hz) and a low pass filter (FIR, order 100, cut-off frequency 80 Hz) is applied to attenuate high frequency noise [9]. Finally, the mother heart activity in the ECGa is attenuated using Approximate Linear Dependency Kernel Recursive Least-Squares (ALD-KRLS) [28] with the ECGt signal as reference (see [9, figure 2] for an example of mother ECG component removal). Considering PCGa, it is first filtered using a band-pass filter (FIR, order 100, pass band [20-200] Hz) to focus on the frequency range of the heart sounds and get the PCGf signal. The absolute value is then filtered with a low pass filter (FIR, order 100, cut-off frequency 15 Hz) to highlight the envelop of the PCGf [13]. The signal output is therefore not exactly the PCGf but a smooth signal for which the periodic aspect is well characterized and from which fHR can be more easily estimated.

2.2.2. Monomodal fHR estimation

After preprocessing steps, fHR estimations are carried out, separately from the ECGf and the PCGf signals. This monomodal fHR estimation step is exactly the same for both modalities, as illustrated on Fig. 3. First, spectrograms of the ECGf and the PCGf are computed with a window (hamming) of 4s and shifted by 250ms (Fig. 3, first line). Choosing a window of 4s allows to highlight the frequential harmonic structure of the ECGf and the PCGf. A shift of 250ms allows, at the end, a fHR estimation at the same sampling frequency than the CTG (i.e. 4 Hz).

Let Z be the spectrogram matrix from one modality. Using the source-filter model, Z can be decomposed as $Z = Z^{(e)} \odot Z^{(\phi)}$ where $Z^{(e)}$ and $Z^{(\phi)}$ are, respectively, the excitation and the filter matrices of the spectrogram [9, 13]. The excitation matrix contains the heart rate information and the filter matrix stores the waveform of each heartbeat. $\odot$ is the Hadamard product. $Z^{(e)}$ and $Z^{(\phi)}$ are factorized using the NMF algorithm [9, 13]:

$$Z = \left(W^{(e)}H^{(e)}\right) \odot \left(W^{(\phi)}H^{(\phi)}\right). \tag{1}$$

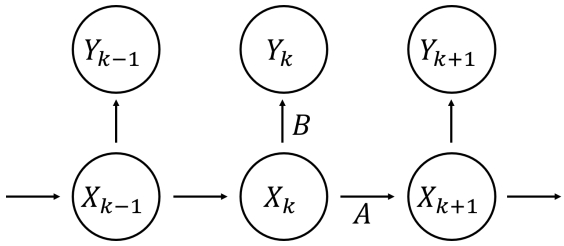


Figure 4: A classic HMM with one latent variable X_k and one observation Y_k. Links represent the conditional dependencies between variables. These dependencies are modelled with a transition matrix A and an emission matrix B.

$W^{(e)}$ is a dictionary of Dirac combs whose fundamental frequencies covered the frequency range of the fHR, and the time evolution of the fHR is given by $H^{(e)}$ (Fig. 3, second line). The vector of fHR estimations (Fig. 3, third line) is finally extracted from the matrix $H^{(e)}$: following [9, 13], at each time j, the fHR estimation is given by the frequency associated to the maximal value in column j of $H^{(e)}$.

The two monomodal fHR estimations (Fig. 3a from ECG modality and Fig. 3b for PCG modality) are denoted $\text{fHR}^{(\text{ecg})}$ and $\text{fHR}^{(\text{pcg})}$, and are used as inputs of the next step, which is the fusion step.

3. Multimodal estimation of the fetal heart rate

This section gathers the main contribution of this paper and is dedicated to the fusion step where $\text{fHR}^{(\text{ecg})}$ and $\text{fHR}^{(\text{pcg})}$ are merged into a new estimation denoted $\text{fHR}^{(\text{fusion})}$.

The considered hidden Markov model (HMM) to model the fHR evolution is first presented in (Sec. 3.1). Then, the Viterbi's algorithm (VA) is described in (Sec. 3.2) and its limits in (Sec. 3.3). Finally, a modified version of VA is proposed in (Sec. 3.4) to overcome limits of the classical VA.

3.1. Hidden Markov models

The HMM is used in this paper to model the evolution of fHR time series. The simplest case of HMM is given in Fig. 4. where $X_{0:T} = \left\{X_k\right\}_{k\in[\![0,T]\!]}$ is the latent path and $Y_{0:T} = \left\{Y_k\right\}_{k\in[\![0,T]\!]}$ is the observation path (where k is the time index). The time evolution of the latent path is characterized by a probability law called the *transition law* and the probability rule of the observation state given the latent state is the *emission law*.

Fig. 5 gives the model used in this study. Let denote:

- $X_{0:T} = \left\{X_k\right\}_{k\in[\![0,T]\!]}$ the true fHR;
- $X^{(1)}_{0:T} = \left\{X^{(1)}_k\right\}_{k\in[\![0,T]\!]}$ and $X^{(2)}_{0:T} = \left\{X^{(2)}_k\right\}_{k\in[\![0,T]\!]}$ the fHR according to, respectively, the first modality and the second modality;
- $Y^{(1)}_{0:T} = \left\{Y^{(1)}_k\right\}_{k\in[\![0,T]\!]}$ and $Y^{(2)}_{0:T} = \left\{Y^{(2)}_k\right\}_{k\in[\![0,T]\!]}$ the estimation of the fHR from both modalities: $\text{fHR}^{(\text{ecg})}$ and $\text{fHR}^{(\text{pcg})}$.

All the fHRs and their variations are modelled in the complete HMM (Fig. 5a). $X_{0:T}$ refers to the true fHR, $X^{(1)}_{0:T}$ and $X^{(2)}_{0:T}$ are, respectively, the representation of the fHR in the electrical modality and the mechanical modality. And $Y^{(1)}_{0:T}$ and $Y^{(2)}_{0:T}$ are, respectively, fHR estimations from these modalities using the ECGf and the PCGf. In practice, the complete HMM cannot be used because $X^{(i)}_{0:T}$ ($i \in \{1,2\}$) as well as probability rules $\Pr\left(X^{(i)}_k | X_k\right)$ for all k are unknown. Because, for each k and $i \in \{1,2\}$, there is no temporal link between $X^{(i)}_{k+1}$ and $X^{(i)}_k$, the complete HMM can be reduced to the HMM given in Fig. 5b where emissions from X_k to $X^{(i)}_k$ and $X^{(i)}_k$ to $Y^{(i)}_k$ are modelled with an emission matrix for both modalities.

