Hierarchical probabilistic framework for fetal R-peak detection, using ECG waveform and heart rate information

Hierarchical probabilistic framework for fetal R-peak detection, using ECG waveform and heart rate information

G.J.J. Warmerdam, *Member, IEEE,* R. Vullings, L. Schmitt, J.O.E.H. Van Laar, and J.W.M. Bergmans, *Senior Member, IEEE*

***Abstract*—The abdominal fetal electrocardiogram (fECG) can provide valuable information about fetal well-being. However, fetal R-peak detection in abdominal fECG recordings is challenging due to the low signal-to-noise ratio (SNR) and the non-stationary nature of the fECG waveform in the abdominal recordings. In this paper, we propose a multichannel hierarchical probabilistic framework for fetal R-peak detection, that combines predictive models of the ECG waveform and the heart rate. The performance of our method was evaluated on set-A of the 2013 Physionet/Computing in Cardiology Challenge and compared to the performance of several methods that have been proposed in the literature. The hierarchical probabilistic framework presented in this study outperforms other methods for fetal R-peak detection with a mean overall detection accuracy for set-A of 99.6%. Even for recordings with low SNR our method enables reliable fetal R-peak detection (Ac 99.4%).**

***Index Terms*—Hierarchical Bayesian model, Kalman filtering, fetal electrocardiography, R-peak detection**

I. Introduction

Since the 1960s, the most widely used technique for fetal monitoring is cardiotocography (CTG) [1]. CTG provides simultaneous information on the fetal heart rate (HR) and uterine activity. However, interpretation of CTG suffers from a low specificity, resulting in unnecessary operative deliveries [2].

To obtain additional information on fetal well-being in case of an abnormal CTG, the fetal electrocardiogram (ECG) could be used [3]. From the fetal ECG (fECG) it is possible to extract beat-to-beat fetal HR information that is required for reliable analysis of fetal HR variability [4]. Moreover, analysis of the fECG waveform could provide information on fetal oxygen deficiency [5]. Both for HR variability analysis and fECG waveform analysis accurate R-peak detection is required.

Generally, the fECG is obtained invasively using an electrode attached to the fetal scalp [5]. Although the signal quality of invasive fECG is good, it can only be used during delivery after the fetal membranes have ruptured. An alternative that can also be used earlier in the pregnancy is to measure the fECG non-invasively by electrodes placed on the maternal abdomen [6].

G.J.J. Warmerdam, R. Vullings, and J.W.M. Bergmans are with the Faculty of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands.

L. Schmitt is with Philips Research, Eindhoven, The Netherlands.

J.O.E.H. Van Laar is with the Máxima Medical Center, Veldhoven, The Netherlands.

The low invasiveness of the abdominal fECG comes at a cost of a reduction in signal-to-noise ratio (SNR) [7]. The abdominal fECG is contaminated by electrical interferences such as the maternal ECG (mECG), muscle activity, power line interference, and measurement noise. Moreover, in the period between 28 to 32 weeks of gestation, an isolating layer (the vernix caseosa) surrounds the fetus and reduces the amplitude and affects the shape of the abdominal fECG [8].

In recent years, abdominal fECG recordings have been extensively studied, most studies focussing on suppression of the mECG, which is the dominant interference [6], [9]–[14]. In 2013, the aim of the PhysioNet/Computing in Cardiology Challenge (further referred to as Challenge) was to extract the fetal HR from abdominal recordings [15]. A variety of algorithms was presented for mECG suppression, such as template subtraction [6], [9], [10], adaptive filtering [11], [12], blind source separation (BSS) [13], [14], [16], or a combination of different algorithms [17]–[19]. For an extensive review see [7] or [20]. To compare different algorithms, a database of abdominal fECG recordings was made publicly available.

After mECG suppression, fetal R-peak detection is often performed by adapting existing algorithms for adult R-peak detection [18], [19], [21]. However, even after mECG suppression, the SNR of the abdominal fECG is generally still much lower than the SNR for adult ECG recordings. Since algorithms for adult R-peak detection are optimized for relatively high SNR, this can lead to numerous mis-detections for the low-SNR abdominal fECG.

