



## Data Article

# Annotated real and synthetic datasets for non-invasive foetal electrocardiography post-processing benchmarking

Giulia Baldazzi<sup>a,b,1,\*</sup>, Eleonora Sulas<sup>a,1</sup>, Monica Urru<sup>c</sup>,  
Roberto Tumbarello<sup>c</sup>, Luigi Raffo<sup>a</sup>, Danilo Pani<sup>a</sup>

<sup>a</sup> Department of Electrical and Electronic Engineering (DIEE), University of Cagliari, Piazza d'Armi, 09122 Cagliari Italy

<sup>b</sup> Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genoa, Via Opera Pia 13, 16145 Genoa Italy

<sup>c</sup> Division of Paediatric Cardiology, San Michele Hospital, Piazzale Alessandro Ricchi 1, 09134 Cagliari Italy

## ARTICLE INFO

## Article history:

Received 12 June 2020

Revised 31 August 2020

Accepted 6 October 2020

Available online 10 October 2020

## Keywords:

Non-invasive fECG

Biopotential recordings

Abdominal ECG

Signal processing

fECG post-processing

Denoising

## ABSTRACT

Non-invasive foetal electrocardiography (fECG) can be obtained at different gestational ages by means of surface electrodes applied on the maternal abdomen. The signal-to-noise ratio (SNR) of the fECG is usually low, due to the small size of the foetal heart, the foetal-maternal compartment, the maternal physiological interferences and the instrumental noise. Even after powerful fECG extraction algorithms, a post-processing step could be required to improve the SNR of the fECG signal. In order to support the researchers in the field, this work presents an annotated dataset of real and synthetic signals, which was used for the study "Wavelet Denoising as a Post-Processing Enhancement Method for Non-Invasive Foetal Electrocardiography" [1]. Specifically, 21 15 s-long fECG, dual-channel signals obtained by multi-reference adaptive filtering from real electrophysiological recordings were included. The annotation of the foetal R peaks by an expert cardiologist was also provided. Recordings were performed on 17 voluntary pregnant women between the 21st

DOI of original article: [10.1016/j.cmpb.2020.105558](https://doi.org/10.1016/j.cmpb.2020.105558)

\* Corresponding author at: Department of Electrical and Electronic Engineering (DIEE), University of Cagliari, Piazza d'Armi, 09122 Cagliari Italy.

E-mail address: [giulia.baldazzi@unica.it](mailto:giulia.baldazzi@unica.it) (G. Baldazzi).

<sup>1</sup> These authors equally contributed.

<https://doi.org/10.1016/j.dib.2020.106399>

2352-3409/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

and the 27th week of gestation. The raw recordings were also included for the researchers interested in applying a different fECG extraction algorithm. Moreover, 40 10 s-long synthetic non-invasive fECG were provided, simulating the electrode placement of one of the abdominal leads used for the real dataset. The annotation of the foetal R peaks was also provided, as generated by the FECGSYN tool used for the signals' creation. Clean fECG signals were also included for the computation of indexes of signal morphology preservation. All the signals are sampled at 2048 Hz. The data provided in this work can be used as a benchmark for fECG post-processing techniques but can also be used as raw signals for researchers interested in foetal QRS detection algorithms and fECG extraction methods.

© 2020 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## Specifications Table

|                                |  |
|--------------------------------|--|
| Subject                        | Biomedical Engineering   |
| Specific subject area          | Non-invasive foetal electrocardiography  |
| Type of data                   | Biosignals: ECG, real (1) and synthetic (2) - .csv files<br>Annotations (fECG QRS position) - .csv files   |
| How data were acquired         | <ol style="list-style-type: none"> <li>1. Portable physiological measurement system</li> </ol> <p>Instruments: hardware, software<br/>Model and maker: Porti7 hardware and PolyBench software, by TMSi, The Netherlands</p> <ol style="list-style-type: none"> <li>2. Physiological signal simulator</li> </ol> <p>Instruments: software<br/>Model and maker: FECGSYN tool [2–4]</p>   |
| Data format                    | Raw<br>Filtered<br>Annotations   |
| Parameters for data collection | <ol style="list-style-type: none"> <li>1. Pregnant women between the 21st and the 27th week of gestation with healthy foetuses</li> <li>2. Maternal ECG at 90 bpm, fECG at 150 bpm, SNR of fECG relative to maternal ECG at –18 dB, only white and pink noise, unique foetal and maternal heart positions and unique electrode map</li> </ol>  |
| Description of data collection | <ol style="list-style-type: none"> <li>1. Real biopotentials acquired with the subject at rest, in a comfortable semi-sitting position. Prior to the electrodes' application on the abdomen and torso, a mild skin treatment was performed to reduce the skin contact impedance</li> <li>2. Synthetic biopotentials were created with the FECGSYN tool [2–4] in order to have access to the fECG signal corrupted by artificial noise but free from the maternal ECG interference</li> </ol> |