Let consider the HMM given in Fig. 5b. The probability law between two successive latent states (the transition rule) and the probability laws between current latent state and observations (the emission rules) are noted:

$$\forall k \geq 1,\ \forall (i,j) \in [\![1,N]\!]^2, \quad \begin{cases} a_{ij} = \Pr\left(X_k = j \middle| X_{k-1} = i\right) \\ b_{ij}^{(1)} = \Pr\left(Y_k^{(1)} = j \middle| X_k = i)\right) \\ b_{ij}^{(2)} = \Pr\left(Y_k^{(2)} = j \middle| X_k = i\right) \end{cases}, \tag{2}$$

where N is the number of states which are directly written with index i and j for simplification. Coefficients $A = \left(a_{ij}\right)_{1\leq i,j\leq N}$, $B^{(1)} = \left(b_{ij}^{(1)}\right)_{1\leq i,j\leq N}$ and $B^{(2)} = \left(b_{ij}^{(2)}\right)_{1\leq i,j\leq N}$ form, respectively, the *transition matrix* and the *emission matrices*. The initial state X_0 is defined by its probability $\Pr(X_0)$. Both modalities are assumed to be subjected to *independent* noise sources, so the joint probability of observations given the state satisfies

$$\forall (i,j,m) \in [\![1,N]\!]^3, \quad \Pr\left(Y_k^{(1)} = j, Y_k^{(2)} = m \middle| X_k = i\right) = \Pr\left(Y_k^{(1)} = j \middle| X_k = i\right) \Pr\left(Y_k^{(2)} = m \middle| X_k = i\right). \tag{3}$$

The transition law between two successive states is modelled by the following probability rule:

$$a_{ij} = \Pr(X_k = j | X_{k-1} = i) = \frac{\Lambda_q(j-i) + \epsilon}{\sum_{j'=1}^{N} \left(\Lambda_q(j'-i) + \epsilon\right)} \tag{4}$$

where ϵ is a small coefficient to avoid numerical problems and $\Lambda_q(.)$ is the triangular function of width q:

$$\Lambda_q(x) = \frac{1}{q} \max\left(q - |x|, 0\right). \tag{5}$$

The emission rules for the ECG and the PCG are assumed to be Gaussian mixture centered around the latent state:

$$\begin{cases} \Pr(Y_k^{(ecg)} = j | X_k = i) = \alpha_i^{(ecg)} \left(\pi^{(ecg)} \mathcal{N}(j; i, (\sigma_1^{(ecg)})^2) + (1 - \pi^{(ecg)}) \mathcal{N}(j; i, (\sigma_2^{(ecg)})^2)\right) \\ \Pr(Y_k^{(pcg)} = j | X_k = i) = \alpha_i^{(pcg)} \left(\pi^{(pcg)} \mathcal{N}(j; i, (\sigma_1^{(pcg)})^2) + (1 - \pi^{(pcg)}) \mathcal{N}(j; i, (\sigma_2^{(pcg)})^2)\right) \end{cases}. \tag{6}$$

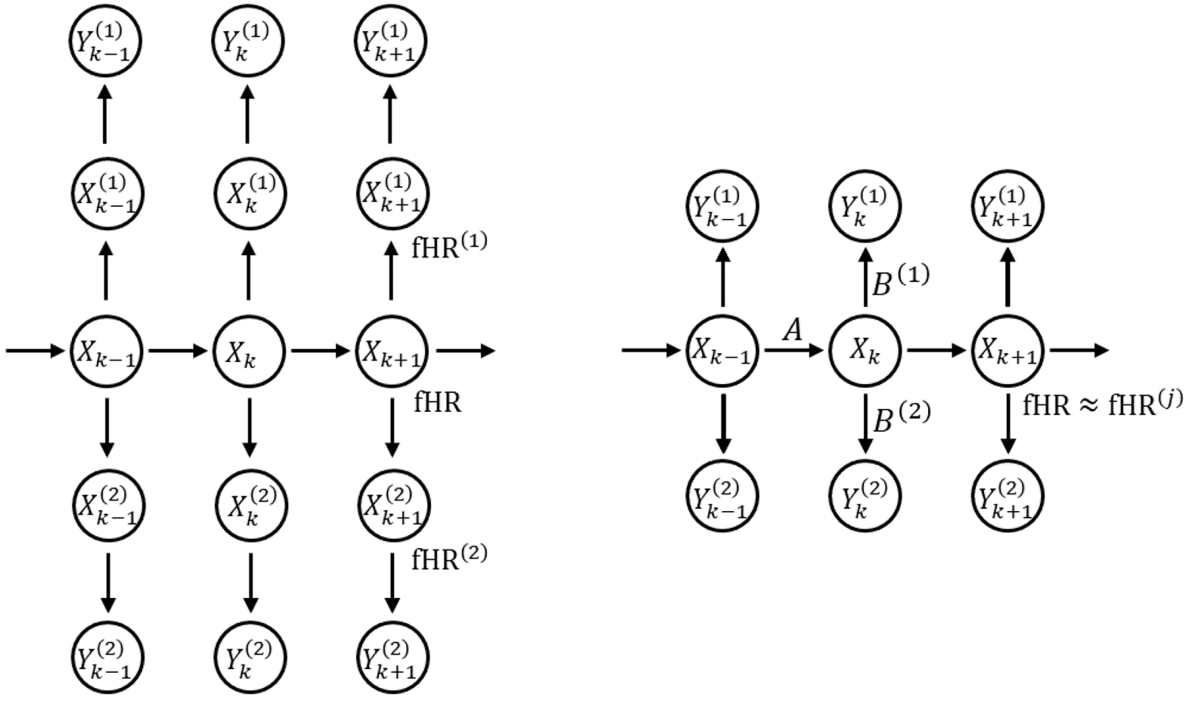
The first standard deviation $\sigma_1^{(ecg)}$ (or $\sigma_1^{(pcg)}$) is small and models the fHR estimation close to the actual fHR. The second standard deviation $\sigma_2^{(ecg)}$ (or $\sigma_2^{(pcg)}$) is large and models the error of fHR estimation. $\pi^{(ecg)}$ (or $\pi^{(pcg)}$) is the ratio between the two Gaussians. Finally, the coefficient $\alpha_i^{(ecg)}$ (or $\alpha_i^{(pcg)}$) is a normalization coefficient to ensure the sum of all probability is equal to one.

Parameter values. Considering clinical specifications of fHR, the set of all possible states for the hidden state and the observation states are fixed to $[\![30, 250]\!]$.

It is worth noting that in the emission laws Eq. 6, only the means of the normal distributions depend on the state value X_k. The emission parameters are estimated from fHR estimations of the volunteer pregnant women who participated to the clinical study presented in Sec. 2.1. Parameter values of the emission matrix for the ECG modality are $\sigma_1^{(ecg)} = 2.4$, $\sigma_2^{(ecg)} = 56.2$ and $\pi^{(ecg)} = 0.6$. Parameter values of the emission matrix for the PCG modality are $\sigma_1^{(pcg)} = 1.6$, $\sigma_2^{(pcg)} = 63.1$ and $\pi^{(pcg)} = 0.5$.

In the same way, only the mean of the transition law Eq. 4 depends on the previous state value X_{k-1}. Moreover, q is the width of the triangular law which describes the variation of fHR. Considering the sampling frequency of fHR estimations at 4Hz, such a transition law discourages then the fHR variation to be larger than q beat per minute (bpm) in 250ms. According to real data from reference CTG, it appears that variations of fHR in 250ms are lower than 3 bpm in 98% of cases, thus the value of q is fixed to 3 bpm. Finally, the ϵ coefficient allows to avoid some numerical problems, it is arbitrary fixed to 10^{-10}.

At last, an initial probability law $\Pr(X_0)$ is needed to merge the two estimations. According to [29], the fHR generally varies between 110-160 bpm. $\Pr(X_0) = \mathcal{N}(135, 50)$ is then proposed to have a law centered around the mean fHR and whose probability $\Pr(X_0 \in [110, 160]) = 40\%$.