Besides low SNR, the position and orientation of the fetus within the abdomen are a priori unknown and can change during a recording. Therefore, the abdominal fECG is typically recorded using multiple electrodes spread across the abdomen [22]. The SNR and waveform of the fECG in each channel depends on the fetal position and orientation. Hence, fetal movement with respect to the abdominal electrodes can cause variations in the SNR and fECG waveform of a certain channel [23]. Despite this fact, several studies performed R-peak detection on each individual channel after which a post-processing step was used to select the channel with the best RR-series [16]–[18]. Although in [19] a multi-channel matched filter approach was used for fetal R-peak detection, changes in fECG waveform were not considered for the matched filter. Both approaches can lead to reduced performance in case of fetal movement.

In short, the low SNR and the non-stationary nature of the

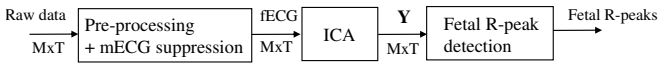


Fig. 1. Block diagram of fetal heart rate detection. M is the number of channels and T is the total time of the recording.

abdominal fECG make fetal R-peak detection challenging. In this study, we propose an adaptive multi-channel R-peak detection method that combines ECG waveform and HR information. A schematic overview of the detection algorithm is shown in Fig. 1. The paper is structured as follows: first, pre-processing and mECG suppression are discussed in section II-A. Then, the fECG and HR model are explained in section II-B, and a hierarchical probabilistic framework for R-peak detection is discussed in sections II-C to II-H. Extension to multiple channels is explained in section II-I. Finally, results and discussion are presented in sections IV and V.

II. Methods

A. Pre-processing and mECG suppression

Analysis of the results of the Challenge showed that the best performing algorithms used a similar approach to suppress the mECG [17]–[19]. In this approach, a matrix of mECG complexes is created, with each row corresponding to one mECG complex. Then, a Principal Component Analysis based approach is used to extract the most significant eigenvectors of this matrix. These eigenvectors contain information about the average mECG complex and morphological variations in the mECG complexes (e.g. due to respiration). The largest eigenvectors are used to estimate the mECG and subtract it from the original signal.

In this study, we used the algorithm of Varanini et al. [17] for mECG suppression, that is online available [15]. Note that in this algorithm the data is pre-processed to suppress baseline wandering, high frequency noise, and power line interference. After pre-processing and mECG suppression, Independent Component Analysis (ICA) is used to further enhance the fECG. Since our study focusses on fetal R-peak detection, the method of Varanini is not discussed in detail and we will use the signals after mECG suppression and ICA ($\boldsymbol{Y}$) as starting point for our fetal R-peak detection method. An example of the original signals, the fECG after pre-processing and mECG suppression, and $\boldsymbol{Y}$ are shown in Fig. 2.

B. fECG model

We now proceed with a single channel approach (sections II-B-II-H) and then extend our model to multiple channels in section II-I. A single ECG signal after pre-processing, mECG suppression, and ICA is denoted as $y(t)$, where t is a time index. The location of the k-th fetal R-peak is denoted as μ_k, and the k-th RR-interval as $w_k = \mu_k - \mu_{k-1}$.

1) Gaussian QRS model: In abdominal fECG recordings, the amplitude of the fetal QRS complex is generally large compared to other segments of the fECG (the P-, and T-wave). Therefore, our model for R-peak detection is limited to describing the fetal QRS complex. It is important that the QRS model is flexible, because it is a priori unknown what the fetal orientation and location is with respect to the abdominal electrodes.

Our QRS model is inspired by a model that was proposed by McSharry et al., which uses Gaussian kernels to describe the fECG [24], [25]. A disadvantage of the model of McSharry is that a relatively large number of parameters is required to described the QRS complex (nine parameters in total). Moreover, the locations of the Q- and S-peak need to be determined with respect to the R-peak.