(continued on next page)

|  |   |
|--|---|
| Data source location                           | 1. Institution: Azienda Ospedaliera Brotzu (AOB)<br>City/Town/Region: Cagliari<br>Country: Italy<br><br>2. Toolbox reference: FECGSYN tool [2–4]  |
| Data accessibility<br>Related research article | With the article, <a href="https://data.mendeley.com/datasets/64zy9f2dkf/1">https://data.mendeley.com/datasets/64zy9f2dkf/1</a><br>Giulia Baldazzi, Eleonora Sulas, Monica Urru, Roberto Tumbarello, Luigi Raffo and Danilo Pani, “Wavelet Denoising as a Post-Processing Enhancement Method for Non-Invasive Foetal Electrocardiography”, Computer Methods and Programs in Biomedicine [1] |

## Value of the Data

- These data are useful to benchmark different fECG denoising algorithms.
- These data are mainly conceived for researchers working on non-invasive fECG signal processing.
- By including the real raw signals acquired with surface ECG electrodes on the maternal body, the data can be also used to test non-invasive fECG extraction algorithms.
- These data are particularly important, considering the lack of annotated fECG datasets in early pregnancy.

## 1. Data Description

The dataset is composed of 63 .csv files for the real data and 120 .csv files for the synthetic data.

Regarding the real recordings, 63 .csv files include:

- (I) 21 .csv files consisting of raw and processed recordings (each data file containing: three non-coplanar maternal ECG leads, horizontal, vertical and oblique abdominal leads, extracted horizontal, vertical and oblique fECG leads).
- (II) 21 .csv files composed of only the extracted fECG leads used in [1].
- (III) 21 .csv files reporting the R peaks annotation for the corresponding data file. The annotation files present the number of samples in which the cardiologist labelled the R peaks.

Regarding the synthetic data, there are:

- (I) 40 .csv file storing the completely clean fECG, the pink and the white noise.
- (II) 40 .csv files reporting the R peak annotations provided by the FECGSYN tool [2–4].
- (III) 40 .csv files storing the simulated extracted abdominal fECG created by the summation of the completely clean fECG, the pink and the white noise, as used in [1].

Table 1 illustrates the structure of the proposed dataset. The signals used in [1] are those identified as Real\_FECG\_n (with the annotations in Real\_FECG\_n\_ann) for the real dataset, and Synth\_FECG\_n (with the annotations in Synth\_ann\_n) for the synthetic dataset. The files Real\_Traces\_n can be useful for the researchers willing to try different fECG extraction algorithms before the post-processing stage. For both the real and synthetic data, sampling frequency is 2048 Hz. The chosen format allows an easy importing by different software processing tools.

All real and synthetic data are meant to be used as a benchmark for fECG post-processing techniques, thus enabling the homogeneous analysis of different denoising algorithms, applied after the fECG extraction process, which is still missing in the scientific literature. Moreover, these data can also be used as raw signals for the evaluation of different foetal QRS detection algorithms and fECG extraction methods, providing the annotation of the foetal QRS complexes.

**Table 1**

Dataset description. For both the real and synthetic data, sampling frequency is 2048 Hz.

| DATA           | FILES                                  | FILE STRUCTURE BY COLUMN  |
|----------------|--|---|
| REAL DATA      | 21 .csv files named Real_Traces_n      | <ol style="list-style-type: none"> <li>1. first maternal ECG lead</li> <li>2. second maternal ECG lead</li> <li>3. third maternal ECG lead</li> <li>4. horizontal abdominal lead</li> <li>5. vertical abdominal lead</li> <li>6. oblique abdominal lead</li> <li>7. fECG extracted from the horizontal lead</li> <li>8. fECG extracted from the vertical lead</li> <li>9. fECG extracted from the oblique lead</li> </ol> |
|                | 21 .csv files named Real_FECG_n        | The selected two fECG signals used in [1]. Table 4 lists the selected leads.  |
|                | 21 .csv files named Real_FECG_n_ann    | Sample numbers related to foetal R peaks (manual annotations)   |
| SYNTHETIC DATA | 40 .csv files named Synth_FECG_n       | 1. noisy fECG signal contributing to the horizontal abdominal lead (free from maternal interference)  |
|                | 40 .csv files named Synth_components_n | <ol style="list-style-type: none"> <li>1. clean fECG signal contribution to the horizontal abdominal lead</li> <li>2. additive pink noise</li> <li>3. additive white noise</li> </ol>   |
|                | 40 .csv files named Synth_ann_n        | Sample numbers related to foetal R peaks (provided by the FECGSYN tool)   |

## 2. Experimental Design, Materials and Methods

The dataset is composed of two parts, real and synthetic signals, in order to analyse all possible effects introduced by fECG post-processing algorithms. In this regard, while the former enables an actual analysis of their outcome in real applications, a ground truth of the noiseless fECG waveforms is provided only by synthetic signals. Thus, only in the latter case an evaluation of possible morphological distortions due to the processing stages can be carried out.