(a) Complete model.

(b) Model for this study.

Figure 5: HMMs to model fHR's evolution. The model given in **(a)** is closer to the reality where the true fHR, the fHR given by the different modalities and the estimations are all represented. Because there are no temporal links between two successive $X^{(1)}$ and two successive $X^{(2)}$, the model in **(a)** can be reduced to the model in **(b)** which is retained in this study.

3.2. Classical Viterbi's algorithm

The VA [30] is an algorithm which estimates the latent path $X_{0:T}$ according to the observation paths $Y_{0:T}^{(1)}$ and $Y_{0:T}^{(2)}$. The aim of the VA is to find the latent path $\widehat{X}_{0:T}$ which maximizes the conditional probability $\Pr\left(X_{0:T}\middle|Y_{0:T}^{(1)}, Y_{0:T}^{(2)}\right)$, *i.e.* :

$$\widehat{X}_{0:T} = \underset{X_{0:T}}{\operatorname{argmax}} \Pr\left(X_{0:T}\middle|Y_{0:T}^{(1)}, Y_{0:T}^{(2)}\right). \tag{7}$$

Because of the conditional probability rule, $\widehat{X}_{0:T}$ can also be written as:

$$\widehat{X}_{0:T} = \underset{X_{0:T}}{\operatorname{argmax}} \Pr\left(X_{0:T}, Y_{0:T}^{(1)}, Y_{0:T}^{(2)}\right). \tag{8}$$

In the HMM, the state X_k is assumed to depend on the previous states. This is why, to solve the previous equation, one can use the recursive following term:

$$\mu(X_k) = \max_{X_{0:k-1}} \Pr\left(X_{0:k}, Y_{0:k}^{(1)}, Y_{0:k}^{(2)}\right) \tag{9}$$

Because of assumptions made on the HMM, one can write $\mu(X_k)$ as a temporal sequence which depends on the previous value $\mu(X_{k-1})$ and the new observation(s).

- If $k > 0$, then $\mu(X_k)$ is given by:

$$\mu(X_k) = \max_{X_{0:k-1}} \Pr\left(X_{0:k-1}, X_k, Y^{(1)}_{0:k-1}, Y^{(1)}_k, Y^{(2)}_{0:k-1}, Y^{(2)}_k\right) \tag{10}$$

By using conditional probability rule and assumptions of the HMM, one can write from Eq. 10:

$$\begin{aligned}\mu(X_k) &= \max_{X_{0:k-1}} \left\{\Pr\left(X_{0:k-1}, Y^{(1)}_{0:k-1}, Y^{(2)}_{0:k-1}\right) \Pr\left(X_k \middle| X_{k-1}\right) \Pr\left(Y^{(1)}_k \middle| X_k\right) \Pr\left(Y^{(2)}_k \middle| X_k\right)\right\},\\ &= \max_{X_{k-1}} \left\{\Pr\left(X_k \middle| X_{k-1}\right) \Pr\left(Y^{(1)}_k \middle| X_k\right) \Pr\left(Y^{(2)}_k \middle| X_k\right) \max_{X_{0:k-2}} \left[\Pr\left(X_{0:k-1}, Y^{(1)}_{0:k-1}, Y^{(2)}_{0:k-1}\right)\right]\right\},\\ &= \max_{X_{k-1}} \left\{\Pr\left(X_k \middle| X_{k-1}\right) \Pr\left(Y^{(1)}_k \middle| X_k\right) \Pr\left(Y^{(2)}_k \middle| X_k\right) \mu(X_{k-1})\right\}. \end{aligned} \tag{11}$$

The passage from the first line to the second one in the above equation is performed using the following property: If $\forall a,\ f(a) \geq 0$ and $\forall (a,b),\ g(a,b) \geq 0$, then:

$$\max_{a,b} f(a)g(a,b) = \max_a \left\{f(a) \max_b g(a,b)\right\},$$

- If $k = 0$, then $\mu(X_0)$ is given by:

$$\mu(X_0) = \Pr\left(X_0\right) \Pr\left(Y^{(1)}_0 \middle| X_0\right) \Pr\left(Y^{(2)}_0 \middle| X_0\right) \tag{12}$$

In practice, logarithms of Eqs. 11 & 12 are used to avoid numerical issues in algorithms.

Once $\mu(X_k)$ is known, then the last state X_T can be estimated by finding the state which maximizes $\mu(X_T)$:

$$\widehat{X}_T = \underset{i}{\text{argmax}}\, \mu(X_T = i). \tag{13}$$

And, all the previous states can be deduced using back tracking and maximum index of each $\mu(X_k)$. Algo 1 summarizes the classical VA.

3.3. Limits of the classical VA

In the classical VA, the first step consists in computing the vector:

$$\log \mu(X_k) = \Big(\log \mu(X_k = 1), \ldots, \log \mu(X_k = N)\Big)^T$$

for each k using Eq. 9 (see Recursion in Algo 1). Then, the path $\widehat{X}_{0:T}$ ($\text{fHR}^{(\text{fusion})}$) is deduced with a back-tracking loop. For fHR estimation, such a selection rule can fail to provide a good estimation. If one of the modality estimations is wrong over a long time, then there is no guarantee that $\text{fHR}^{(\text{fusion})}$ will stay on the fHR. This is in particular true for the ECG modality where the mHR can be estimated instead of the fHR if the mother component is not correctly removed from the ECGa signal.

One illustration on real data recording is given in Fig. 6. The classical VA has been applied on 30 minutes of fHR estimations and the result of the fusion is displayed on a 50s-window for better visualization. Fig. 6a gives $\text{fHR}^{(\text{ecg})}$, $\text{fHR}^{(\text{pcg})}$ and $\text{fHR}^{(\text{fusion})}$ according to the VA. Estimations are compared with the reference fHR from the CTG ($\text{fHR}^{(\text{ref})}$) and the reference mHR estimated from the ECGt ($\text{mHR}^{(\text{ref})}$) by R-peaks detection. In this example, the algorithm to remove the mother component on ECGa failed and $\text{fHR}^{(\text{ecg})}$ modality gives the mHR all the time. On the other hand, $\text{fHR}^{(\text{pcg})}$ gives the expected heart rate. With such inputs, $\text{fHR}^{(\text{fusion})}$ with the classical VA alternates between the fHR and the mHR.

Fig. 6b gives values of $\log \mu(X_k)$ which allows to estimate $\text{fHR}^{(\text{fusion})}$ in Fig. 6a. The Y-axis values are the $\log \mu(X_k)$ and the X-axis corresponds to the states of $\text{fHR}^{(\text{fusion})}$. For each k, values of $\log \mu(X_k)$ decrease in time: $\log \mu(X_k = i) > \log \mu(X_{k+1} = 1)$, $\forall i \in [\![1, N]\!]$. It is then possible to read the time evolution of $\log \mu(X_k)$ in the figure be reading it from the top to the bottom. Local maxima of $\log \mu(X_k)$ are given by the "*" points for each k. Two main paths of local maxima can be observed in the figure: the mHRs' path (between 50-100 bpm) and the fHRs' path (between 100-150 bpm). Dashed lines are $\log \mu(X_k)$ synchronized with instant k where $\text{fHR}^{(\text{fusion})}$ starts and ends to follow the mHRs' path. This example shows that the VA does not ensure that the $\text{fHR}^{(\text{fusion})}$ always follows the fHRs' path despite the presence of local maxima on this path.

