



# Fetal heart rate estimation by non-invasive single abdominal electrocardiography in real clinical conditions

Nafissa Dia, Julie Fontecave-Jallon, Mariel Resendiz, Marie-Caroline Faisant, Veronique Equy, Didier Riethmuller, Pierre-Yves Gumery, Bertrand Rivet

## ► To cite this version:

Nafissa Dia, Julie Fontecave-Jallon, Mariel Resendiz, Marie-Caroline Faisant, Veronique Equy, et al.. Fetal heart rate estimation by non-invasive single abdominal electrocardiography in real clinical conditions. Biomedical Signal Processing and Control, 2022, 71, part 2 (January), pp.103187. 10.1016/j.bspc.2021.103187 . hal-03544789

**HAL Id: hal-03544789**

**<https://hal.science/hal-03544789>**

Submitted on 26 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Fetal heart rate estimation by non-invasive single abdominal electrocardiography in real clinical conditions

Nafissa Dia<sup>a,b</sup>, Julie Fontecave-Jallon<sup>b,\*</sup>, Mariel Resendiz<sup>a,b</sup>, Marie-Caroline Faisant<sup>c</sup>,  
Véronique Equy<sup>c</sup>, Didier Riethmuller<sup>c</sup>, Pierre-Yves Guméry<sup>b</sup>, Bertrand Rivet<sup>a</sup>

<sup>a</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP\*, GIPSA-lab, Grenoble, France

<sup>b</sup>Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP\*, TIMC-IMAG, Grenoble, France

<sup>c</sup>Grenoble Alpes University Hospital, Grenoble, France

---

## Abstract

Fetal heart rate (FHR) is the main feature for monitoring fetal well-being as recommended by the International Federation of Gynecology and Obstetrics (FIGO). The ultrasound based cardiotocography (CTG) is nowadays the world-wide non-invasive clinical reference technique for intrapartum FHR monitoring, as stated by FIGO. However, some limitations of CTG have justified the interest for alternative solutions, based on abdominal electrocardiography (ECG). Their aim is to extract a fetal ECG signal (fECG), from which R peaks are detected to deduce FHR. Nevertheless, most published methods require a large number of abdominal sensors, which is not suitable for an expected use in real clinics.

A novel methodology using a single abdominal ECG derivation is then investigated for FHR estimation in real clinical conditions. After maternal ECG attenuation, a source-filter model is considered to characterize the resulting fetal signal. Based on a non-negative matrix factorization of its spectrogram, the fundamental frequency of the source part is estimated, corresponding to FHR. Validation of the proposed FHR estimation is carried out on a real clinical database.

Compared to a selection of other FHR estimation methodologies, our proposition presents the best performance in terms of reliability and similarity to the reference CTG. Evaluated for 8 subjects, corresponding to a total of 167 minutes of recordings, the mean ratio of outliers is lower than 25%. These good results in real clinical conditions makes the proposed methodology a promising solution for use in routine practice by hospital

---

\*Corresponding author.

Julie Fontecave-Jallon  
Laboratoire TIMC  
Faculté de Médecine, Pavillon Taillefer  
38706 LA TRONCHE Cedex, France

Email addresses: nafissa.dia@univ-grenoble-alpes.fr (Nafissa Dia),  
julie.fontecave@univ-grenoble-alpes.fr (Julie Fontecave-Jallon),  
mariel.resendiz@univ-grenoble-alpes.fr (Mariel Resendiz), MCFaisant@chu-grenoble.fr  
(Marie-Caroline Faisant), VEquy@chu-grenoble.fr (Véronique Equy),  
DRiethmuller@chu-grenoble.fr (Didier Riethmuller),  
pierre-yves.gumery@univ-grenoble-alpes.fr (Pierre-Yves Guméry),  
bertrand.rivet@gipsa-lab.grenoble-inp.fr (Bertrand Rivet)

departments.

*Keywords:* Fetal heart rate (FHR), abdominal electrocardiography, non-negative matrix factorization (NMF)

---

## 1. Introduction

Analyzing the fetal heart rate (FHR) and its variability allows to follow the fetal well-being during labor and birth in clinical setting. Indeed, FHR monitoring brings important information about possible fetal sufferings [1] and is essential in the prevention from serious neonatal abnormalities thanks to quick cares [2]. The non-invasive reference tool in clinical routine is nowadays the cardiotocography (CTG) technique, as required by the International Federation of Gynecology and Obstetrics (usually referred as FIGO) [2]. CTG is based on Doppler ultrasounds and measures simultaneously the FHR and the uterine activity motions. It has been used since the sixties and by 1989 was present at 99% of the pregnancies in France [1]. According to FIGO [2], the CTG external FHR monitoring is the recommended initial method for routine intrapartum monitoring. Although CTG is currently the world-wide reference technique, it has been shown that using this latter contributed to the rise of the number of cesareans and instrumental deliveries [3, 4], especially because it presents frequent confusions between the maternal and the fetal heart rates during labor and maternal pushing efforts [5]. An alternative to external FHR monitoring with CTG is the internal one using scalp electrocardiography (scalp ECG). It consists of screwing an electrode on the fetal scalp to assess the fetal ECG. Scalp ECG is considered to provide an accurate FHR estimation for fetal well-being monitoring [6]. However this method has important drawbacks: it requires water breaking and ruptured membranes, it is highly invasive, carrying possible maternal or fetal infections and it has many established contraindications [2]. All these considerations have then lead to investigate new techniques for FHR estimation combining the non-invasive property of CTG and the reliability and accuracy of scalp ECG. In particular, some studies have shown the interest of substituting CTG by abdominal ECG (aECG) for FHR monitoring [7].

Since the beginning of the 20th century, extraction of fetal information from aECG is possible [8] and in particular, numerous signal processing methodologies of FHR estimation from aECG signals have then been proposed ([9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19] among others). Generally, they first need to extract the fetal ECG (fECG) from aECG signals before detecting the *R* peaks to estimate FHR.

The fECG extraction step is not always easy due to the low power of fetal cardiac beats compared to the maternal ones. To this end, blind source separation [11, 12] is an interesting and effective tool for fECG extraction but this technique most often requires the use of a high number of sensors (typically at least 5 or 6 placed on the thorax and the abdomen). Regarding ergonomic concerns for the mother and the clinicians during labour and delivery, this is therefore not suitable for clinical practice, limiting thus its potential use in real situations. Consequently, usable devices and related algorithms have to minimize the number of abdominal leads for an expected use in real clinics. In this aim, other methodologies that use less sensors have been proposed (e.g. [16, 14]).

For example time-scale analysis as wavelets transform [16] focuses on the ECG remarkable points detection in order to distinguish maternal and fetal ones. Wavelets are then used to reconstruct the fECG from its singular features. In [14], an other approach based on Kalman filtering was proposed to monitor the FHR in a mobile way at home from a single-channel abdominal ECG. A robust version of the extended Kalman filter, namely strong extended Kalman filter, is used for the tracking of the maternal ECG. This tracking allows to remove the maternal component from the aECG and to enhance the fetal *QRS* complexes. The FHR is finally estimated classically by detecting the *R* peaks. Although results show the effectiveness of the algorithm, its validation has been carried out on controlled data or over short duration ( $\approx 10s$ ).

For FHR estimation from fECG signals, several methods have been proposed [17, 18, 19], mainly based on temporal events detection, usually *R* peaks which are more prominent than other ECG waves. The well-known Pan & Tompkins algorithm [20], originally implemented for *R* peaks detection from thoracic ECG signals in adults, has been adapted for fetus [17]. This algorithm consists in first enhancing the *R* peaks by using classical filters (band-pass, derivative, ...) and then in detecting them using decision rules. In [19], authors proposed to robustly detect the *QRS* complexes on the derivative signal using the Hilbert transform. Wavelets transform has been considered in [18] since it has the interesting advantage to well highlight specific features; this allows to detect the specific points of the fECG and then to discriminate the *QRS* complexes from the other waves.

Finally, if many literature approaches have shown satisfactory results, their evaluations were, in most cases, limited on either synthetic data or controlled realistic signals (e.g., with high signal-to-noise ratio (SNR) over short durations). However, in real clinical conditions, signals are highly interfered by different types of noise as drifts, mother motions or electromyogram (EMG) leading then to a difficult detection of characteristic points, such as fetal *R* peaks, even visually.

Regarding the limitations of existing methodologies detailed before and to ensure a clinical use, a new FHR estimation methodology is thus proposed in this paper. Firstly, in relation to the clinical ergonomic constraints, a *single* ECG signal is considered on the abdomen to ensure a good ergonomics for the mother and the clinicians. It is associated to a thoracic ECG signal, already often used in clinics for maternal monitoring. Second, to avoid fetal *R* peaks detection, the proposed algorithm is based on the non-negative matrix decomposition (NMF) well adapted to extract specific features from the spectrogram of physiological signals. Finally, to overcome the limitations of publicly available data, the evaluation is performed on new signals recorded in real clinical conditions.