As noted by Biglari et al., the abdominal fECG is recorded at far-field and the morphology of the QRS complex is not as diverse and complex as for the adult ECG [26]. In our approach, the QRS complex is modeled by the sum of a Gaussian (mainly modeling the R-wave), its first derivative (mainly modeling the Q- and S-wave), and its second derivative (modeling the Q-, R-, and S-wave). Examples of these functions are shown in Fig. 3. The combination of these functions can be used to describe the QRS complex for most abdominal fECG recordings [26].

Assuming that the Gaussian, its first, and second derivative are centered around time μ_k, our QRS model is written as

$$G(t, \mu_k, \boldsymbol{z}) = \Big(a_1 + a_2(t-\mu_k) + a_3\big(1 - \frac{(t-\mu_k)^2}{b^2}\big)\Big) e^{\frac{-(t-\mu_k)^2}{2b^2}}, \tag{1}$$

where $\boldsymbol{z} = [a_1, a_2, a_3, b]$, and $a_{1,2,3}$ are the amplitudes of the Gaussian, its first and its second derivative. Note that for the first and second derivative a factor $1/b^2$ is included in terms a_2 and a_3, respectively. The QRS model in Eq. 1 depends on four parameters and does not require knowledge on the location of the Q- and S-peak.

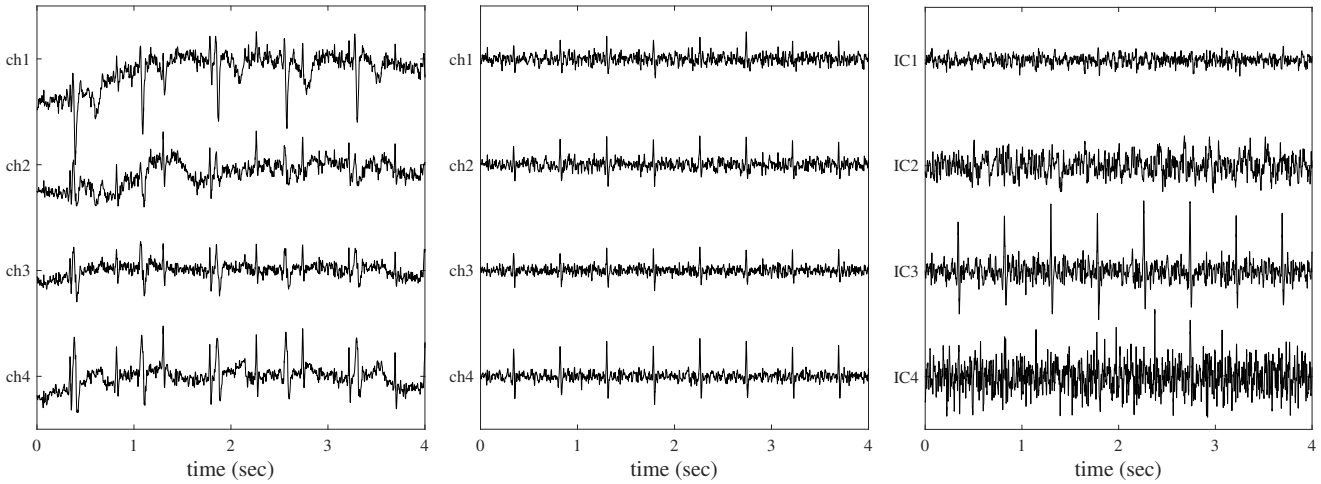
Besides containing the fetal QRS complex, $y(t)$ will be contaminated by electrical interferences, such as muscle artifacts, remainders of the mECG, and measurement noise. These inferences are represented by an additive noise term ξ_t, which will be referred to as the observation noise of the QRS model. Because ξ_t is a combination of several signals, including measurement noise, we will assume ξ_t to have a zero-mean Gaussian distribution with variance λ_t. Given the location of the $k+1$-th R-peak, μ_{k+1}, we can write $y(t)$ during the $k+1$-th QRS complex as:

$$y(t) = G(t, \mu_{k+1}, \boldsymbol{z}) + \xi_t \qquad \xi_t \sim \mathcal{N}(0, \lambda_t). \tag{2}$$