The VA is therefore modified to overcome this limitation by changing the termination and the back-tracing steps of the classic VA. The main idea of the modified VA consists in using the location of all local maxima of $\log \mu(X_k)$ and the estimated state X_{k-1} to select the new state X_k at each k.

Algorithm 1 Classical Viterbi's algorithm for bimodal HMM

Require: $Y^{(1)}_{0:T}$, $Y^{(2)}_{0:T}$ and $\pi_0 = \Pr\left(X_0\right)$.
  $j_1 \leftarrow Y^{(1)}_0$ ▷ Initialization
  $j_2 \leftarrow Y^{(2)}_0$
  for $i \in [\![1, N]\!]$ **do**
    $\log\left(\mu(X_0 = i)\right) \leftarrow \log \pi_0 + \log\left(b^{(1)}_{ij_1}\right) + \log\left(b^{(2)}_{ij_2}\right)$
    $\psi_0(X_0 = i) \leftarrow 0$
  end for
  for $k \in [\![1, T]\!]$ **do** ▷ Recursion
    $j_1 \leftarrow Y^{(1)}_k$
    $j_2 \leftarrow Y^{(2)}_k$
    for $i \in [\![1, N]\!]$ **do**
      $\log \mu(X_k = i) \leftarrow \max_n \left\{\log \mu(X_{k-1} = n) + \log\left(a_{ni}\right) + \log\left(b^{(1)}_{ij_1}\right) + \log\left(b^{(2)}_{ij_2}\right)\right\}$
      $\psi_k(X_k = i) \leftarrow \operatorname{argmax}_n \left\{\log \mu(X_{k-1} = n) + \log\left(a_{ni}\right) + \log\left(b^{(1)}_{ij_1}\right) + \log\left(b^{(2)}_{ij_2}\right)\right\}$
    end for
  end for
  Compute $\widehat{X}_T = \operatorname{argmax}_i \log \mu(X_T = i)$. ▷ Termination
  for $k \in [\![0, T-1]\!]$ **do** ▷ Path backtracking (k from $T-1$ to 0)
    $\widehat{X}_k = \psi_{k+1}(\widehat{X}_{k+1})$
  end for
  Return: $\widehat{X}_{0:T}$.

3.4. Modified Viterbi's algorithm

In the proposed modified VA, the termination step and the path backtracking step (*i.e.* the selection rule of the state X_k given in Algo 1) are modified. The initialization step and the recursion step (see Algo 1) remain the same. The transition rule (*i.e.* the triangular law) of the state (see Eq. 4) discourages large variations of the state ($\pm q$ bpm between two successive states). However, the presence of noise can modify the location of the global maximum of $\log \mu(X_k)$ with a distance to the previous state greater than q. To ensure the variation of the state remains small, the selection rule of the state X_k is modified by choosing the closest local maximum with the previous state X_{k-1}. If none of the local maxima has a distance with the previous state X_{k-1} lower than q, then the current state is rejected (X_k = NaN or "not a number") and the next state X_{k+1} is now researched. For successive missing states, the distance to reject the closest local maximum is proportional to the time distance between the current state to evaluate and the latest known state. For example, let $\forall i \in [\![k-(p+1), k-1]\!]$, X_i = NaN and $X_{k-p} \neq$ NaN. The distance between X_k and the closest local maximum at time X_{k-p} has to be lower than pq to be not rejected. In the case, where one of the fHR estimation input gives the mHR, the rule to ignore a state at a given time allows not to shift directly to the mHR if the fHR evolution is lost.

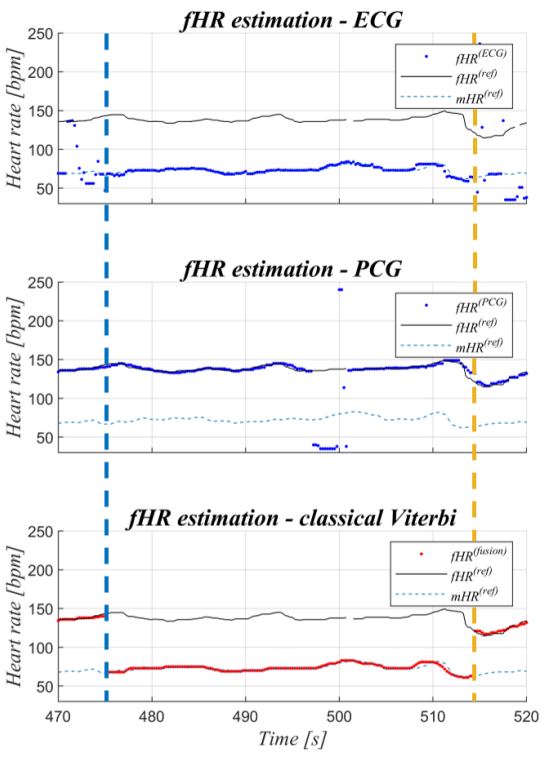
Algo 2 gives the modified VA and Fig. 7 is an illustration of the modified VA functioning.

4. Experiments

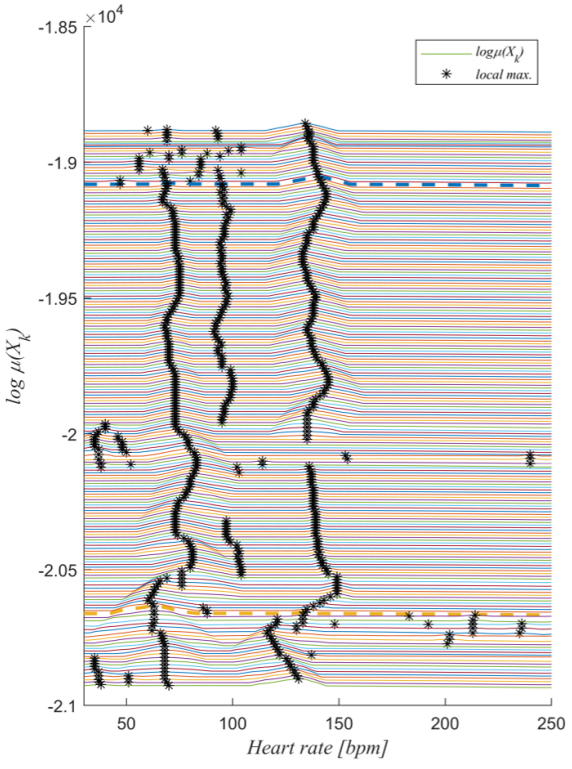
Evaluation is carried out on real data (Sec. 2.1) by comparison of fHR estimation performances between the modified VA and the classical VA.

4.1. Testing configurations

The classical VA and the modified VA are compared on six patients with the bimodal HMM (Fig. 5b). The HMM's parameters are given in the last paragraph of Sec. 3.1. For the six patients, the initial probability law is centered around the expected fHR ($\Pr(X_0) = \mathcal{N}(135, 50)$).



(a) Heart rate estimation, real example.



(b) Vector values ($\log \mu(X_k)$) evolution.