The paper is organized as follows. First, the signals database considered in the study is described. Then the various steps of the proposed algorithm are explained, as well as the evaluation methodology. Finally, results are presented and discussed.

## 2. Fetal heart rate estimation methodology

### 2.1. Signal acquisitions

A clinical protocol was established and a prospective open label study started in May 2019 in the University Hospital of Grenoble (study N<sup>o</sup> RCB : 2018-A03182-53).



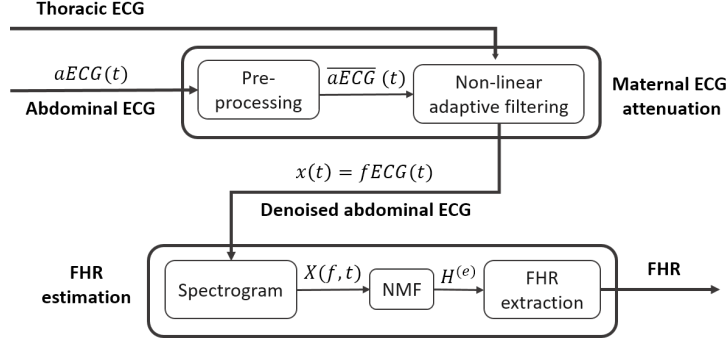


Figure 1: Block diagram of the proposed methodology for FHR estimation

Pregnant women older than 18 years old with a single pregnancy were included among the 9th month, provided none maternal or fetal complication. Data acquisitions were carried out in the obstetrics ward of the university hospital of Grenoble. After signature of informed consent, the volunteer was laying on her back, in a comfortable position, in order to minimize movements and electrical interference.

Signal acquisitions were carried out using a PowerLab data acquisition system (ADInstruments), with sampling frequency at 1 kHz. Each acquisition session consisted in the recording of one thoracic ECG and one abdominal ECG (BioAmp, ADInstruments). For thoracic ECG, bipolar electrodes were placed across the mother heart axis. For abdominal ECG, first the location of the fetal heart was defined by obstetrical ultrasound, and then two abdominal bipolar electrodes were laid on the mother's abdomen, each side of the fetal heart following the fetus' spine.

One reference electrode was also placed on the woman wrist. A simultaneous reference CTG monitoring (Avalon F20/F30, Philips) was recorded, allowing the acquisition of a reference fetal heart rate. Once the sensors were placed, every monitoring session lasted about 30 minutes. Eight volunteer women, between 37(+6) and 40(+4) weeks(+days) of gestation, participated for now in the study, with a mean duration of gestation of 38(+6) weeks and a standard deviation of 1(+0) week.

As said in Section 1, FHR estimation from abdominal ECG is usually carried out according to two main parts: the attenuation of mother ECG (Section 2.2), followed by the step of FHR estimation (Section 2.3). Fig.1 presents the block diagram of the proposed methodology. Each step is detailed in the following subsections.

## 2.2. Maternal ECG attenuation

This first step consists in denoising the abdominal ECG signal ( $aECG(t)$ ) so as to reduce the contribution of maternal ECG and to estimate a fetal signal, which allows us to estimate the fetal heart rate. Even if our primary aim is not to estimate the fetal ECG signal and to analyze its morphology, by sake of simplicity, this denoised signal will be denoted  $fECG(t)$ . A pre-processing step of the raw abdominal ECG consists of (i) the baseline removal using a 10 Hz high-pass filter (Finite Impulse Response filter

of order 100), (ii) the 50 Hz and its harmonics removal using a notch filter at 50 Hz (Infinite Impulse Response filter of order 2) and a 80 Hz low-pass filter (FIR of order 100). To enhance the fetal cardiac beats, a non-linear kernel adaptive filter [21] is then applied on the obtained abdominal ECG,  $\overline{aECG}(t)$ , taking as reference the thoracic ECG ( $tECG(t)$ ). The resulting signal is considered as an estimation of fetal ECG and denoted  $fECG(t)$ . This filter is a non-linear extension of the classic linear adaptive filter, as in [9], using the non-linear square exponential kernel (length scale  $\lambda = 240$ ) to map the reference to the observed signals

$$k(\mathbf{u}(t), \mathbf{u}(t')) = e^{\left(-\frac{\|\mathbf{u}(t) - \mathbf{u}(t')\|_F^2}{2\lambda^2}\right)}, \quad (1)$$

where  $\mathbf{u}(t) = [tECG(t - M + 1), \dots, tECG(t)]^T$  is a time-moving window of length  $M = 900$  samples.

Although the predominance of the maternal component in  $aECG(t)$ , the proposed methodology does not need to perfectly remove the maternal ECG for an effective FHR extraction, contrary to many classical methodologies based on  $R$  peaks detection, as it will be presented in Section 4.

In Fig.2, examples of signals are presented across the different steps of maternal ECG attenuation over 10 s of recording.

### 2.3. NMF-based algorithm of FHR estimation

#### 2.3.1. Temporal modeling of ECG signal and spectrogram

As many physiological signals, ECG signals are non-stationary but quasi-periodic. They can be described as a succession of cardiac beats, each of them corresponding to several cardiac waves, namely P, Q, R, S and T. Therefore one ECG signal can be seen as a time-varying shape of the cardiac beat repeated quasi-periodically. This can be modeled as a source-(time-varying) filter model defined by

$$x(t) = \sum_i \delta(t - \tau_i) * s_i(t), \quad (2)$$

where  $\delta(t)$  is the Dirac delta function,  $\tau_i$  are the cardiac beats instants and  $s_i(t)$  materializes the  $i$ th cardiac beat pattern that can differ one from others. The filter input is  $e(t) = \sum_i \delta(t - \tau_i)$  and the time-dependent impulse response of the filter is  $s_i(t)$ .

This temporal modeling of ECG signal is also valid for fetal ECG. Note that a similar type of excitation-filter modeling has also been investigated [22, 23] for phonocardiographic (PCG) signals, that share the same temporal properties with ECG.

Commonly applied to physiological signals for their analysis, the spectrogram, (*i.e.* the time-frequency representation based on the short-time Fourier transform (STFT)) allows to stand out specific features with a certain spectral content. Its application to the temporal model defined in (2) leads to

$$\forall(f, t), \quad X(f, t) = X^{(e)}(f, t)X^{(\varphi)}(f, t), \quad (3)$$

where  $X^{(e)}(f, t)$  (resp.  $X^{(\varphi)}(f, t)$ ) is the modulus of the STFT of the excitation (resp. the filter). This can be rewritten as

$$\mathbf{X} = \mathbf{X}^{(e)} \odot \mathbf{X}^{(\varphi)}, \quad (4)$$

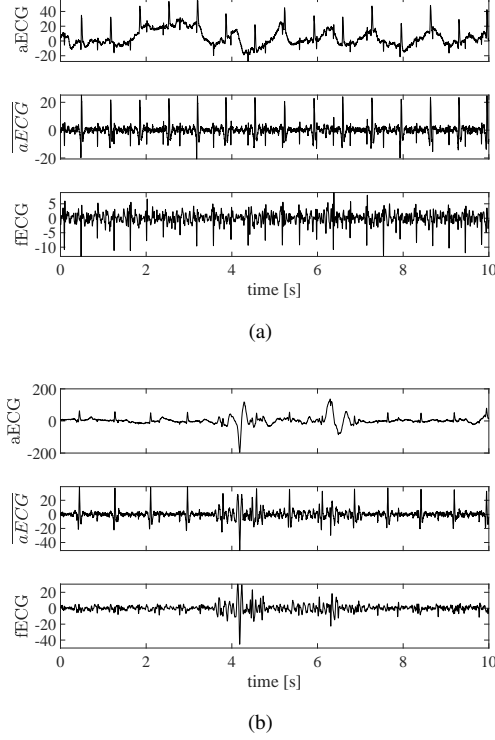


Figure 2: Temporal ECG signals for two different volunteers. From top to bottom : raw abdominal ECG,  $aECG(t)$ , abdominal ECG after pre-processing,  $\overline{aECG}(t)$ , and estimated fetal ECG after non-linear adaptive filtering,  $fECG(t)$ .

where the  $(f, t)$ -th entry of  $\mathbf{X}$  is  $X(f, t)$ ,  $\mathbf{X} \in \mathbb{R}_+^{F \times N}$ ,  $\mathbf{X}^{(e)} \in \mathbb{R}_+^{F \times N}$  and  $\mathbf{X}^{(\varphi)} \in \mathbb{R}_+^{F \times N}$ , with  $F$  the number of frequency bins and  $N$  the number of time windows.  $\odot$  is the Hadamard product (*i.e.* the element-wise multiplication) and  $\mathbb{R}_+$  is the non-negative real numbers set.