Since for each (regular) heartbeat the electrical activity of the fetal heart propagates in a similar way through the heart, the electrical activity will be similar across heartbeats. However, the electrical activity measured by the abdominal electrodes can vary over time due to fetal movement with respect to the electrodes. Assuming that during a ventricular contraction the fetal orientation remains constant, we can express the model parameters of the $k+1$-th QRS complex as

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \boldsymbol{\eta}_k \qquad \boldsymbol{\eta}_k \sim \mathcal{N}(0, \boldsymbol{\Sigma}_k). \tag{3}$$

Here, variations in $\boldsymbol{z}_k$ are assumed to be described by a zero-mean Gaussian random walk process $\boldsymbol{\eta}_k$, with $[4 \times 4]$ covariance matrix $\boldsymbol{\Sigma}_k$. We refer to $\boldsymbol{\eta}_k$ as the process noise of the QRS model.



(a) Raw signals (mECG and fECG). (b) fECG. (c) fECG after ICA ($\boldsymbol{Y}$).

Fig. 2. Signals before and after pre-processing, mECG suppression, and ICA.

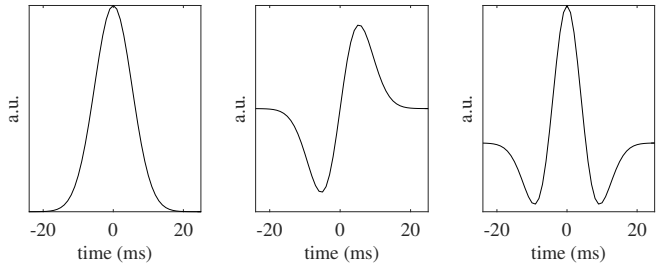


Fig. 3. Example QRS model. From left to right: Gaussian, first derivative, and second derivative.

2) Time-varying autoregressive HR model: To describe the fetal HR, we use a p-th order time-varying autoregressive (TVAR) model [27], [28]. In the TVAR model the $k+1$-th RR-interval w_{k+1} is described as a linear combination of p previously detected RR-intervals $\boldsymbol{w}_k = [w_k, w_{k-1}, ..., w_{k-p+1}]^T$. Since $w_{k+1} = \mu_{k+1} - \mu_k$, we can use this model to predict the location of μ_{k+1}, given previous RR-intervals and μ_k. The order of the TVAR model was empirically determined and set to $p = 6$. The TVAR model is written as

$$w_{k+1} = \boldsymbol{w}_k^T \boldsymbol{\theta} + v_{k+1} \qquad v_{k+1} \sim \mathcal{N}(0, R_{k+1}), \tag{4}$$

with $\boldsymbol{\theta} = [\theta_1, ..., \theta_p]^T$ the TVAR model parameters. In Eq. 4, any variation in the HR that cannot be captured by the linear TVAR model is represented by v_{k+1}. We refer to v_{k+1} as the observation noise of the HR model and it is assumed to be zero-mean Gaussian with variance R_{k+1}.

Because variations in the HR are regulated by the autonomic nervous system and the autonomic regulation varies over time (e.g. due to changes in fetal behavioral states [29]), parameters $\boldsymbol{\theta}$ are also time-varying. We will assume that variations in $\boldsymbol{\theta}$ can be modeled by a zero mean Gaussian $\boldsymbol{d}_k$, with $[p \times p]$ covariance $\boldsymbol{Q}_k$:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \boldsymbol{d}_k \qquad \boldsymbol{d}_k \sim \mathcal{N}(0, \boldsymbol{Q}_k). \tag{5}$$

The noise term $\boldsymbol{d}_k$ is referred to as the process noise of the HR model.