Figure 6: Limit of fHR estimation using the classical VA illustrated on 50s. **(a)** gives $\text{fHR}^{(\text{ecg})}$ on the top, $\text{fHR}^{(\text{pcg})}$ at the middle and $\text{fHR}^{(\text{fusion})}$ at the bottom. $\text{fHR}^{(\text{ref})}$ and $\text{mHR}^{(\text{ref})}$ are, respectively, the reference fHR given by the CTG and the reference mHR estimated from the ECGt. **(b)** gives the vector values $\log \mu(X_k)$ which allows to get $\text{fHR}^{(\text{fusion})}$ in **(a)**. Each line corresponds to $\log \mu(X_k)$ values for all states at a given date k. Because the values of $\log \mu(X_k)$ decrease in time, one can read the evolution of $\log \mu(X_k)$ in time by shifting from the top to the down of the figure. For each instant k, local maxima are displayed on the figure. Dashed lines in **(a)** and **(b)** are synchronized and give $\log \mu(X_k)$ with k corresponding to the location when $\text{fHR}^{(\text{fusion})}$ starts and ends to give the mHR and not the fHR.

For real data, the true fHR and the true mHR are unknown. However, fHR estimation from CTG can be considered as the reference fHR and the reference mHR can be estimated thanks to the detection of mother R peaks on the thoracic ECGt signal. The quality of ECGt allows to have an accurate estimation of the mHR. Results of fusion algorithms $\text{fHR}^{(\text{fusion})}$ are compared to the reference fHR and the reference mother heart rate (mHR) denoted, respectively, $\text{fHR}^{(\text{ref})}$ and $\text{mHR}^{(\text{ref})}$.

4.2. Evaluation criteria

The two first evaluation criteria to compare algorithm are based on outliers estimations. The set of outliers is defined as the set of indices of $\text{fHR}^{(\text{fusion})}$ values that differ from $\text{fHR}^{(\text{ref})}$ by more than 12.5 bpm. The choice of 12.5 bpm is adapted to the clinical criterion for the fHR normal variability (see [31]). The same threshold for the definition of outliers with $\text{mHR}^{(\text{ref})}$ is kept to remain comparable, but without any clinical justification.

Let $k \in [\![0, T]\!]$ be the discrete time vector. The fetal non-outlier's ratio $\overline{out}^{\text{fHR}}$ is defined by:

$$\overline{out}^{\text{fHR}} = \frac{\sum_{k=0}^{T} \delta\left(\left|\text{fHR}_k^{(\text{ref})} - \text{fHR}_k^{(\text{fusion})}\right| < 12.5\right)}{T+1}, \tag{14}$$

Algorithm 2 Modified Viterbi's algorithm for bimodal HMM

Require: q, $Y_{0:T}^{(1)}$, $Y_{0:T}^{(2)}$ and $\pi_0 = \Pr\left(X_0\right)$.
Initialization: identical to Algo 1.
Recursion: identical to Algo 1.
$\widehat{X}_0 \leftarrow \max_i \log \mu(X_0)$. ▷ Termination
$k^* \leftarrow 0$ ▷ Last date for which the state is defined ($\neq$ NaN)
for $k \in [\![1, T]\!]$ **do**
  $\widehat{X}_k^{temp} \leftarrow \text{localmax}(\log \mu(X_k))$ ▷ Search all local max. ($\widehat{X}_k^{temp}$ is a vector).
  if $\min |\widehat{X}_{k^*} - \widehat{X}_k^{temp}| > q \times (k - k^*)$ **then** ▷ If the distance between the closest local max. and the previous state is greater than $\pm q \times (k - k^*)$
    $\widehat{X}_k \leftarrow \text{NaN}$ ▷ State is ignored
  else ▷ If the distance between the closest local max. and the previous state is lower than $\pm q \times (k - k^*)$
    $\widehat{X}_k \leftarrow \text{argmin}_i |\widehat{X}_{k^*} - \widehat{X}_k^{temp}|$
    $k^* \leftarrow k$
  end if
end for
Return: $\widehat{X}_{0:T}$.

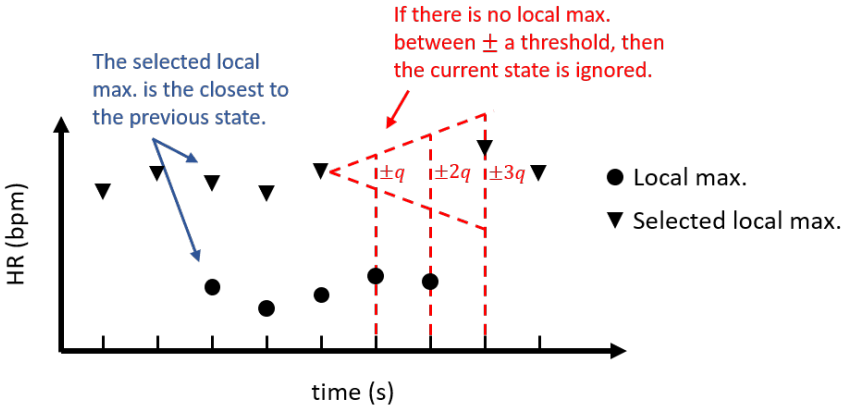


Figure 7: Explanatory diagram of the modified VA. The two new rules are described here. First, when there is more than one local maximum in $\log \mu(X_k)$, then the local maximum corresponding to the closest heart rate to the previous heart rate sample is retained. Second, if there is no local maximum for which the distance to previous heart rate sample is lower than q bpm, then the current state is ignored, *i.e.* $X_k = \text{NaN}$. The threshold q increases each time a state is ignored.

where $*_k$ refers to the value of $*$ at time k and $\delta(x) = 1$ if x is true, 0 else. And, the mother non-outlier's ratio $\overline{out}^{\text{mHR}}$ is equal to:

$$\overline{out}^{\text{mHR}} = \frac{\sum_{k=0}^{T} \delta\left(\left|\text{mHR}_k^{(\text{ref})} - \text{fHR}_k^{(\text{fusion})}\right| < 12.5\right)}{T+1}. \tag{15}$$

It is worth noting that these criteria values are between 0 and 1: the higher $\overline{out}^{\text{fHR}}$, the better, while the lower $\overline{out}^{\text{mHR}}$, the better.

In addition to the outlier's ratio, the signal loss ratio is also computed. For the classical VA, there is no removed data, but for the modified VA an estimation at a given time can be rejected (if Algo 2). The signal loss evaluates the quantity of removed data. For clinical uses, the signal loss should be the lower as possible ($\leq$ 20% to be acceptable

Table 1
Non-outliers ratio using classical VA and modified VA.

$\Pr(X_0)$	Patient number	Classical VA			Modified VA		
		$\overline{out}^{\text{fHR}}$	$\overline{out}^{\text{mHR}}$	signal loss	$\overline{out}^{\text{fHR}}$	$\overline{out}^{\text{mHR}}$	signal loss
$\mathcal{N}(135,50)$	01	**92%**$^{(*)}$	1%	0%	87%	1%	7%
	02	75%	8%	0%	**84%**	1%$^{(**)}$	13%
	03	90%	2%	0%	**93%**	1%	6%
	04	**38%**	36%	0%	23%	45%	20%
	05	**92%**	5%	0%	90%	0%	6%
	06	41%	59%	0%	**96%**	0%	4%
$\mathcal{N}(60,50)$	06	**41%**	59%	0%	0%	81%	6%

$^{(*)}$: Bold texts give best $\overline{out}^{\text{fHR}}$ between the two algorithms.
$^{(**)}$: Underlined texts give best $\overline{out}^{\text{mHR}}$ between the two algorithms.