Examples of estimated fetal ECG spectrograms from two volunteers are shown in Fig.3 using a 4s STFT window size to gather several heart beats. The harmonic structure is noticed as predicted, due to quasi-periodicity of cardiac signals. A fluctuated fundamental frequency, around 140 bpm for Fig.3a) and 155 bpm for Fig.3b), and its harmonics are well highlighted. Even when the signal is disturbed by noise (obviously between 6 and 12 min for Fig.3a)), the trend of the fetal rhythm remains visible.

### 2.3.2. Harmonic estimation using NMF

The aim is to estimate the excitation component  $\mathbf{X}^{(e)}$  [22, 23] which contains the harmonic structure of the fECG spectrogram as the STFT of  $\sum_i \delta(t - \tau_i)$ . The NMF [24, 25] is a matrix decomposition methodology allowing to factorize a positive matrix  $\mathbf{A}$

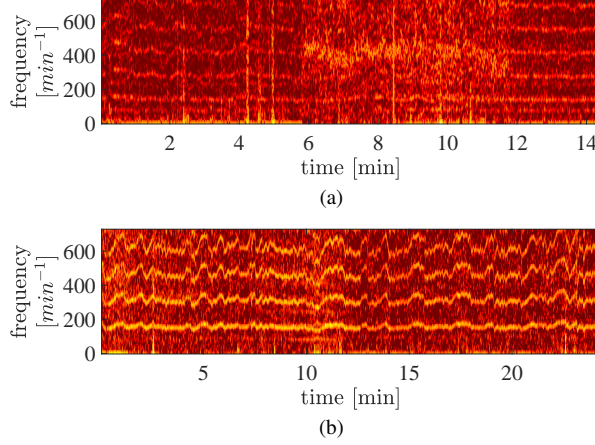


Figure 3: Spectrogram of estimated fetal ECG,  $fECG(t)$  ( $F_s = 1kHz$ , window = 4s, shift = 32ms, zero-padding ratio = 4) for two different volunteers.

into a product of positive matrices  $\mathbf{W}$  and  $\mathbf{H}$  of lower rank than  $\mathbf{A}$  so that  $\mathbf{A} \simeq \mathbf{WH}$ . When NMF is applied to  $\mathbf{X}$  from (4), the following expression is obtained

$$\mathbf{X} \simeq \mathbf{V} = (\mathbf{W}^{(e)}\mathbf{H}^{(e)}) \odot (\mathbf{W}^{(\phi)}\mathbf{H}^{(\phi)}), \quad (5)$$

where  $\mathbf{W}^{(e)}$  and  $\mathbf{W}^{(\phi)}$  are the frequency patterns of respectively the source and the filter,  $\mathbf{H}^{(e)}$  and  $\mathbf{H}^{(\phi)}$  are their related temporal amplitudes.

Matrices  $\mathbf{H}^{(e)}$ ,  $\mathbf{W}^{(\phi)}$  and  $\mathbf{H}^{(\phi)}$  are estimated by minimizing the criterion

$$C(\mathbf{H}^{(e)}, \mathbf{W}^{(\phi)}, \mathbf{H}^{(\phi)}) = \frac{1}{2} \|\mathbf{X} - \mathbf{V}\|_F^2 + \gamma L(\mathbf{W}^{(\phi)}, \mathbf{H}^{(e)}, \mathbf{W}^{(\phi)}), \quad (6)$$

where  $\|\cdot\|_F$  is the Frobenius norm and

$$L(\mathbf{W}^{(\phi)}, \mathbf{H}^{(e)}, \mathbf{H}^{(\phi)}) = \sum_{k_e, k_\phi, n} (\mathbf{H}_{k_e, n}^{(e)})^2 (\mathbf{H}_{k_\phi, n}^{(\phi)})^2 \sum_{f=2}^F \left( \mathbf{W}_{f, k_\phi}^{(\phi)} - \mathbf{W}_{(f-2), k_\phi}^{(\phi)} \right)^2$$

is a smoothness penalization term added to avoid some ambiguities, as detailed in [22].

To this end, multiplicative updates based on a majoration-minimization (MM) algorithm [26] is used [22]. It is worth noting that, in our proposition, the frequency dictionary  $\mathbf{W}^{(e)}$  is chosen fixed and is modeled as a set of Dirac combs. It contains a large range of cardiac frequencies from 30 to 240 bpm (beats per minute) with 1 bpm of precision, covering the possible fetal heart frequencies. Therefore, the matrix  $\mathbf{H}^{(e)}$  has the role to consistently select the right cardiac frequency in  $\mathbf{W}^{(e)}$  at each instant making the matrix  $\mathbf{H}^{(e)}$  to contain the temporal evolution of FHR estimation. The filter part  $\mathbf{W}^{(\phi)}\mathbf{H}^{(\phi)}$  corresponds to the modulated envelop. Matrix  $\mathbf{H}^{(e)}$  is initialized with strictly positive values near 0, and matrices  $\mathbf{W}^{(\phi)}$  and  $\mathbf{H}^{(\phi)}$  are initialized around 1.

One example of  $\mathbf{H}^{(e)}$  estimation for one volunteer is shown in the top plot of Fig. 4, on which the time-varying cardiac frequency is clearly visible and should be extracted, as it will be explained in the following section.

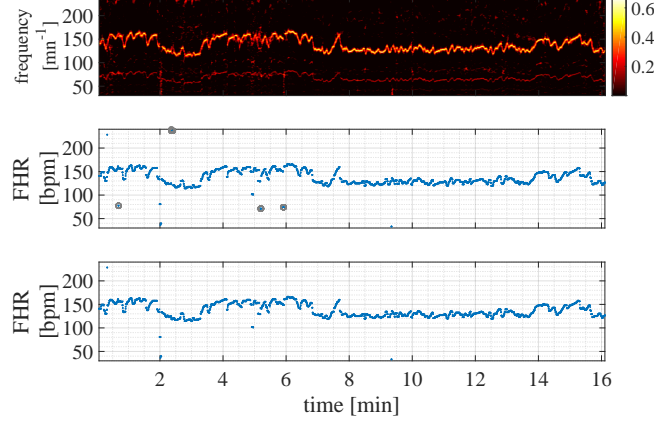


Figure 4: FHR estimation for volunteer  $V_2$ . From top to bottom : matrix  $\mathbf{H}^{(e)}$  estimated from the proposed NMF-based algorithm (represented as an image, where each element of  $\mathbf{H}^{(e)}$  is specified by a color), extracted FHR using  $\mathbf{H}^{(e)}$  (7) and estimated FHR after post-processing step.

### 2.3.3. FHR extraction

Considering that  $\mathbf{H}^{(e)}$  carries all information about FHR, this latter is then estimated by taking the maximum of amplitude in each column of  $\mathbf{H}^{(e)}$ .

$$FHR(t) = \nu \left( \arg \max_k H_{k,t}^{(e)} \right), \quad (7)$$

where  $H_{k,t}^{(e)}$  is the  $(k,t)$ -th entry of  $\mathbf{H}^{(e)}$  and  $\nu \in \mathbb{R}^{K_e}$  is the vector of the  $K_e$  fundamental frequencies of each Dirac comb of  $\mathbf{W}^{(e)}$ , *i.e.* cardiac frequency between 30 bpm and 240 bpm. However, detecting the maximum of amplitude of  $\mathbf{H}^{(e)}$  may sometimes induce the detection of sub-harmonics or harmonics of the fundamental frequency, leading to wrong FHR estimations. A post-processing step is proposed to detect too large FHR derivatives which can be associated to multiple or sub-multiple FHR estimations as detailed in [22]. In such cases, these detected FHR estimations are replaced by their sub-multiples or multiples, accordingly. The result of this process is shown for one volunteer in Fig.4. The circled samples correspond to detected wrong values due to (sub-)harmonics detection which are then corrected by the post-processing step.

## 3. Evaluation methodology

For the validation of the proposed NMF methodology for FHR estimation, numerical assessments of FHR estimations are realized related to FHR estimated from cardiotocography, noted  $FHR_{CTG}$  and considered as reference. All FHR estimations are resampled every 250 ms (*i.e.* sampling frequency of 4 Hz) to match the usual sampling period of CTG devices [27].

The proposed NMF methodology is compared to previous methods of the literature introduced in Section 3.1 based on quantitative criteria defined in Section 3.2.

### 3.1. Comparative methodologies

Results of our NMF-based proposition, referred as *NMF*, are compared to those obtained from three comparative methodologies [17, 18, 19], referenced as *PT*, *Wave*, *Hilb*, respectively. Among the many methods proposed for FHR estimation, these three approaches have first been selected, since they only require a single abdominal sensor, in order to allow a fair comparison with our proposed solution. Moreover, the purpose of this choice is also to select a panel of various algorithmic principles. Indeed, [17] adapts the well-known Pan & Tompkins algorithm [20] for fetal ECG characteristics. [19], based on a Hilbert transform, allows a R-peaks detection robust to motion and noise artifacts, and baseline drift. Finally, [18] proposes a wavelet transform allowing an accurate detection of various QRS morphologies.