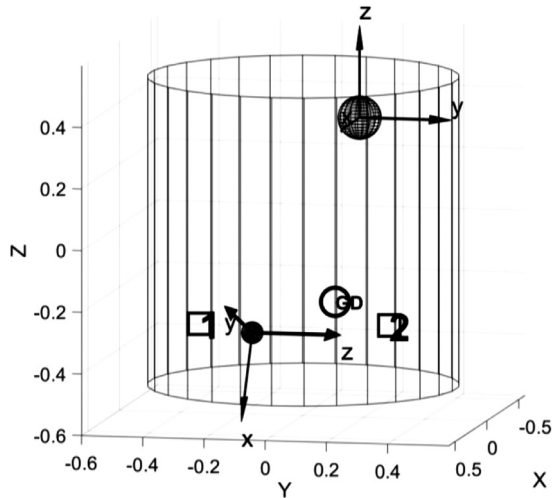
### 2.1. Synthetic Dataset

An open-source simulator for non-invasive fECG research was adopted to create a small synthetic dataset [2–4]. The tool can be freely downloaded.<sup>1</sup>

By means of this tool, it is possible to generate several maternal-fECG mixtures, selecting the beat-to-beat variability, heart rate baseline value and changes, the SNR of the fECG relative to the maternal ECG ( $SNR_{fm}$ ), the SNR of the maternal ECG relative to noise ( $SNR_{mn}$ ), the position and orientation of the foetal and maternal hearts, foetal movements and uterine contractions. The toolbox calibrates the physiologic and noise components with respect to the maternal ECG signal by choosing the  $SNR_{fm}$  and the  $SNR_{mn}$ .

For the creation of the dataset presented in this work, all parameters were selected in order to simulate physiologically plausible recordings and, wherever possible, in accordance with the real dataset acquisition setup. Specifically, the sampling frequency was set to 2048 Hz, as in the real dataset, and the signal duration to 10 s. The maternal and foetal heart rates were assumed to be fixed at respectively 90 and 150 bpm. The maternal heart coordinates were imposed equal to  $[\frac{2}{3}\pi, 0.2, 0.4]$  and those of the foetus to  $[-0.412, 0.284, -0.300]$ . The coordinates of the virtual surface electrodes detecting the abdominal ECGs were set to  $[-0.524, 0.500, -0.250]$  and  $[0.785, 0.500, -0.250]$ , which correspond to the schematic representation depicted in Fig. 1.

<sup>1</sup> Available at <http://fernandoandreotti.github.io/fecgsyn/>



**Fig. 1.** Graphical representation, provided by the FECGSYN tool, illustrating the virtual maternal torso, the location of the maternal heart (upper sphere) and the foetal heart (lower sphere). Squared boxes, located on the maternal abdomen, indicate electrode positions and are numbered for the corresponding output channel. 1-2 represents the horizontal lead provided in this dataset, which simulates the horizontal lead adopted in the real dataset. The reference electrode is indicated by the GD circle, approximately placed on the back.

Virtual electrodes capture unipolar signals, referenced to a common GD electrode, so that a single differential lead was digitally created by subtracting the signal detected by the electrode 2 to the one detected by the electrode 1.

By construction, the differential abdominal lead was obtained completely free of the maternal ECG components, as it would happen when the signal is produced by a perfect fECG extraction stage. For the same reason, we avoided adding realistic noise sources mimicking muscular artefacts, baseline wander and electrode movements. Conversely, white and pink noises created by using the OSET toolbox [5] were added to the abdominal lead, with amplitudes chosen according to the desired SNR. Since the FECGSYN tool does not allow to set directly the SNR of fECG relative to noise, we set the  $SNR_{fm}$  to  $-18$  dB, whereas  $SNR_{mn}$  was linearly varied in the range between 3 dB and 15 dB producing the 40 synthetic signals released in this dataset. Table 2 lists the SNR level characterizing each synthetic trace.

The signals accompanying the synthetic dataset used in [1] (Synth\_FECG\_n, i.e. the abdominal fECG signal with added pink and white noises scaled as in Table 2, and the associated foetal QRS annotations provided by the FECGSYN toolbox, Synth\_ann\_n) are provided along with the clean fECG signal on the abdominal lead and the noise sources (white and pink) properly scaled to give rise, by summing them to the clean fECG, to the noisy synthetic fECG signal. Such signals, stored in the Synth\_components\_n files, can be useful for the computation of some performance metrics characterizing the capability of the post-processing algorithms to preserve the fECG morphology, such as root-mean-square error and linear or non-linear correlation coefficients.