[1]). The signal loss definition is:

$$\text{signal loss} = \frac{\sum_{k=0}^{T} \delta\left(\text{fHR}_k^{(\text{fusion})} = \text{NaN}\right)}{T+1} \tag{16}$$

where NaN (not a number) refers to a missing value.

4.3. Results

Test results are summarized in Tab. 1, which is separated in two parts. The first one is dedicated to results on six pregnant women with the initial probability law centered around the expected fHR. And, the second part gives the result for patient n°06 with the initial probability law centered around the expected mHR. For each experiment, $\overline{out}^{\text{fHR}}$, $\overline{out}^{\text{mHR}}$ and the signal loss are given for both the classical VA and the modified VA.

Accuracy of fHR estimation It is provided by the criterion $\overline{out}^{\text{fHR}}$. In tests where $\Pr(X_0) = \mathcal{N}(135,50)$, values of $\overline{out}^{\text{fHR}}$ are larger for the modified VA in 50% of cases, resulting to a better estimation of the fHR with the modified VA than with the classical VA. The use of the modified VA never decreases $\overline{out}^{\text{fHR}}$ by more than 5%. The only exception is with patient n°04 where both algorithms fail to provide a high $\overline{out}^{\text{fHR}}$: 38% and 23% for, respectively the classical VA and the modified VA.

Confusion with mHR When the fHR is not correctly estimated, it is interesting to compare the estimated HR with the maternal one to know if there is a confusion with this latter or if the wrong estimation of the fHR comes from another kind of noise. This is evaluated with $\overline{out}^{\text{mHR}}$. The use of the modified VA decreases $\overline{out}^{\text{mHR}}$ for all patients but one. For the classical VA, $\overline{out}^{\text{mHR}}$ varies from 1% to 59%. For the modified VA, $\overline{out}^{\text{mHR}}$ is lower or equal to 1% for all patients except patient n°04 where both algorithms gives high $\overline{out}^{\text{mHR}}$: 36% and 45% for, respectively the classical VA and the modified VA.

Patient n°06 is the most spectacular case where fetal heart rate accuracy, $\overline{out}^{\text{fHR}}$, increases from 41% with the classical VA to 96% with the modified VA. The maternal confusion, $\overline{out}^{\text{mHR}}$, decreases from 59% with the classical VA to 0% with the modified VA. Additionally, to test the influence of the initial probability law, this latter is chosen as $\Pr(X_0) = \mathcal{N}(60,50)$, so that the mean of this probability is now centered about the classical mHR. As one can see in the last line of Tab. 1, the modified VA now provides an heart estimation related to the mHR for 81% of the time while $\overline{out}^{\text{mHR}} = 0\%$ with $\Pr(X_0) = \mathcal{N}(135,50)$. This highlights the importance of initial state, but a rough value of fHR is sufficient for a good quality estimation, as shown previously.

Signal loss The signal loss induced by the modified VA is between 4% and 20%. This result seems, at the first sight, disappointing compared to the classical VA for which the signal loss is always 0%. It has to be noted that a 0% signal

loss does not mean necessary that the fetal heart rate estimation is accurate: it has to be analyzed with regard to the 2 other criteria $\overline{out}^{\text{mHR}}$ and $\overline{out}^{\text{fHR}}$.

Fig. 8 illustrates results of fHR estimations from the two algorithms, for 3 subjects (Patients n°01, 05 and 06). In each figure, the two plots on the left are the input estimations from the ECG modality (top) and the PCG modality (bottom). And the two right plots are the output estimations from the classical VA (top) and the modified VA (bottom). Results are given for the initial probability rules centered around the expected fHR, $\Pr(X_0) = \mathcal{N}(135, 50)$. In Fig. 8a, for patient n°01, the modified VA does not increase $\overline{out}^{\text{fHR}}$ and does not decrease $\overline{out}^{\text{mHR}}$. The signal loss (with the modified VA) is equal to 7%. This is shown with less isolated points in the plot of fHR estimation with the modified VA. Fig. 8b are results of patient n°05 where the modified VA decreases the confusion with the maternal heart rate, $\overline{out}^{\text{mHR}}$, from 5% to 0%. This can be shown around 210-240s and about 520s. Finally, results of patient n°06 are given in Fig. 8c. In this experiment, one can see that the fetal heart rate estimation from the ECG modality follows the mHR. The fetal heart rate estimation with the classical VA switches between the two heart rates (maternal and fetal ones) while the fetal heart rate estimation with the modified VA succeeds to avoid confusion with the maternal heart rate.

5. Discussion

The multimodal analysis, by merging ECG and PCG, is a promising way to have a robust fHR estimation from noisy observations. Our previous work on fusion of estimations using multimodal HMM [26] showed that merging estimations from different modalities helps to get better estimation in many cases. However, if one modality is very noisy, the estimation can fail. In particular, the presence of the mother component in ECGa can make more difficult the estimation of the fHR: indeed, if the mother component is not correctly removed, the mHR might be estimated instead of the fHR. This issue can lead to a misestimation after the fusion of monomodal fHR estimations if mHR is estimated from the ECGa for several seconds. In this paper, the classical VA is modified in order to be adapted to the fHR estimation application and in particular, to avoid the confusion with the mHR. The proposition is to discourage large variations between successive estimated values.

5.1. Advantages of the proposed modified VA

The modified VA succeeds to reduce the confusion with the mHR in almost all cases with the mother non-outlier's ratio $\overline{out}^{\text{mHR}}$ lower or equal to 1%. This improvement is particularly important in the clinical context [2, 3] where confusion between fetal and maternal heart rates can lead to emergency cesarean sections. In addition, the accuracy of the fHR estimation, evaluated with the fetal non-outliers' ratio $\overline{out}^{\text{fHR}}$, is improved for 50% of patients, with the modified VA in comparison to the classical VA. Moreover, the accuracy of the fHR estimation never decreases by more than 5%.

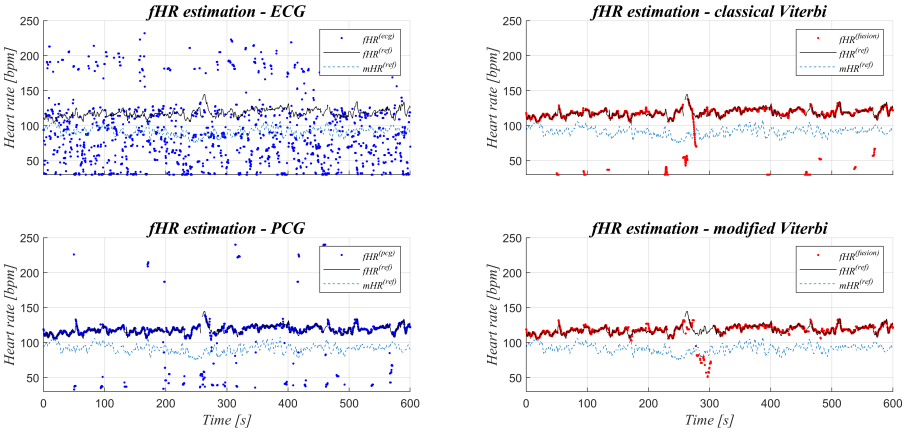
The design of the modified VA allows to ignore some estimations when all possible estimations at a given time are aberrant. The signal loss induced by the new algorithm remains low and always acceptable for clinical uses ($\leq$ 20% for all patients [1]). On the contrary, the classical VA is never subject to signal loss: this means that an estimation is provided at each time even if the estimation is obviously wrong. The philosophy behind the design of the modified VA is to limit outliers to help the interpretation of the estimated fetal heart rate by the clinicians: it is better to provide no estimation than giving a wrong one. Indeed, according to clinicians, having some missing isolated fHR values over fHR estimation has a very few impact when reading the fHR signal, while aberrant values can make the reading more difficult.