These comparative methodologies are applied to the estimated fetal ECG,  $fECG(t)$ , obtained after the pre-processing and non-linear adaptive filtering of abdominal ECG signal,  $aECG(t)$ , similarly as our proposed NMF-methodology (Fig. 1). FHR estimations have then been carried out on the proposed signals database with the 4 methodologies and compared to  $FHR_{CTG}$ . It is worth noting that  $FHR_{meth}$  refers to the FHR estimations based on methodology *meth* where  $meth \in \{NMF, PT, Hilb, Wave\}$ .

### 3.2. Quantitative criteria

The following criteria are considered for quantitative evaluation of the proposed NMF methodology as well as the comparative methodologies.

#### 3.2.1. Outliers

To measure the reliability of the estimated FHR compared to the FHR provided by the CTG reference (when available, see Section 4.2), the set  $\mathcal{O}$  of outliers is defined as the set of indexes of estimated  $FHR_{meth}$  values that differ from  $FHR_{CTG}$  by more than 12.5 bpm. This confidence margin to define outliers has been chosen according to the clinical fetal heart rate variability, which is judged normal when comprising peak-to-peak between 6 and 25 bpm [6].

$$\mathcal{O}_{meth} = \{i \mid |FHR_{meth}(i) - FHR_{CTG}(i)| \geq 12.5\}.$$

One can then define the set of non-outliers  $\overline{\mathcal{O}}_{meth}$  as the complementary set of  $\mathcal{O}_{meth}$ :

$$\overline{\mathcal{O}}_{meth} = \{1, \dots, N\} \setminus \mathcal{O}_{meth},$$

where  $N$  is the total number of FHR estimations given by CTG. Finally, to quantify the number of outliers, the ratio of outlier (*RO*) is defined as

$$RO_{meth} = \frac{\text{card}(\mathcal{O}_{meth})}{N}, \quad (8)$$

where  $\text{card}(\cdot)$  is the cardinality of the input set.

Good reliability between FHR estimation and FHR reference is characterized by a small ratio of outliers (*RO* near to 0%) or by  $1 - RO$  near to 100% as depicted in Table 2.

### 3.2.2. Median deviation with reference (MD)

It corresponds to the median of the differences between estimated  $FHR_{meth}$  and reference  $FHR_{CTG}$

$$MD_{meth} = \text{med}(FHR_{meth} - FHR_{CTG}), \quad (9)$$

where  $\text{med}(\cdot)$  is the median. This is computed over the recording duration and expressed in bpm. A small value near to 0 bpm is expected.

### 3.2.3. Pearson correlation coefficient (R)

It quantifies the similarity between  $FHR_{CTG}$  and  $FHR_{meth}$  after removal of outliers. It is computed as

$$R_{meth} = \frac{\sum_{i \in \bar{\mathcal{O}}_{meth}} \widetilde{FHR}_{meth}(i) \widetilde{FHR}_{CTG}(i)}{\sqrt{\sum_{i \in \bar{\mathcal{O}}_{meth}} \widetilde{FHR}_{meth}^2(i)} \sqrt{\sum_{i \in \bar{\mathcal{O}}_{meth}} \widetilde{FHR}_{CTG}^2(i)}}, \quad (10)$$

where  $\widetilde{FHR}(i) = FHR(i) - \overline{FHR}$  and  $\overline{FHR}$  is the mean value of  $FHR$  computed as

$$\overline{FHR} = \frac{1}{\text{card}(\bar{\mathcal{O}}_{meth})} \sum_{i \in \bar{\mathcal{O}}_{meth}} FHR(i).$$

The nearer to 100% the  $R$  value is, the better is the similarity between FHR estimations and FHR reference given by CTG.

## 4. Results

In the following section, qualitative illustrations are first presented to show the behaviour of the proposed FHR estimation. Then quantitative evaluations of the proposed methodology are carried out, based on the CTG reference and comparative methodologies of FHR estimation.

### 4.1. Qualitative observations of FHR estimation

FHR estimation using NMF are first visually analyzed, starting from the algorithm results for one volunteer. In Fig.5, the fundamental frequency is stood out and fluctuated around an average frequency of 160 bpm which is in the range for a fetus. The figure also highlights for different moments of estimation (grey rectangles) the corresponding temporal fetal ECG. In the signal portion a), fetal cardiac beats are roughly recognizable in fECG signal, however only a few values are badly estimated on FHR estimation, in b) the signal is mostly noisy with quite impossible  $R$  peaks visualization, leading to a higher number of false FHR estimations and in c) the fetal heart beats are easily identifiable with a very good FHR estimation. This example allows to show the effectiveness of our methodology when fetal cardiac beats are easily detectable but also when the signal is hard with noise so that classical algorithms of  $R$  peaks detection may fail. However, when the signal is flooded in noise, the proposed methodology has difficulties to well estimate FHR.

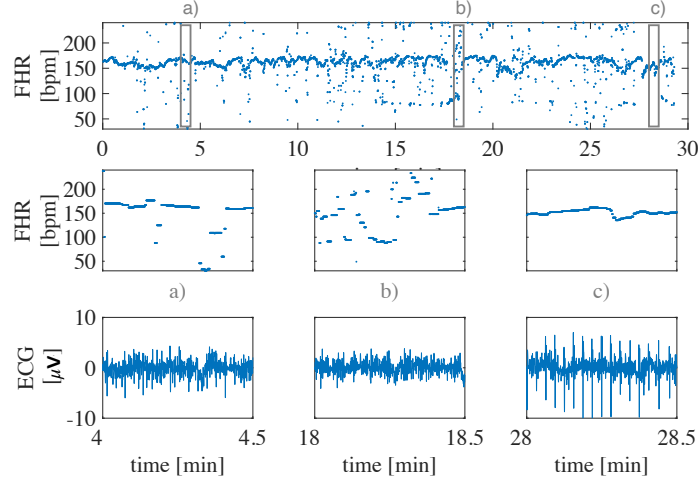


Figure 5: Qualitative analysis of FHR estimation for volunteer  $V_5$ . FHR estimation from our proposed algorithm (top row). Illustration of estimation quality (second row) with the corresponding temporal fECG (third row): a) with fetal cardiac beats hardly identifiable, b) with important noise and c) with fetal cardiac beats easily recognizable.

#### 4.2. Quality of CTG reference

During acquisitions, CTG was recorded (at 4 Hz of sampling frequency) for FHR reference as the non-invasive technique for clinical routine [2]. Generally, when the signal loss ratio (SLR) is greater than 20%, the CTG is not exploitable according to the International Federation of Gynecology and Obstetrics [28]. In Table 1 are mentioned SLR of each CTG recording from the 8 different volunteers. SLR is null for 2 subjects over 8, however all CTG outlines are exploitable since the SLR remains smaller than 20% and CTG can then be used for all subjects for the comparative assessments. For each subject, the recording duration (RD) is then considered as the usable duration of the electrocardiographic signals, once signal losses on CTG have been excluded. The values of RD for all subjects are reported on Table 2. These are comprised between 12 and 30 minutes.

Table 1: Quality of the CTG reference. Signal loss ratio (SLR) for the 8 volunteers.

Volunteer	$V_1$	$V_2$	$V_3$	$V_4$	$V_5$	$V_6$	$V_7$	$V_8$
SLR [%]	0	0	13	10	4	3	15	3

#### 4.3. Comparison with other methodologies

##### 4.3.1. Qualitative comparison

Fig. 6 illustrates, for one volunteer, FHR estimations from the 4 methods superimposed to the reference  $FHR_{CTG}$ .  $FHR_{NMF}$  from the algorithm proposed in this article is visually comparable to  $FHR_{CTG}$  for this volunteer all along the 15 minutes



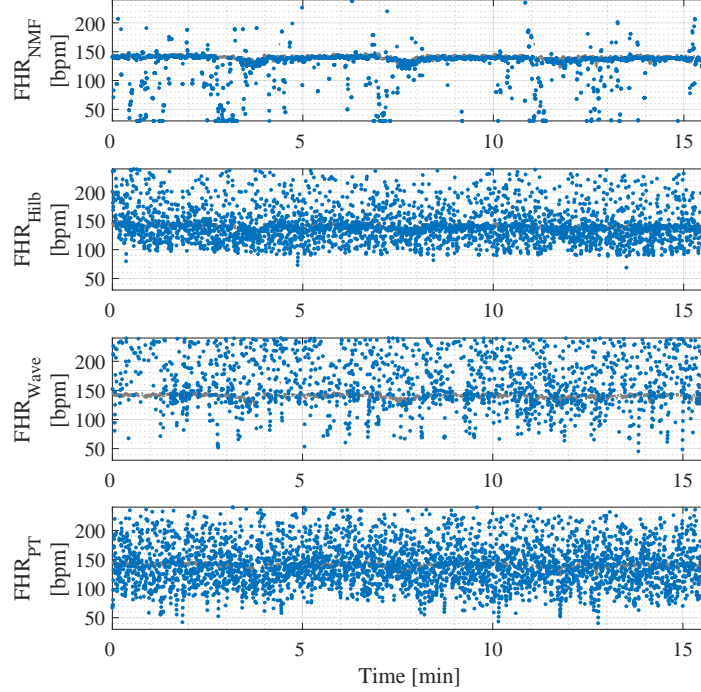


Figure 6: Comparison of FHR estimations (in blue points) from NMF-based algorithm and comparative algorithms with reference  $FHR_{CTG}$  (in grey line) for volunteer  $V_4$ . From top to bottom:  $FHR_{NMF}$ ,  $FHR_{Hilb}$  [19],  $FHR_{Wave}$  [18] and  $FHR_{PT}$  [17].

of recording. There are some values badly estimated but these do not prevent to follow FHR across time.