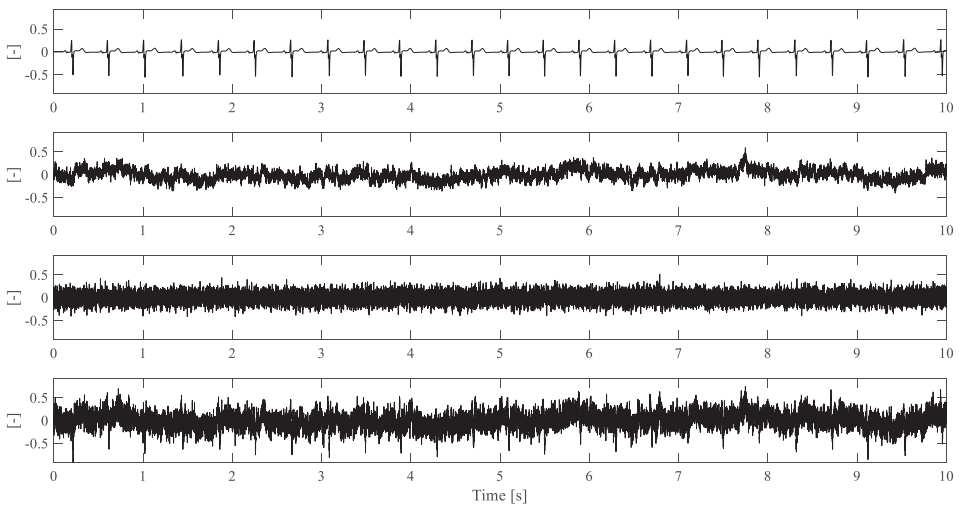
Two examples of the abdominal lead for two virtual foetuses included in this dataset are shown in Figs. 2 and 3.

## 2.2. Real dataset

The real dataset was derived from 21 recordings that were acquired at the Division of Paediatric Cardiology, San Michele Hospital, Cagliari (Italy) by the authors. The recording protocol was approved by the Independent Ethical Committee of the Cagliari University Hospital (AOU

**Table 2**  
 $SNR_{mn}$  values for all the signals composing the synthetic dataset.

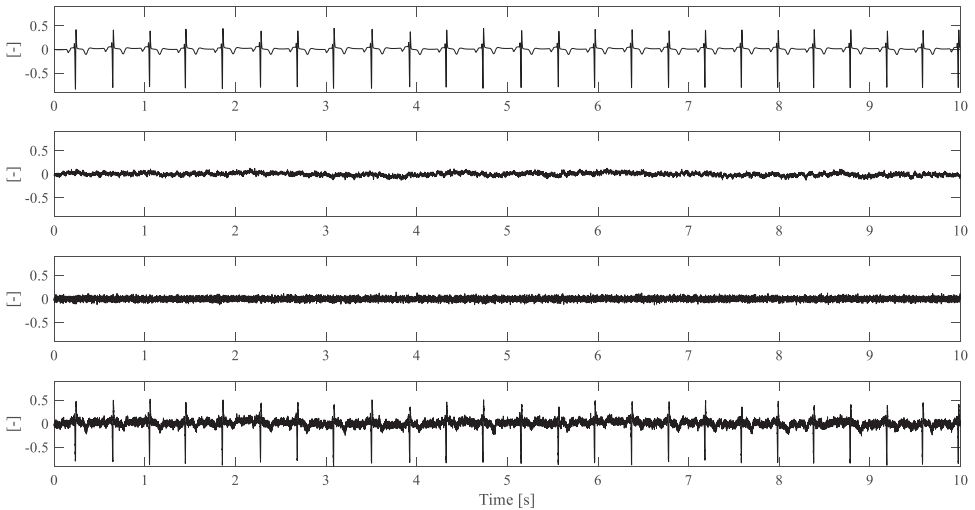
| SIGNAL        | $SNR_{mn}$ [dB] | Signal        | $SNR_{mn}$ [dB] |
|---------------|-----------------|---------------|-----------------|
| Synth_FECG_1  | 3.00            | Synth_FECG_21 | 9.15            |
| Synth_FECG_2  | 3.31            | Synth_FECG_22 | 9.46            |
| Synth_FECG_3  | 3.62            | Synth_FECG_23 | 9.77            |
| Synth_FECG_4  | 3.92            | Synth_FECG_24 | 10.10           |
| Synth_FECG_5  | 4.23            | Synth_FECG_25 | 10.38           |
| Synth_FECG_6  | 4.54            | Synth_FECG_26 | 10.69           |
| Synth_FECG_7  | 4.85            | Synth_FECG_27 | 11.00           |
| Synth_FECG_8  | 5.15            | Synth_FECG_28 | 11.31           |
| Synth_FECG_9  | 5.46            | Synth_FECG_29 | 11.62           |
| Synth_FECG_10 | 5.77            | Synth_FECG_30 | 11.92           |
| Synth_FECG_11 | 6.08            | Synth_FECG_31 | 12.23           |
| Synth_FECG_12 | 6.39            | Synth_FECG_32 | 12.54           |
| Synth_FECG_13 | 6.69            | Synth_FECG_33 | 12.85           |
| Synth_FECG_14 | 7.00            | Synth_FECG_34 | 13.15           |
| Synth_FECG_15 | 7.31            | Synth_FECG_35 | 13.46           |
| Synth_FECG_16 | 7.62            | Synth_FECG_36 | 13.77           |
| Synth_FECG_17 | 7.92            | Synth_FECG_37 | 14.08           |
| Synth_FECG_18 | 8.23            | Synth_FECG_38 | 14.38           |
| Synth_FECG_19 | 8.54            | Synth_FECG_39 | 14.69           |
| Synth_FECG_20 | 8.85            | Synth_FECG_40 | 15.00           |



**Fig. 2.** Example of the signals involved in the creation of a horizontal abdominal lead of the synthetic dataset, characterized by an  $SNR_{mn}$  of 3 dB. From top to bottom, the clean fECG signal, the pink noise, the white noise, and the noisy fECG signal obtained by their sum. Amplitudes are dimensionless.