C. Hierarchical probabilistic framework

Given μ_k and the estimated $k+1$-th RR-interval $\hat{w}_{k+1}$, we will look for the next R-peak μ_{k+1} within a predefined search window. The search window is limited to the interval $\boldsymbol{t} = \mu_k + \hat{w}_{k+1} \pm \frac{T}{2}$, with T the width of the search window (set to 400 ms). The search window is further limited by a minimum and maximum RR-interval (RR_{min} and RR_{max}, respectively), that are defined based on a range of the fetal HR between 50 and 210 beats per minute (bpm) [5]. The observation noise in $\boldsymbol{y}_{k+1}$ is described by $\boldsymbol{\xi}_{k+1}$, with covariance $\boldsymbol{\Lambda}_{k+1}$. Notice that with respect to Eq. 2, we have changed from scalar to vector notation.

Based on this description of $\boldsymbol{y}_{k+1}$ and 4 we can define a state-space model for the $k+1$-th QRS complex:

$$\begin{aligned} \mu_{k+1} &= \mu_k + \boldsymbol{w}_k^T \boldsymbol{\theta}_{k+1} + v_{k+1} & v_{k+1} &\sim \mathcal{N}(0, R_{k+1}) \\ \boldsymbol{y}_{k+1} &= G(\boldsymbol{t}, \mu_{k+1}, \boldsymbol{z}_{k+1}) + \boldsymbol{\xi}_{k+1} & \boldsymbol{\xi}_{k+1} &\sim \mathcal{N}(0, \boldsymbol{\Lambda}_{k+1}), \end{aligned} \tag{6}$$

Given $\boldsymbol{y}_{k+1}$, we are interested in estimating the new R-peak location μ_{k+1}. The uncertainties in the state-space model in Eq. 6 suggest that a probabilistic approach can be used to solve the estimation problem for μ_{k+1}. To find a tractable solution for inferring μ_{k+1}, we propose a hierarchical pseudo-Bayesian framework that consists of three inference levels:

Level 1: State estimation. In this level, we estimate μ, while assuming that $\boldsymbol{\theta}$, $\boldsymbol{z}$, and covariances $\boldsymbol{Q}$, R, $\boldsymbol{\Sigma}$, and $\boldsymbol{\Lambda}$ are known.

Level 2: QRS and HR model estimation. In this level, we estimate $\boldsymbol{\theta}$ and $\boldsymbol{z}$, while assuming μ, $\boldsymbol{Q}$, R, $\boldsymbol{\Sigma}$, and $\boldsymbol{\Lambda}$ to be known.

Level 3: Noise estimation. In this level, we estimate $\boldsymbol{Q}$, R, $\boldsymbol{\Sigma}$, and $\boldsymbol{\Lambda}$, while assuming μ, $\boldsymbol{\theta}$, and $\boldsymbol{z}$ to be known.

A schematic overview of the inference is presented in Fig. 4. Note that we used different inference strategies to estimate the model parameters (as summarized in Table II-C), hence our hierarchical model is not fully Bayesian.

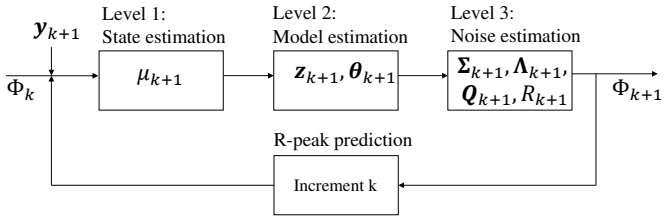


Fig. 4. Schematic overview of the hierarchical model.

TABLE I
OVERVIEW OF MODEL PARAMETERS (Φ) AND INFERENCE STRATEGY FOR PARAMETER ESTIMATION.