During delivery, clinicians control the fetal heart rate in sequences of time duration of a few minutes [1]: e.g. the baseline (the mean level of the most constant fHR) is estimated on segments of 10 minutes while the variability (oscillations in the fHR) is estimated on 1-minute segments. Both the classical VA and the modified VA merge 10 minutes of fHR estimations in less than 1 second on a processor Intel(R) Core(TM) i7-10610U CPU @1.80GHz-2.30GHz with Matlab 2022a. Such results show that these algorithms may be adapted for clinical uses in real time.

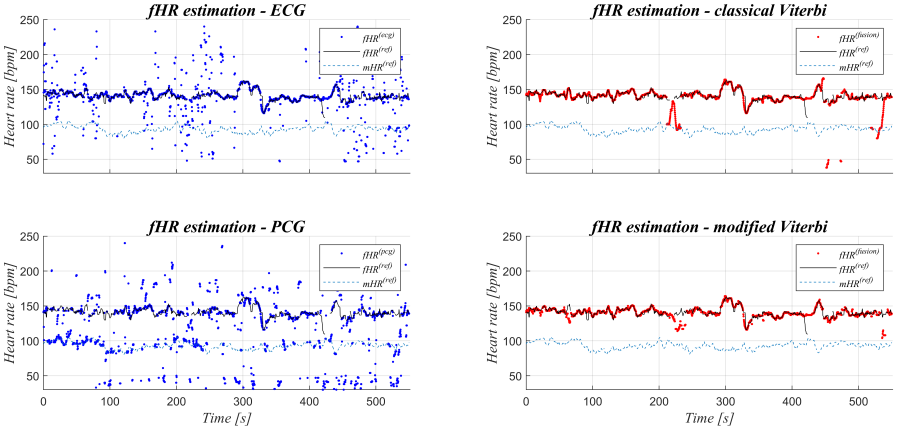
5.2. Limits

The fHR estimation after fusion depends on the input estimations. If input estimations are too noisy (as illustrated with patient n°04), both algorithms fail to provide a good estimation.

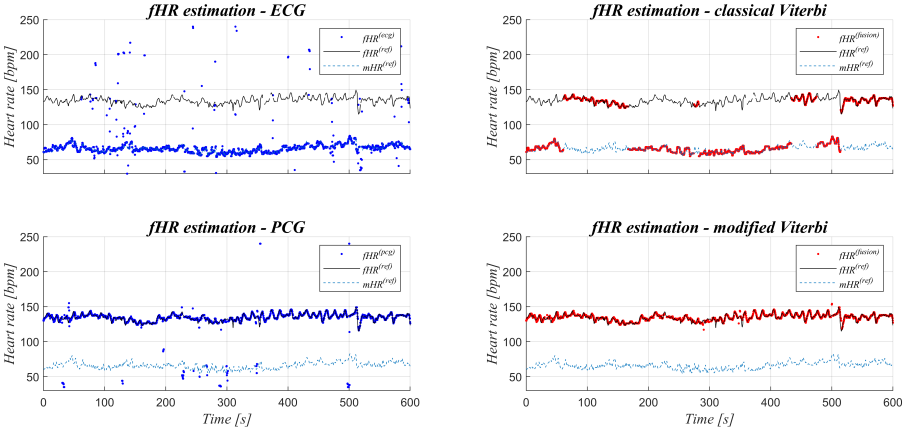
The classical VA provides better results than the modified VA when the initial probability rule is centered around the expected mHR. In that case, the classical VA allows the estimation to switch back to the expected heart rate, while



(a) Patient n°01.



(b) Patient n°05.



(c) Patient n°06.

Figure 8: Result examples on patients n°01, n°05 and n°06. On each figure, the left plots are the input estimations from the modalities separately. The top right plot gives the result of fusion estimation using the classical VA and the bottom right figure gives the result with the modified VA. For each example, the initial probability rule $\Pr(X_0) = \mathcal{N}(135, 50)$ is centered around the expected fHR.

the modified VA prevents the estimation from large variations (Fig. 7), so that the heart rate estimation can stay stack around the maternal heart rate. However, this situation can be simply avoid by choosing an initial rough probability.

In the configuration where one monomodal estimation is centered around the fHR and the other estimation is centered around mHR, if the merged estimation at previous time is centered around mHR, then the modified VA algorithm will force the next estimation to stay around the mHR.

5.3. Perspectives

From a methodological point of view, there are many perspectives to improve the fusion algorithm. Modifying the transition law and/or emission laws to be closer from reality is one of them, as well as adding a control of the prominence of each local maximum to reject local maximum likely associated with noise. Some changes have been tested but not presented in this paper, since most changes only slightly improve some criteria and degrade the other with small percentages.

As previously mentioned, the quality of input estimations has an impact on the fusion output. Changes on the preprocessing step and the monomodal fHR estimation should not be neglected and are always under development [32, 33].

From a clinical point of view, this paper presents a proof of concept to estimate the fHR from two modalities. Tests have been performed on a small number of pregnant women and without considering the health status. Additional signal acquisitions are needed on more patients for a clinical validation of the proposed methodology.

6. Conclusion

In this study, two fetal heart rate estimations, obtained separately from an ECGf and a PCGf, are merged into a new estimation thanks to a modified Viterbi's algorithm. Because of the presence of the maternal heart rate in the fetal heart rate estimation, even the fusion of estimations using the classical Viterbi's algorithm can lead to confusion between the maternal and fetal heart rates. Therefore a modified Viterbi's algorithm has been proposed to limit large variations in the fetal heart rate estimation.

Results on real pregnant women data showed that the proposed modified Viterbi's algorithm succeeds to improve the fetal heart rate estimation by reducing both the presence of outliers estimation and the confusion with the maternal heart rate. Thus, fusion of estimations from two modalities seems a promising approach for more robust fetal heart rate monitoring in clinical situations.

Acknowledgement

This work is supported by the French National Research Agency, as part of the SurFAO project (ANR-17-CE19-0012).

The authors thank the Clinical Investigation Centre for Innovative Technology of Grenoble, for their support to data acquisition.

References

[1] D. Ayres-de Campos, C.Y. Spong, and E. Chandraharan. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. *International Journal of Gynecology & Obstetrics*, 131(1):13–24, 2015.

[2] K. Bhogal and J. Reinhard. Maternal and fetal heart rate confusion during labour. *British Journal of Midwifery*, 18(7):424–428, 2010.