Cagliari) and performed following the principles outlined in the Helsinki Declaration of 1975, as revised in 2000. The volunteers provided their signed informed consent to the protocol. The recordings were carried out with the patient in a comfortable semi-sitting position. Before applying the electrodes, mild skin treatment was performed on the maternal abdomen by using an abrasive gel (NuPrep, by Weaver and Company, USA) to reduce the skin contact impedance.

The signals originated from 17 pregnant women satisfying the inclusion criteria during the dataset acquisition time frame, i.e. healthy foetuses between the 21st and 27th weeks of gestation. For each voluntary woman, only one recording session was performed, from which the good-quality abdominal signals' segments were extracted. The relevant anamnestic information



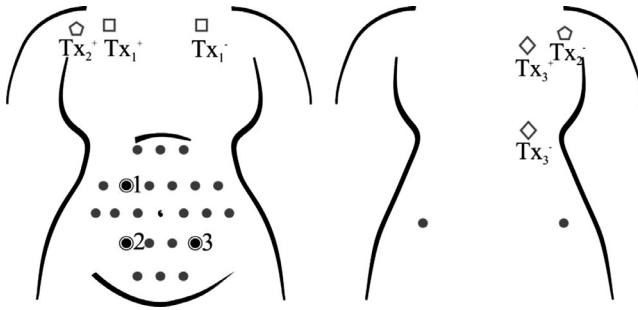
**Fig. 3.** Example of a 15 dB horizontal abdominal lead of the synthetic dataset (lowest plot) along with the three components giving rise to it: the pure foetal ECG signal (upper plot), the pink noise (second row) and white noise (third row). Amplitudes are dimensionless.

**Table 3**

Real dataset relevant anamnestic information (foetal presentation: L: left, R: right, O: occiput, S: sacrum, T: transverse, P: posterior A: anterior). Abdominal signals acquired from the same foetus are grouped together in the same row.

| REAL SIGNAL | week of gestation | foetus presentation |
|-------------|-------------------|---------------------|
| 1           | 24                | breech,             |
| 2           |                   | LST                 |
| 3           |                   |                     |
| 4           | 21                | breech,             |
| 5           |                   | RSP                 |
| 6           | 24                | breech, RSP         |
| 7           | 26                | vertex, ROP         |
| 8           | 25                | breech, RSP         |
| 9           | 25                | vertex, ROP         |
| 10          | 24                | vertex, LOP         |
| 11          | 21                | breech, RSP         |
| 12          | 25                | breech, RSP         |
| 13          | 24                | breech, LSP         |
| 14          | 24                | breech, LSP         |
| 15          | 21                | vertex,             |
| 16          |                   | OP                  |
| 17          | 22                | breech, LSP         |
| 18          | 24                | vertex, LOP         |
| 19          | 26                | vertex, ROP         |
| 20          | 25                | breech, LSP         |
| 21          | 23                | vertex, LOP         |

is provided in Table 3. The gestational epoch was selected to have the best morphological accuracy for the transabdominal signals. The lower limit is due to the small size of the foetus heart, that cannot guarantee a good signal acquisition, whereas the upper limit allows reducing the influence of the *vernix caseosa*, a waxy insulating biological layer completely covering the foetal skin between the 28th to the 32nd week of gestation [6]. During this time frame, the fECG signal propagation towards the maternal abdomen is hampered, making the foetal signal hardly detectable with surface electrodes. After the 32nd week of gestation, this layer is usually



**Fig. 4.** Electrode positioning for the real dataset acquisition. Circles indicate the electrodes utilized for unipolar recording while squares, pentagons and rhombuses the electrodes for thoracic bipolar ones. Specifically, black circles represent the unipolar electrodes selected for the construction of the differential abdominal leads.

partially removed by the foetal movements but the morphology of the fECG signal that can be non-invasively acquired through surface electrodes is questioned [7].