The comparative methods, as already said in Section 1, based their implementation on time events localization and detection. They were first implemented and intended for adults ECG and they are not always directly applicable to fetal ECG since some parameters have to be adjusted. Algorithms based on Hilbert and Wavelets transforms [18, 19] are less physiologically constrained than Pan & Tompkins algorithm [17]. The advantage of the Hilbert-based methodology is that the threshold is fixed only on the current window without impacts on the remain of the signal if presence of noise and this can be seen in the second row of Fig. 6.  $FHR_{Hilb}$  is visually comparable to  $FHR_{CTG}$  at certain instants.  $FHR_{Wave}$  is less accurate and barely comparable to  $FHR_{CTG}$  because of a threshold based on the maximum amplitude of the whole processed signal. Pan & Tompkins algorithm proved its effectiveness on adults ECG signals with a low SNR but it is difficult to adapt on fetal ECG extracted from abdominal ECG although physiological values are reviewed. The main issue is that

the algorithm is not able to adapt to strong variations in the signal despite adaptive thresholds. This leads to a result of  $FHR_{PT}$  unusable.

#### 4.3.2. Quantitative evaluation of FHR estimation

Quantitative criteria, as introduced in Section 3.2, have been computed for each volunteer and for each considered methodology, the proposed NMF solution and 3 comparative methodologies (Hilbert, Wavelets and Pan & Tompkins). All values are depicted in Table 2.

Recordings last between 12 and 30 min, with a median value of recording duration  $RD$  equal to 20 min and a mean value of 20.9 min. Four recordings last more than 20 min. The total of recording duration for the eight volunteers is 167 min.

Considering the number of  $FHR_{meth}$  distant more than  $\pm 12.5$  bpm from  $FHR_{CTG}$  along  $RD$ , outliers ratio are computed and  $1 - RO$  (8) is displayed in Table 2. From all volunteers, our proposed NMF methodology presents the best reliability between FHR estimations and reference. For 6 volunteers over 8, more than 75 % of  $FHR_{meth}$  is reliable. This is confirmed by the median value: 78.2% against 45.5% for Hilbert-based methodology, 15% for the Wavelets-based algorithm and 18.5% for Pan & Tompkins algorithm.

While considering the differences between  $FHR_{meth}$  and  $FHR_{CTG}$ , the median deviations with reference  $MD$  (9) depicted in Table 2 also show that our proposed methodology allows the best FHR estimations, since  $MD$  values for NMF are lower to other methodologies for most subjects. Indeed,  $MD$  median value for the 8 volunteers is  $-1.2$  bpm with our NMF approach, which is quite similar to  $MD$  from the Hilbert-based methodology ( $-2.5$  bpm) but smaller than  $MD$  from Pan & Tompkins algorithm and Wavelets approach (respectively  $-7.1$  bpm and  $59.7$  bpm).

Correlation coefficient  $R$  (10) between  $FHR_{NMF}$  and  $FHR_{CTG}$  is higher or equal to  $R$  from other methodologies for all volunteers.  $R$  values from NMF are included between 58 and 96 %, against 49 and 95 % for Hilbert, 35 and 95 % for Wavelets and 45 and 92 % for Pan & Tompkins. Note that  $R$  could not be computed for  $FHR_{PT}$  for  $V_6$  due to inoperable estimation ( $1 - RO = 0$ ).  $R$  median value (resp. 93.0 %, 85.5 %, 84.0 % and 77.0 %) for respectively NMF, Hilbert, Wavelets and Pan & Tompkins are close but slightly better for NMF, highlighting a better similarity of  $FHR_{NMF}$  with CTG reference than other methods.

Globally, the proposed NMF algorithm for FHR estimation shows the best performances in terms of reliability and similarity to the reference CTG. The results for median deviation to the reference  $MD$  and correlation coefficient  $R$  are close and comparable to those brought by the Hilbert-based methodology. However, the reliability between estimated FHR and  $FHR_{CTG}$  is weaker for Hilbert, regarding  $1 - RO$  values, leading to a lower global performance for this comparative methodology. While considering Wavelets-based methodology or Pan & Tompkins algorithm, the results provided are systematically worse than the 2 previous solutions.

Figure 7 shows the results for one subject ( $V_5$ ) of FHR estimation  $FHR_{NMF}$  from our proposed algorithm, superimposed on reference  $FHR_{CTG}$ . The outliers are clearly visible. Indeed, they correspond to points that deviate from  $FHR_{CTG}$  by more than

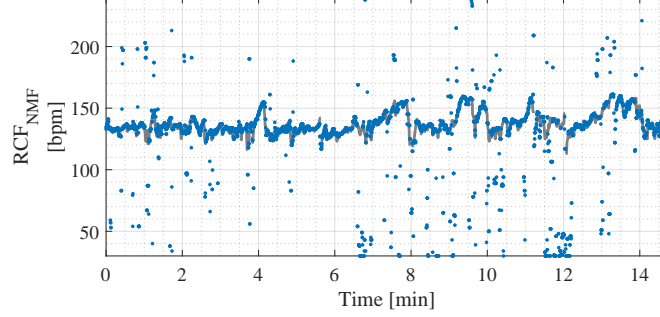


Figure 7:  $FHR_{NMF}$  estimations superimposed on reference  $FHR_{CTG}$  for volunteer V5

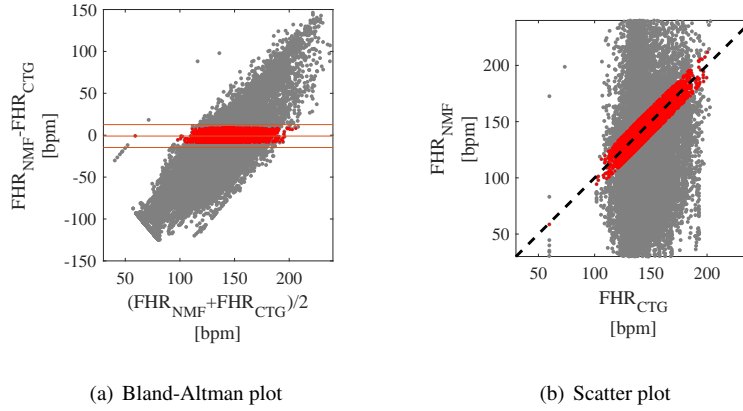


Figure 8: Overall FHR estimations from proposed NMF-methodology for the 8 volunteers. (a) Bland-Altman plot of the whole 8 estimation sets including outliers (in grey) and without outliers (in red), the horizontal red lines correspond to the mean value of the differences and the limits of agreement without outliers, (b) scatter plot of  $FHR_{NMF}$  and  $FHR_{CTG}$  with outliers (in grey) and without outliers (in red), the dashed black line corresponds to  $y = x$ .

12.5 bpm and this plot confirms the numerical value of outliers reported in Tab. 2. One can find in appendix (for more visibility) similar plots for all 8 subjects.

To illustrate the overall performance of the proposed FHR estimation algorithm, we consider pooling the results from all 8 individual subjects into one Bland-Altman plot and one scatter plot, as shown in Fig. 8. The Bland-Altman plot [29] displays the difference  $\Delta FHR$  between the  $FHR_{NMF}$  and the  $FHR_{CTG}$  according to their average value  $((FHR_{NMF} + FHR_{CTG})/2)$ . Outliers are points with a large  $|\Delta FHR|$  corresponding thus to points mainly distributed along the line with a slope equal to 2 (*i.e.* the points on the diagonal line). Moreover, on this Bland-Altman plot, the mean of  $\Delta FHR$  without outliers, as well as the limits of agreement (mean difference  $\pm 1.96$  standard

deviation of the difference) are also reported, as the horizontal straight lines. All points without outliers (plotted in red) are contained inside the interval defined by the limits of agreement, showing that the proposed methodology for FHR estimation is interchangeable with CTG estimation. As reported in Tab. 2, the values of the median deviation with reference  $MD$  are negative showing that the  $FHR_{NMF}$  is underestimated compared to the reference FHR ( $FHR_{CTG}$ ) for all volunteers except  $V_5$  and  $V_6$  for whom  $MD_{NMF}$  is zero. The individual Bland-Altman plots for all 8 subjects are also reported in appendix.