[3] S. Buisson. *Confusion des enregistrements cardiaques fœtal et maternel au cours de l'expulsion : impact sur l'état néonatal et le taux d'extraction instrumentale*. PhD thesis, Faculté de médecine de Grenoble, 2009.

[4] D. Devane, J.G Lalor, S. Daly, W. McGuire, A. Cuthbert, and V. Smith. Cardiotocography versus intermittent auscultation of fetal heart on admission to labour ward for assessment of fetal wellbeing. *Cochrane Database of Systematic Reviews*, 2017.

[5] N. Feinstein, K.L. Torgersen, and J. Atterbury. *Fetal Heart Monitoring: Principles and Practices*. Kendall Hunt, 1993.

[6] J. Westgate, RDF. Keith, JSH. Curnow, EC. Ifeachor, and KR. Greene. Suitability of fetal scalp electrodes for monitoring the fetal electrocardiogram during labour. *Clinical Physics and Physiological Measurement*, 11(4):297, 1990.

[7] V. Kariniemi, J. Ahopelto, PJ. Karp, and TE. Katila. The fetal magnetocardiogram. *J Perinat Med*, 1974.

[8] R. Oweis, E. Abdulhay, A. Alhaddad, F. Sublaban, M. Radwan, and H. Almasaeed. Non-invasive fetal heart rate monitoring techniques: Review article. *Biomedical Science and Engineering*, 2, 01 2014.

[9] N. Dia, J. Fontecave-Jallon, M. Reséndiz, M.C. Faisant, V. Equy, D. Riethmuller, P.Y. Guméry, and B. Rivet. Fetal heart rate estimation by non-invasive single abdominal electrocardiography in real clinical conditions. *Biomedical Signal Processing and Control*, 2021.

[10] R. Kahankova, R. Martinek, R. Jaros, K. Behbehani, A. Matonia, M. Jezewski, and J.A. Behar. A review of signal processing techniques for non-invasive fetal electrocardiography. *IEEE reviews in biomedical engineering*, 13:51–73, 2019.

[11] F. Andreotti, J. Behar, S. Zaunseder, J. Oster, and G.D. Clifford. An open-source framework for stress-testing non-invasive foetal ECG extraction algorithms. *Physiological measurement*, 37(5):627, 2016.

[12] A. Matonia, J. Jezewski, T. Kupka, M. Jezewski, K. Horoba, J. Wrobel, R. Czabanski, and R. Kahankowa. Fetal electrocardiograms, direct and abdominal with reference heartbeat annotations. *Scientific data*, 7(1):1–14, 2020.

[13] N. Dia, J. Fontecave-Jallon, P.Y. Guméry, and B. Rivet. Fetal heart rate estimation from a single phonocardiogram signal using non-negative matrix factorization. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5983–5986. IEEE, 2019.

[14] A. Prashanth Chetlur, S. Ravi, M. Wilfrido Alejandro, and H. Stuart. Trends in fetal monitoring through phonocardiography: Challenges and future directions. *Biomedical Signal Processing and Control*, 33:289–305, 2017.

[15] W. P Holsinger, K.M Kempner, and M.H Miller. A QRS preprocessor based on digital differentiation. *IEEE Transactions on Biomedical Engineering*, BME-18(3):212–217, 1971.

[16] S. Mallat and W.L Hwang. Singularity detection and processing with wavelets. *IEEE transactions on information theory*, 38(2):617–643, 1992.

[17] S. Ismail, I. Siddiqi, and U. Akram. Localization and classification of heart beats in phonocardiography signals—a comprehensive review. *EURASIP Journal on Advances in Signal Processing*, 2018(1):1–27, 2018.

[18] R. Vullings, CHL. Peters, RJ. Sluijter, M. Mischi, SG. Oei, and JWM. Bergmans. Dynamic segmentation and linear prediction for maternal ECG removal in antenatal abdominal recordings. *Physiological measurement*, 30(3):291, 2009.

[19] DA. Ramli, YH. Shiong, and N. Hassan. Blind source separation (BSS) of mixed maternal and fetal electrocardiogram (ECG) signal: A comparative study. *Procedia Computer Science*, 176:582–591, 2020.

[20] R. Sameni, C. Jutten, and MB. Shamsollahi. Multichannel electrocardiogram decomposition using periodic component analysis. *IEEE transactions on biomedical engineering*, 55(8):1935–1940, 2008.

[21] W. Zhong, L. Liao, X. Guo, and G. Wang. Fetal electrocardiography extraction with residual convolutional encoder–decoder networks. *Australasian physical & engineering sciences in medicine*, 42(4):1081–1089, 2019.

[22] R. Sameni and G.D. Clifford. A review of fetal ECG signal processing; issues and promising directions. *The open pacing, electrophysiology & therapy journal*, 3:4, 2010.

[23] S. Noorzadeh, B. Rivet, and P-Y. Guméry. A multi-modal approach using a non-parametric model to extract fetal ecg. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 832–836. IEEE, 2015.

[24] A. Shokouhmand, C. Antoine, B. K Young, and N. Tavassolian. Multi-modal framework for fetal heart rate estimation: Fusion of low-snr ecg and inertial sensors. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 7166–7169. IEEE, 2021.

[25] A. Khandoker, E. Ibrahim, S. Oshio, and Y. Kimura. Validation of beat by beat fetal heart signals acquired from four-channel fetal phonocardiogram with fetal electrocardiogram in healthy late pregnancy. *Scientific reports*, 8(1):1–11, 2018.

[26] R. Souriau, J. Fontecave-Jallon, and B. Rivet. Fusion of estimations from two modalities using the Viterbi's algorithm: application to fetal heart rate monitoring. In *the 29th European Symposium on Artificial Neural Networks (ESANN)*, pages 623–628, 2021.

[27] M-C. Faisant, J. Fontecave-Jallon, B. Genoux, N. Dia, M. Reséndiz, B. Rivet, D. Riethmuller, V. Equy, and P. Hoffmann. Non-invasive fetal monitoring: Fetal Heart Rate multimodal estimation from abdominal electrocardiography and phonocardiography. *Journal of Gynecology Obstetrics and Human Reproduction*, page 102421, 2022.

[28] S. Van Vaerenbergh and I. Santamaría. A comparative study of kernel adaptive filtering algorithms. In *2013 IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, pages 181–186. IEEE, 2013.

[29] S.P. Von Steinburg, A-L. Boulesteix, C. Lederer, S. Grunow, S. Schiermeier, W. Hatzmann, K-T.M. Schneider, and M. Daumer. What is the "normal" fetal heart rate? *PeerJ*, 1:e82, 2013.

[30] L.R Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[31] A. Martin. Fetal heart rate during labour: definitions and interpretation. *Journal de gynecologie, obstetrique et biologie de la reproduction*, 37:S34–45, 2008.

[32] H. Lafaye de Micheaux, M. Reséndiz, B. Rivet, and J. Fontecave-Jallon. Residual convolutional autoencoder combined with a non-negative matrix factorization to estimate fetal heart rate. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 1292–1295. IEEE, 2022.

[33] M. Reséndiz, J. Fontecave-Jallon, and B. Rivet. Hidden markov model in nonnegative matrix factorization for fetal heart rate estimation using physiological priors. *Physiological Measurement*, 2022.