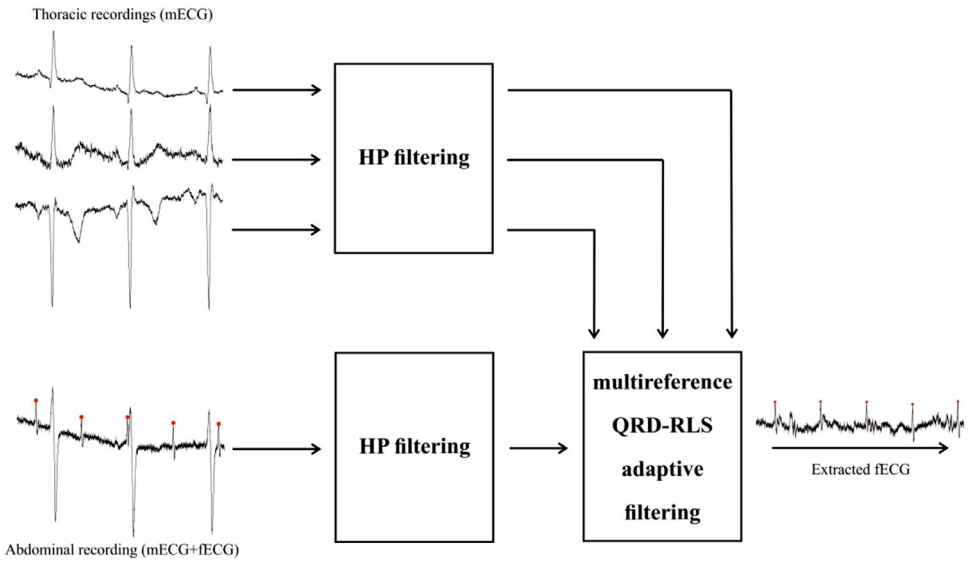
The biopotentials were recorded with the Porti7 portable physiological measurement system (TMSi, The Netherlands), that has 22 bits of resolution (71.5 nV amplitude resolution), and choosing a sampling frequency set at 2048 Hz. The biopotential measurement system is characterized by an input bandwidth limited by a digital decimation filter to approximately 550 Hz (0.27x the sampling frequency). Moreover, it featured both differential and unipolar channels, and every unipolar channel is acquired with respect to the average of all the unipolar channels applied on the body so that the SNR of the raw signal is influenced by the number of used channels in the measurement. Three differential channels were used on the maternal thorax for the acquisition of a reference maternal ECG, whereas 24 unipolar channels were used on the abdomen, in order to have the highest possible SNR on the abdominal signals in terms of common-mode signal rejection. Among the unipolar channels, only three of them, clearly indicated in Fig. 4, were selected to give rise to three abdominal differential leads digitally, such as 1–2 to obtain a horizontal lead, 2–3 to obtain a vertical lead and 1–3 to obtain an oblique lead.

In order to obtain the fECG, among the various methods available for the extraction of the fECG from non-invasive recordings, we chose to exploit a high-performance multireference QR-decomposition with back-substitution recursive least-squares (QRD-RLS) adaptive filter, available from the authors of [8], given its numerical stability and good performance [9]. Compared to other techniques, such as blind source separation [10,11], adaptive filters are real-time algorithms able to extract the fECG from a reduced number of abdominal electrodes, which can be useful for wearable systems [12] and real-time embedded platforms [11]. The three thoracic leads were used as reference signals in order to produce a good representation of the maternal ECG on the abdomen. The forgetting factor  $\lambda$  was set to 0.999 while the number of filter coefficients was set to 20, in accordance with previous studies [13]. Considering that the recording device adopted for the raw data collection was DC-coupled, all the leads were pre-processed to reduce the baseline wandering artefact in order to facilitate the QRD-RLS adaptive filter processing. In fact, such an artefact appears in different forms in the various leads and cannot be reduced by the application of the adaptive filter. Pre-processing consisted of a high-pass linear-phase equiripple finite impulse response filter with a cut-off frequency of 1 Hz. Compared with the current guidelines that would impose a 0.67 Hz limit for the cut-off frequency of ECG signals [14], some studies on non-invasive fECG proved that a higher cut-off frequency of the high-pass filter allows achieving better results, still preserving the low-frequency components of the ECG, mainly related to P and T and to the ST segments [13].

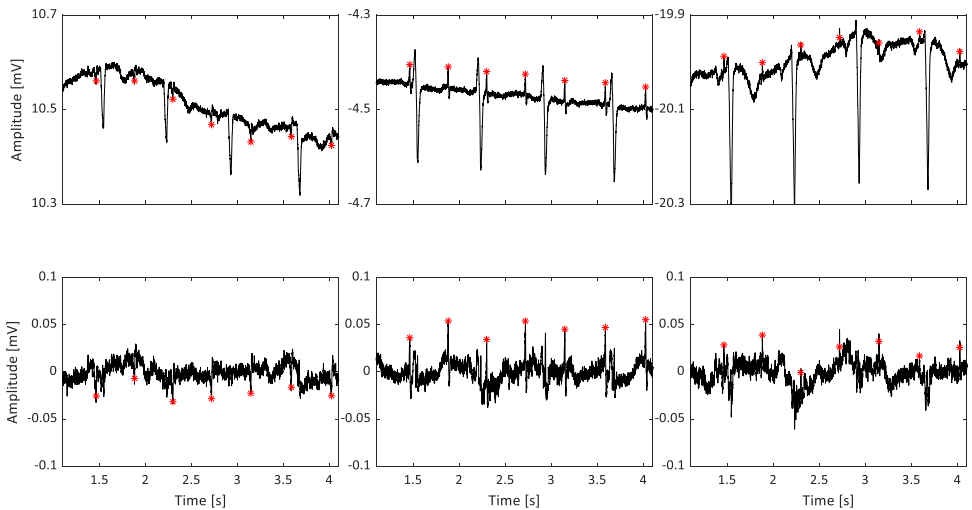
The whole processing chain used to extract the fECG is presented in Fig. 5.