Finally, Figure 8(b) shows the scatter plot between  $FHR_{NMF}$  and  $FHR_{CTG}$  with (in grey) and without outliers (in red). The rectangular shape of the scatter plot with outliers (grey points) shows (i) that the FHR estimation errors appear for the full range of FHR values (100 to 200 bpm) and (ii) that these estimated FHR values are not correlated with the reference FHR as shown by the vertical orientation of the rectangular shape. While considering the scatter plot after removal of outliers (red points), we highlight the good behaviour of the FHR estimated by NMF compared to the reference one ( $R = 0.9$ ). The individual performance for the 8 subjects are available graphically in appendix. Moreover Tab. 2 reports Pearson correlation coefficients without outliers  $R \geq 0.8$  for all volunteers but  $V_4$ . For this subject, the points on the scatter plot remain close to the line  $y = x$  but with a small dispersion (due to a quite flat FHR along time). One can conclude that the FHR estimation is better than suggested by the Pearson correlation coefficient value.

## 5. Discussion and conclusion

The proposed methodology based on NMF from abdominal ECG for FHR estimation outperforms the comparative methods as shown in the results section. Indeed, the NMF-based algorithm estimates directly the FHR contrary to more classical methods that are based on  $R$  peaks detection. This difference is fundamental since it can be difficult to detect accurately the fetal  $R$  peaks due to the very low SNR in clinical conditions (long recording, large artifacts, variations of the baseline, etc).

The proposed method has also the great advantage to require a single abdominal ECG sensor, compared to many methods of literature using several abdominal sensors. Indeed, from clinical perspectives, it is very important to limit the number of abdominal sensors during the labor and the delivery, since the abdomen should be free to allow surgery if necessary. In the proposed solution, an extra thoracic ECG sensor is considered to attenuate the maternal ECG contribution on abdominal ECG signal. But, one can imagine getting rid of this additional sensor, by considering other ways of maternal ECG attenuation than adaptive filtering. A first attempt of a simple template subtraction [30] on abdominal ECG has provided very similar performance results. Indeed, the comparison for the eight volunteers between the template subtraction and the non-linear adaptive filtering respectively leads to  $1 - RO = 81 \pm 6\%$  and to  $1 - RO = 79 \pm 5\%$ , highlighting no significant difference. However, adding an extra thoracic ECG sensor is not a clinical limitation at all, since maternal heart rate monitoring is already done in delivery rooms.

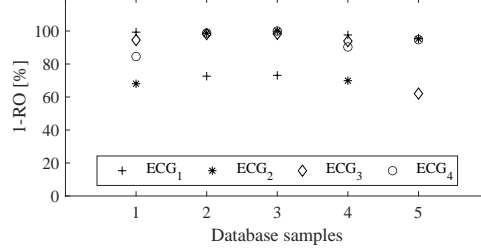


Figure 9: Consistency of FHR estimation from NMF-proposed methodology applied on the five samples of “Abdominal and Direct Fetal ECG Database” [31, 32]. For each sample,  $1 - RO$  is computed for FHR separately estimated from each of the 4 abdominal ECG signals.

In addition to the reported results obtained from the recorded database, the proposed method for estimating the FHR was also applied to the publicly available database “Abdominal and Direct Fetal ECG Database” (also known as Silesia database) [31, 32] for comparison purposes. This database consists of 5 segments of 5 minutes composed of 4 abdominal recordings. A spiral electrode attached to the fetal head provides a reference of the fECG, that is used to extract the reference FHR based on R peaks detection (instants publicly available on Silesia database). Our proposed FHR estimation method was applied separately on each of the 4 abdominal signals, *i.e.* no spatial filtering as independent component analysis was applied as pre-processing to follow the proposed methodology. The maternal ECG attenuation was performed by a simple template subtraction as discussed in the previous paragraph. The FHR quality estimation is reported on Figure 9 by the ratio of outliers ( $RO$ ): for each of the 5 database samples, at least one of the abdominal signal leads to more than 95% of good FHR estimations, for which the median deviations with reference (MD) are -0.05, -0.06, -0.02, -0.04 and -0.02 and the correlation coefficients 99%, 99%, 98%, 99% and 99%, highlighting the effectiveness of the proposed methodology. Among the publicly available database, this is the only one that comes close to the clinical conditions: long enough for allowing a clinical diagnostic, avoiding too short signals (about 1 minute of signal). Indeed, as already mentioned and as reported in the FIGO guidelines [2], a duration longer than 10 minutes is necessary for FHR recordings interpretation, since most clinical criteria definitions are based on 10 minutes time periods (FHR baseline, tachycardia, bradycardia, accelerations, deceleration...). Similar to the reported results obtained on our database, the Bland-Altman plot and the scatter plot have been computed on the overall data of the Silesia database (all 4 channels of the 5 subjects taken together). These are shown on Fig. 10(a) and Fig. 10(b). For both Bland-Altman and scatter plots, behaviors are similar to those obtained for our database as presented on Fig. 8. It is worth noting that the available time instants of the fetal R-peaks of the scalp ECG signals in the Silesia database present a few errors of detection, shown in Fig. 10(b) as the two sets of horizontal points around 125 and 145 bpm of the  $FHR_{NMF}$ . This result highlights that our methodology is reproducible on other databases and is thus not specific to our signals.

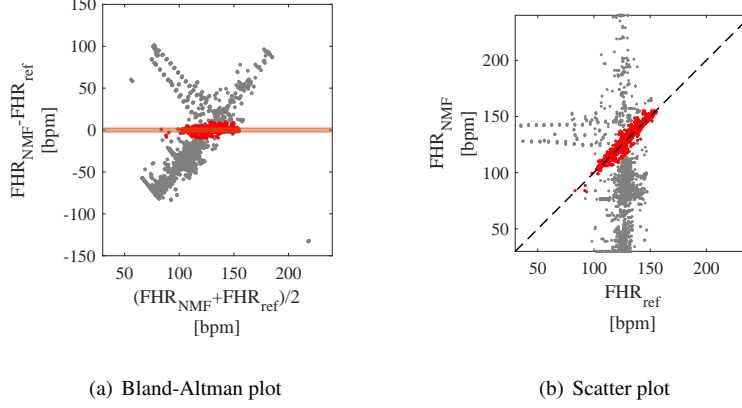


Figure 10: Overall FHR estimations from proposed NMF-methodology for the 5 volunteers of the Silesia database. (a) Bland-Altman plot of the whole estimations from the 4 channels including outliers (in grey) and without outliers (in red), the horizontal red lines correspond to the mean value of the differences and the limits of agreement without outliers, (b) scatter plot of  $FHR_{NMF}$  and  $FHR_{ref}$  (from R-peaks instants of the reference scalp ECG) with outliers (in grey) and without outliers (in red), the dashed black line corresponds to  $y = x$ .

Some remarks should also be made on CTG as the reference technology to measure FHR. As recommended by the International Federation of Gynecology and Obstetrics, CTG is nowadays the world-wide reference for intrapartum FHR monitoring. Therefore, it makes sense to compare results of new FHR monitoring methodological propositions to CTG reference. However, in our case, it has to be noticed that the two technologies (ECG and CTG) do not measure the same physiological phenomena: the ECG measures directly the electrical cardiac activity whereas the CTG provides, according to ultrasounds, an indirect measure of FHR. As indicated in [2], the CTG process results in an approximation of the true FHR, considered sufficiently accurate for analysis. So, when comparing estimations from these 2 technologies, even if the FHR trends are similar, *i.e.* the same kind of variations can be seen regarding the FHR baseline and acceleration/deceleration (Fig. A.11 and A.12, first row), fine variations of the FHR, *i.e.* FHR variability, will not be necessary the same. This can explain the lower  $R$  coefficient of  $V_4$  (Fig. A.11, last row) but a quite good fitting between  $FHR_{NMF}$  and  $FHR_{CTG}$  (Fig. A.11, first row). Moreover, it has to be noted that  $FHR_{NMF}$  values are often lower than  $FHR_{CTG}$  ones, especially for high FHR values: this is noticeable on  $V_2$  and  $V_3$ . This also allows to explain the value of  $MD = -6.0$  for  $V_3$  in Table 2, who has the highest average FHR. To investigate the differences between  $FHR_{NMF}$  and  $FHR_{CTG}$ , a manual detection of well identifiable  $R$  peaks in signal portions has been carried out. This leads to a related FHR estimation closer to  $FHR_{NMF}$  than to  $FHR_{CTG}$ , which suggests an overestimation of FHR by the considered CTG device. As a consequence, the comparison values of FHR from the considered fECG-based methodologies,  $FHR_{meth}$ , are impacted by these overestimated  $FHR_{CTG}$  values.