The annotation of the foetal QRS complexes (Real\_FECG\_n\_ann) was performed manually and validated by the cardiologist with the help of a simultaneous cardiac foetal pulsed-wave Doppler signal recorded during the procedure, thus providing a mechanical reference for the electrophys-





**Fig. 5.** Schematic representation of the whole processing chain for fECG extraction. Three thoracic leads along with each abdominal trace are pre-processed with a high-pass filter with a cut-off frequency of 1 Hz and passed to the QRD-RLS adaptive filter for fECG extraction. Red dots mark the foetal QRS complexes.



**Fig. 6.** Example of raw abdominal traces (upper plots) and the corresponding extracted foetal ECG signals (lower plots). From the left to the right: horizontal, vertical and oblique leads. Red markers indicate the manually-annotated foetal R peaks.

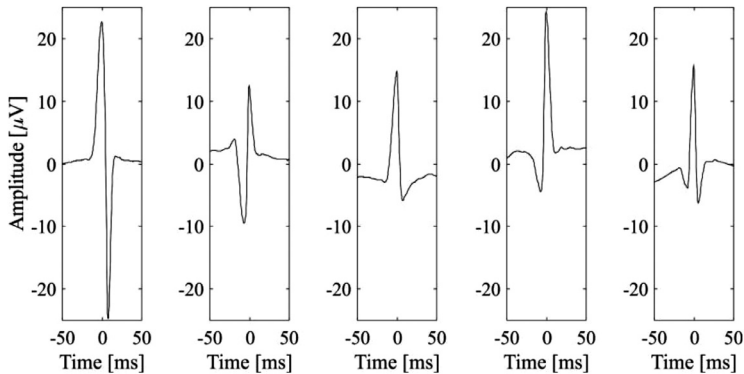
iological signal. Despite the low signal-to-noise ratio of some real traces, the presence of three different simultaneous abdominal leads (horizontal, vertical and oblique) helps the R peaks annotation, enabling the identification of clear foetal R peaks at least in one good-quality trace. The annotation procedure was verified before and after the fECG extraction process to provide the most accurate indication of the foetal QRS complexes.

Some examples of the extracted fECG are shown in Fig. 6. From this inspection, it emerged that some extracted fECG channels were unreadable, from a clinical perspective. In order to use

**Table 4**

The abdominal leads chosen for the study [1] and included in the real dataset.

| SIGNAL       | fecg channel 1 | FECG CHANNEL 2 |
|--------------|----------------|----------------|
| Real_FECG_1  | horizontal     | oblique        |
| Real_FECG_2  | horizontal     | oblique        |
| Real_FECG_3  | horizontal     | oblique        |
| Real_FECG_4  | horizontal     | vertical       |
| Real_FECG_5  | horizontal     | vertical       |
| Real_FECG_6  | horizontal     | oblique        |
| Real_FECG_7  | vertical       | oblique        |
| Real_FECG_8  | horizontal     | oblique        |
| Real_FECG_9  | horizontal     | oblique        |
| Real_FECG_10 | horizontal     | vertical       |
| Real_FECG_11 | horizontal     | oblique        |
| Real_FECG_12 | horizontal     | oblique        |
| Real_FECG_13 | horizontal     | vertical       |
| Real_FECG_14 | horizontal     | oblique        |
| Real_FECG_15 | vertical       | oblique        |
| Real_FECG_16 | vertical       | oblique        |
| Real_FECG_17 | vertical       | oblique        |
| Real_FECG_18 | vertical       | oblique        |
| Real_FECG_19 | horizontal     | oblique        |
| Real_FECG_20 | vertical       | oblique        |
| Real_FECG_21 | vertical       | oblique        |



**Fig. 7.** Main different foetal heartbeat morphologies from the real dataset. Specifically, the represented foetal QRS complexes were obtained by synchronized averaging of the highly correlated beats (Pearson's correlation coefficient  $> 0.6$ ) after wavelet post-processing exploiting Stationary Wavelet Transform with 7-level decomposition and Han et al. threshold [1].

the same number of leads for each recording, the best two fECG traces per recording were selected to be released in the real dataset (Real\_FECG\_ $n$ ) used in [1]. For the sake of completeness, they are reported in Table 4.

The morphology of the foetal QRS complexes in the real dataset exhibits a high variability, as clearly visible in Fig. 7, depending on the foetal presentation and the recording lead (Table 4), allowing to test the post-processing algorithms in different conditions.

Along with the real fECG dataset, the raw thoracic maternal ECG differential leads, the three abdominal recordings and the fECG signals extracted by the method presented above were included in a separate file (Real\_Traces\_ $n$ ) for those scientists interested in applying and benchmarking different fECG extraction algorithms before the post-processing stage.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

## Acknowledgments

The authors wish to thank the team headed by Dr. Roberto Tumbarello, Division of Paediatric Cardiology, AOB Hospital (Cagliari, Italy), for their support. Eleonora Sulas is grateful to Sardinia Regional Government for supporting her PhD scholarship (P.O.R. F.S.E., European Social Fund 2014–2020).

Part of this research was supported by the Italian Government—Progetti di Interesse Nazionale (PRIN) under the grant agreement 2017RR5EW3 - ICT4MOMs project.

## Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.dib.2020.106399](https://doi.org/10.1016/j.dib.2020.106399).

## References

- [1] G. Baldazzi, E. Sulas, M. Urru, R. Tumbarello, L. Raffo, D. Pani, Wavelet denoising as a post-processing enhancement method for non-invasive foetal electrocardiography, *Comput. Methods Programs Biomed.* 195 (2020) 10558, doi:[10.1016/j.cmpb.2020.105558](https://doi.org/10.1016/j.cmpb.2020.105558).
- [2] J. Behar, F. Andreotti, S. Zauneder, Q. Li, J. Oster, G.D. Clifford, An ECG simulator for generating maternal-foetal activity mixtures on abdominal ECG recordings, 2014, doi:[10.1088/0967-3334/35/8/1537](https://doi.org/10.1088/0967-3334/35/8/1537).
- [3] F. Andreotti, J. Behar, S. Zauneder, J. Oster, G.D. Clifford, An open-source framework for stress-testing non-invasive foetal ECG extraction algorithms, *Physiol. Meas.* 37 (2016) 627.
- [4] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiological signals, *Circulation* 101 (2000) e215–e220, doi:[10.1161/01.cir.101.23.e215](https://doi.org/10.1161/01.cir.101.23.e215).
- [5] R. Sameni, The Open-Source Electrophysiological Toolbox (OSET), Version 3.14, 2018, <https://gitlab.com/rsameni/OSET/>.
- [6] T.F. Oostendorp, A. Van Oosterom, H.W. Jongsma, The effect of changes in the conductive medium on the fetal ECG throughout gestation, *Clin. Phys. Physiol. Meas.* (1989), doi:[10.1088/0143-0815/10/4B/002](https://doi.org/10.1088/0143-0815/10/4B/002).
- [7] T.F. Oostendorp, A. van Oosterom, H.W. Jongsma, Electrical properties of tissues involved in the conduction of foetal ECG, *Med. Biol. Eng. Comput.* 27 (1989) 322–324, doi:[10.1007/BF02441492](https://doi.org/10.1007/BF02441492).
- [8] J.A. Apolinario JR, *QRD-RLS Adaptive Filtering*, Springer, 2009.
- [9] E. Sulas, M. Urru, R. Tumbarello, L. Raffo, D. Pani, Systematic analysis of single- and multi-reference adaptive filters for non-invasive fetal electrocardiography, *Math. Biosci. Eng.* 17 (n.d.) 286–308, doi:[10.3934/mbe.2020016](https://doi.org/10.3934/mbe.2020016).
- [10] D. Pani, S. Argiolas, L. Raffo, A DSP algorithm and system for real-time fetal ECG extraction, *Comput. Cardiol.* (2008) 1065–1068, doi:[10.1109/CIC.2008.4749229](https://doi.org/10.1109/CIC.2008.4749229).
- [11] D. Pani, G. Barabino, L. Raffo, NinFEA: an embedded framework for the real-time evaluation of fetal ECG extraction algorithms, *Biomed. Technol. Eng.* 58 (2013) 13, doi:[10.1515/bmt-2012-0018](https://doi.org/10.1515/bmt-2012-0018).
- [12] M.G. Signorini, G. Lanzola, E. Torti, A. Fanelli, G. Magenes, Antepartum fetal monitoring through a wearable system and a mobile application, *Technologies* 6 (2018) 44, doi:[10.3390/technologies6020044](https://doi.org/10.3390/technologies6020044).
- [13] J. Behar, A. Johnson, G.D. Clifford, J. Oster, A comparison of single channel fetal ECG extraction methods, *Ann. Biomed. Eng.* 42 (2014) 1340–1353, doi:[10.1007/s10439-014-0993-9](https://doi.org/10.1007/s10439-014-0993-9).
- [14] P. Kligfield, L.S. Gettes, J.J. Bailey, R. Childers, B.J. Deal, E.W. Hancock, G. van Herpen, J.A. Kors, P. Macfarlane, D.M. Mirvis, O. Pahlm, P. Rautaharju, G.S. Wagner, M. Josephson, J.W. Mason, P. Okin, B. Surawicz, H. Wellens, Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology a scientific statement from the American heart association electrocardiography and arrhythmias committee, council on clinical cardiology, *J. Am. Coll. Cardiol.* 49 (2007) 1109–1127, doi:[10.1016/j.jacc.2007.01.024](https://doi.org/10.1016/j.jacc.2007.01.024).