As already mentioned, FHR provided by CTG and ECG have the same trends, but

different FHR variabilities. A point-by-point comparison is therefore unfair, without a confidence margin. Instead of a classical relative threshold, expressed in percentage of the instantaneous FHR, we prefer absolute tolerance, independent of the FHR values. As already said, the clinical variability is judged normal when comprising peak-to-peak between 6 and 25 bpm [6]. This justifies our choice of 12.5 bpm for the confidence margin used for the computation of  $RO$ . Moreover, 12.5 bpm is about 10% of the classical FHR values range (120 – 160 bpm). In perspectives of this, it may be of interest to propose relevant clinical criteria in relation with the baseline rhythm, accelerations/decelerations and variability, in addition to the quantitative criteria proposed in this paper. Moreover, a concurrent indexation by clinicians of  $FHR_{CTG}$  and  $FHR_{NMF}$  outlines from a large signals database would be useful to compare the ability of the two technologies to reliably identify the clinical FHR variability, as the main feature for fetal well-being monitoring.

Finally, although our NMF-based proposition shows in the whole the best results compared to the state-of-art methodologies [17, 18, 19], some improvements should be investigated. Indeed the post-processing step, which allows to correct the bad fundamental frequencies detection as mentioned in Section 2.3.3, can be included directly in the NMF algorithm. Moreover, the excitation-filter model used in Section 2.3 to model the fECG contains only one harmonic part supposed to be the cardiac beat events. However, as already noticed, in clinical situations, some artifacts interfere with the fECG. Such specific components could be added in the model to improve the accuracy of the NMF. In addition, our proposition should be able to detect when the fECG signal is not available. For example,  $FHR_{NMF}$  of  $V_1$  depicted in first row of Fig.A.11 shows bad FHR estimations between 5 and 10 min probably due to loss of fECG, maybe because of fetal motions. At such instants, the method must not estimate FHR values, and thus does not display any FHR value as the CTG does in such situations: according to the International Federation of Gynecology and Obstetrics [28], a FHR outline (by CTG) remains exploitable with a signal loss ratio lower than 20%.

To conclude, a novel approach of FHR estimation from a single-channel abdominal ECG for an expected use in clinics has been investigated in this paper. Unlike most methods to estimate the FHR, the proposed one does not need neither to detect the R-peaks nor to train parameters. Evaluation of FHR estimations from ECG has been qualitatively and quantitatively carried out on a real signals database of 8 volunteers, related to the CTG clinical reference and comparative methodologies. The average duration of signals is of 21 min near of what is done in clinics (30 min is the classical monitoring duration during pregnancy). The results highlight robust performance of the proposed methodology, which may be considered as a promising solution for FHR estimation in clinical situations and a potential alternative to CTG in the future.

## Appendix A. Individual results

Figure A.11 and Figure A.12 illustrate the FHR estimation performance for each subject (i) by superimposing the  $FHR_{NMF}$  on reference  $FHR_{CTG}$  (first row), (ii) by displaying the Bland-Altman plots [29] of all the estimations with and without outliers

(second row) and (iii) by plotting the  $FHR_{NMF}$  according to the  $FHR_{CTG}$  after removal of outliers (third row).

### Acknowledgment

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] Intérêt et indications des modes de surveillance du rythme cardiaque foetal au cours de l'accouchement normal. Technical report, Agence nationale d'accréditation et d'évaluation en santé, 2003.
- [2] Diogo Ayres-de Campos, Catherine Y. Spong, Edwin Chandrachan, and FIGO Intrapartum Fetal Monitoring Expert Consensus Panel. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. *International Journal of Gynecology & Obstetrics*, 131(1):13–24, 2015.
- [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*. 2009.
- [4] D. Devane. Cardiotocography versus intermittent auscultation of fetal heart on admission to labour ward for assessment of fetal wellbeing. *Journal of Evidence-Based Medicine*, 5(2):100–100, 2012.
- [5] K. Bhogal and J. Reinhard. Maternal and fetal heart rate confusion during labour. *British Journal Midwifery*, 18(7):424–428, 2010.
- [6] A. Martin. Fetal heart rate during labour: definitions and interpretation. *Journal de gynécologie, obstétrique et biologie de la reproduction*, 37:S34–45, 2008.
- [7] Barbara S. Kisilevsky and C. Ann Brown. Comparison of fetal and maternal heart rate measures using electrocardiographic and cardiotocographic methods. *Infant Behavior and Development*, 42:142 – 151, 2016.
- [8] M. Cremer. Über die Ursache der elektromotorischen Eigenschaften der Gewebe, zugleich ein Beitrag zur Lehre von Polyphasischen Elektrolytketten. *Z. Biol*, 47, 1906.
- [9] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, JR. E. Dong, and R. C. Goodlin. Adaptive noise cancelling: Principles and applications. *Proceedings of the IEEE*, 63(12):1692–1716, 1975.
- [10] E. R. Ferrara and B. Widrow. Fetal electrocardiogram enhancement by time-sequenced adaptive filtering. *IEEE Transactions on Biomedical Engineering*, BME-29(6):458–460, 1982.



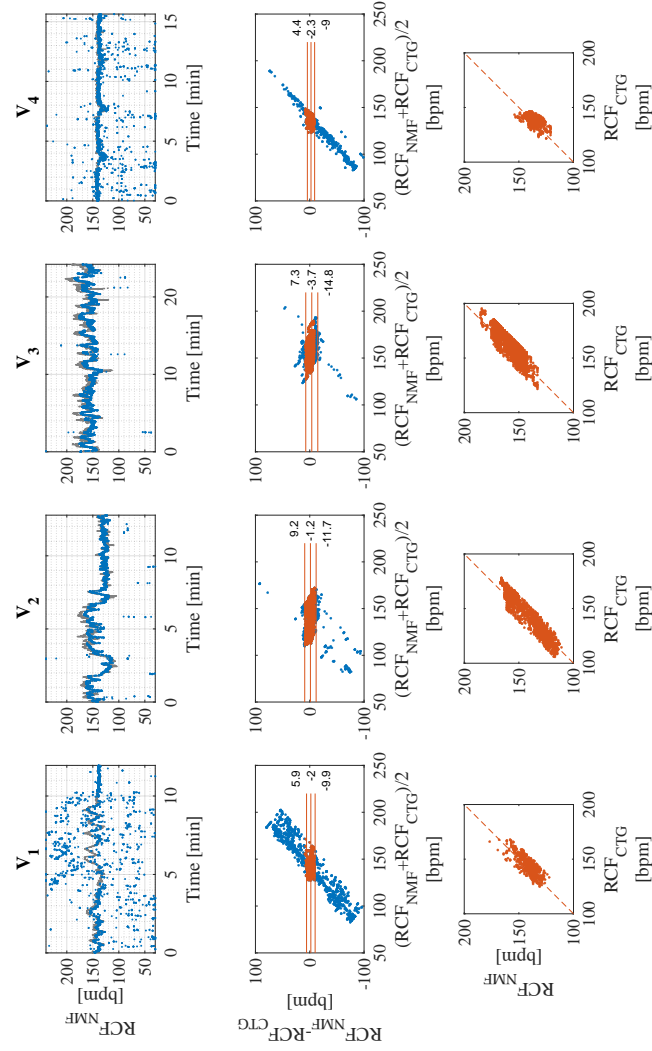


Figure A.11: FHR estimation results from proposed NMF-methodology for volunteers V1 to V4. From top to bottom: (i)  $FHR_{NMF}$  estimations superimposed on reference  $FHR_{CTG}$ , (ii) Bland-Altman plots of the whole recording including outliers (in blue) and without outliers (in red), the lines correspond to the mean value of the differences and the limits of agreement without outliers, (iii) correlations of  $FHR_{NMF}$  and  $FHR_{CTG}$  after removal of outliers, the line corresponds to  $y = x$ .

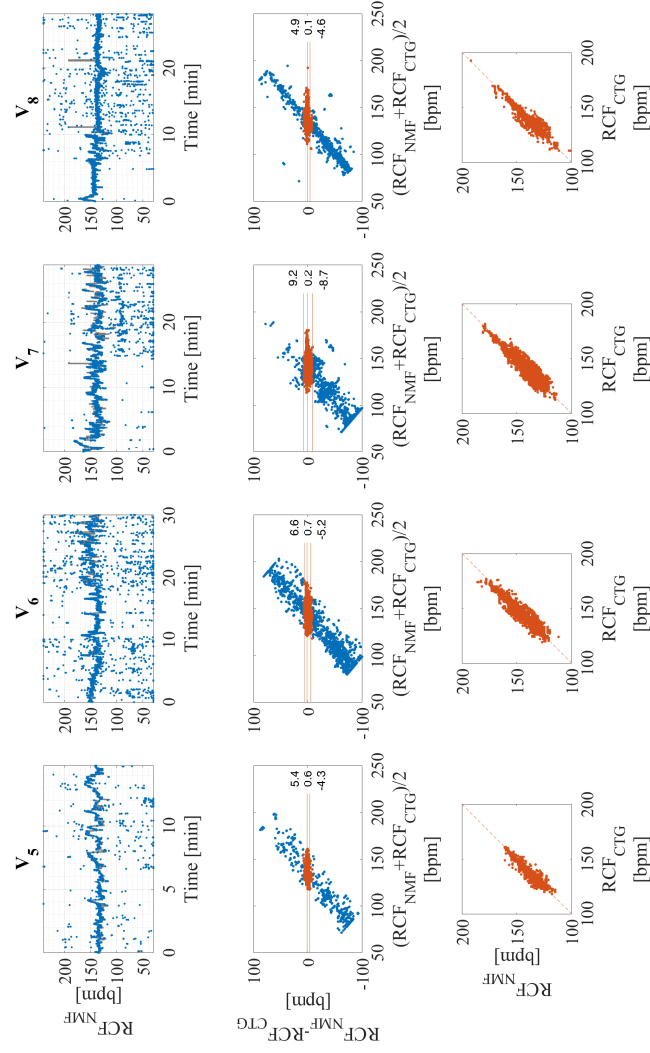


Figure A.12: FHR estimation results from proposed NMF-methodology for volunteers V5 to V8. From top to bottom: (i)  $FHR_{NMF}$  estimations superimposed on reference  $FHR_{CTG}$ , (ii) Bland-Altman plots of the whole recording including outliers (in blue) and without outliers (in red), the lines correspond to the mean value of the differences and the limits of agreement without outliers, (iii) correlations of  $FHR_{NMF}$  and  $FHR_{CTG}$  after removal of outliers, the line corresponds to  $y = x$ .

- [11] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. Fetal electrocardiogram extraction by blind source subspace separation. *IEEE Transactions on Biomedical Engineering*, 47(5):567–572, 2000.
- [12] V. Zarzoso and A. K. Nandi. Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation. *IEEE Transactions on Biomedical Engineering*, 48(1):12–18, 2001.
- [13] S. MM Martens, C. Rabotti, M. Mischi, and R. J. Sluijter. A robust fetal ECG detection method for abdominal recordings. *Physiological measurement*, 28(4):373, 2007.
- [14] G. Guojun, X. Zeng, X. Zhou, and Q. Liu. Robust adaptive fetal heart rate estimation for single-channel abdominal ECG recording. In *2012 5th International Conference on BioMedical Engineering and Informatics*, pages 587–591, 2012.
- [15] K. J. Lee and B. Lee. Sequential total variation denoising for the extraction of fetal ECG from single-channel maternal abdominal ECG. *Sensors*, 16(7), 2016.
- [16] A. Khamene and S. Negahdaripour. A new method for the extraction of fetal ECG from the composite abdominal signal. *IEEE Transactions on Biomedical Engineering*, 47(4):507–516, 2000.
- [17] A. Agostinelli, I. Marcantoni, E. Moretti, A. Sbröllini, S. Fioretti, F. Di Nardo, and L. Burattini. Noninvasive fetal electrocardiography part I: Pan-Tompkins’ algorithm adaptation to fetal R-peak identification. *The Open Biomedical Engineering Journal*, 11:17–24, 2017.
- [18] Z. Zidelmal, A. Amirou, M. Adnane, and A. Belouchrani. QRS detection based on wavelet coefficients. *Computer Methods and Programs in Biomedicine*, 107(3):490–496, 2012.
- [19] Ds S. Benitez, P. A. Gaydecki, A. Zaidi, and A. P. Fitzpatrick. A new QRS detection algorithm based on the hilbert transform. In *Computers in Cardiology 2000. Vol. 27 (Cat. 00CH37163)*, pages 379–382, 2000.
- [20] J. Pan and W. J. Tompkins. A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, 1985.
- [21] W. Liu, J. Principe, and S. Haykin. *Kernel Adaptive Filtering: A Comprehensive Introduction*. Wiley, 2010.
- [22] N. Dia, J. Fontecave-Jallon, P-Y. Gumery, and B. Rivet. Heart rate estimation from phonocardiogram signals using non-negative matrix factorization. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1293–1297, 2019.
- [23] N. Dia, J. Fontecave-Jallon, P-Y. Gumery, and B. Rivet. Fetal heart rate estimation from a single phonocardiogram signal using non-negative matrix factorization. *Proc. IEEE EMBC*, 2019. (Accepted for publication).

- [24] P. Paatero and U. Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994.
- [25] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788, 1999.
- [26] C. Févotte and J. Idier. Algorithms for nonnegative matrix factorization with the  $\beta$ -divergence. *Neural computation*, 23(9):2421–2456, 2011.
- [27] Wayne R. Cohen, Sophia Ommani, Sarmina Hassan, Fadi G. Mirza, Molham Solomon, Raymond Brown, Barry S. Schiffrin, John M. Himsworth, and Barrie R. Hayes-Gill. Accuracy and reliability of fetal heart rate monitoring using maternal abdominal surface electrodes. *Acta Obstetrica et Gynecologica Scandinavica*, 91(11):1306–1313, 2012.
- [28] T. Fuchs, M. Pomorski, K. Grobelak, M. Tomiałowicz, and M. Zimmer. Signal loss during fetal heart rate monitoring using maternal abdominal surface electrodes between 28 and 42 weeks of pregnancy. *Advances in clinical and experimental medicine: official organ Wroclaw Medical University*, 23(5):813–819, 2014.
- [29] J. M. Bland and D. Altman. Statistical-methods for assessing agreement between 2 methods of clinical measurement. *International journal of nursing studies*, 47, 2010.
- [30] A. Matonia, J. Jezewski, K. Horoba, A. Gacek, and P. Labaj. The maternal ECG suppression algorithm for efficient extraction of the fetal ECG from abdominal signal. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3106–3109, Aug 2006.
- [31] Janusz Jezewski, Adam Matonia, Tomasz Kupka, Roj Dawid, and Robert Czabanski. Determination of fetal heart rate from abdominal signals: evaluation of beat-to-beat accuracy in relation to the direct fetal electrocardiogram. *Biomedizinische Technik/Biomedical Engineering*, 57(5):383–394, 2012.
- [32] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23):e215–e220, 2000.

Table 2: Quantitative evaluation between estimated  $FHR_{meth}$  and reference  $FHR_{CTG}$ . Recording duration ( $RD$ ), consistency of the estimation  $1 - RO$  ( 8), median deviation from the reference  $MD$  ( 9) and Pearson correlation coefficient  $R$  ( 10). Values are depicted for the 8 volunteers and the 4 compared methodologies NMF, Hilbert, Wavelets and Pan & Tompkins. Values in bold cases highlight which method performs best in terms of either  $1 - RO$ ,  $MD$  and  $R$ . For each methodology, the outliers are FHR estimations that differ from CTG by more than 12.5 bpm. Note that  $R$  is computed after outliers' removal.

Volunteer	$RD$ [min]	Methodologies											
		NMF			Hilbert			Wavelets			PT		
		$1 - RO$ [%]	$MD$ [bpm]	$R$ [%]	$1 - RO$ [%]	$MD$ [bpm]	$R$ [%]	$1 - RO$ [%]	$MD$ [bpm]	$R$ [%]	$1 - RO$ [%]	$MD$ [bpm]	$R$ [%]
$V_1$	12	<b>52</b>	<b>-2.0</b>	<b>83</b>	43	-3.2	72	14	61.6	71	16	-42.9	68
$V_2$	13	<b>90</b>	-1.6	<b>95</b>	82	-1.6	<b>95</b>	66	<b>-1.1</b>	<b>95</b>	47	-3.9	92
$V_3$	24	<b>79</b>	<b>-6.1</b>	<b>88</b>	58	-8.3	85	15	147.3	83	36	-10.3	86
$V_4$	16	<b>77</b>	-3.3	<b>58</b>	49	<b>-2.7</b>	49	12	143.4	35	18	41.4	46
$V_5$	15	<b>82</b>	<b>0.0</b>	<b>96</b>	70	0.3	92	36	21.7	85	19	6.6	78
$V_6$	30	<b>68</b>	<b>0.0</b>	<b>96</b>	12	-62.3	86	22	57.7	85	0	-141.7	—
$V_7$	29	<b>74</b>	<b>-0.7</b>	<b>92</b>	37	-2.2	88	15	-44.1	91	4	-58.3	77
$V_8$	28	<b>82</b>	<b>0.2</b>	<b>94</b>	32	-1.0	79	7	157.3	79	23	-2.7	76
<i>Median</i>	20	<b>78.2</b>	<b>-1.2</b>	<b>93.0</b>	46.0	-2.5	85.5	15.0	59.7	84.0	18.5	-7.1	77.0
<i>Mean</i>	20.9	<b>75.6</b>	<b>-1.7</b>	<b>87.6</b>	47.9	-10.1	80.8	23.4	67.9	78.0	20.4	-25.8	74.7