Parameter	Estimation strategy	Section
μ	MAP	II-D, II-G
$\boldsymbol{z}$	EKF	II-E1
$\boldsymbol{\theta}$	KF	II-E2
$\boldsymbol{\Sigma}$	empirically determined*	II-F1
$\boldsymbol{\Lambda}$	from signal	II-F2
$\boldsymbol{Q}$	maximum likelihood	II-F3
$\boldsymbol{R}$	empirically determined*	II-F4

* Values were empirically determined based on the optimization of the detection performance for set-A of the Challenge.

D. Level 1: State estimation

Given the estimated model parameters after observing the k-th QRS complex and new observation $\boldsymbol{y}_{k+1}$, we want to predict the location of the $k+1$-th QRS complex. Using Bayes' rule we can describe the *posterior* probability density function for μ_{k+1} as

$$p(\mu_{k+1}|\boldsymbol{y}_{k+1}, \Phi_k) = \frac{p(\boldsymbol{y}_{k+1}|\mu_{k+1}, \Phi_k)p(\mu_{k+1}|\boldsymbol{y}_k, \Phi_k)}{p(\boldsymbol{y}_{k+1}|\boldsymbol{y}_k, \Phi_k)} \quad (7)$$

with $\Phi_k = \{\mu_k, \boldsymbol{z}_k, \boldsymbol{\theta}_k, \boldsymbol{w}_k, \boldsymbol{\Sigma}_k, \boldsymbol{\Lambda}_k, \boldsymbol{Q}_k, R_k\}$ our prior information about the model parameters from the previous iteration.

Because $p(\boldsymbol{y}_{k+1}|\boldsymbol{y}_k, \Phi_k)$ is a normalization factor that is independent of μ_{k+1}, the maximum *a posteriori* (MAP) estimate for μ_{k+1} can be written as:

$$\hat{\mu}_{k+1} = \underset{\mu_{k+1}}{\operatorname{argmax}} \big(\ln p(\boldsymbol{y}_{k+1}|\mu_{k+1}, \Phi_k) + \ln p(\mu_{k+1}|\boldsymbol{y}_k, \Phi_k) \big), \quad (8)$$

where the use of the log-posterior is justified by the monotonic behavior of the logarithm, which will not influence the MAP solution. In the next levels $\hat{\mu}_{k+1}$ is used to estimate the model parameters and noise covariances.

E. Level 2: QRS and HR model estimation

Given observation $\boldsymbol{y}_{k+1}$ and the location of the $k+1$-th R-peak at $\hat{\mu}_{k+1}$, we want to update our model parameters $\boldsymbol{\theta}$ and $\boldsymbol{z}$. We will assume that the HR does not influence the QRS waveform, and consider $\boldsymbol{\theta}$ and $\boldsymbol{z}$ to be independent.

1) Extended Kalman filter to estimate QRS model: Using the QRS model that is explained in section II-B1, we can describe the waveform of the $k+1$-th QRS complex by the following state-space equations:

$$\begin{aligned} \boldsymbol{z}_{k+1} &= \boldsymbol{z}_k + \boldsymbol{\eta}_k & \boldsymbol{\eta}_k &\sim \mathcal{N}(0, \boldsymbol{\Sigma}_k) \\ \boldsymbol{y}_{k+1} &= G(\boldsymbol{t}, \hat{\mu}_{k+1}, \boldsymbol{z}_{k+1}) + \boldsymbol{\xi}_{k+1} & \boldsymbol{\xi}_{k+1} &\sim \mathcal{N}(0, \boldsymbol{\Lambda}_{k+1}) \end{aligned} \quad (9)$$

A tractable approximation for the MAP estimate $\hat{\boldsymbol{z}}_{k+1}$ for $\boldsymbol{z}_{k+1}$ of this non-linear state-space model can be calculated using an Extended Kalman Filter (EKF) [30]. In the EKF formalism, a first order approximation is used to describe the non-linear function $G(\boldsymbol{t}, \hat{\mu}_{k+1}, \boldsymbol{z}_k)$. Note that the EKF provides the minimum mean square error estimate, which is not necessarily identical to the MAP estimate in case of a non-linear model. The EKF update equations are given by [30]:

$$\hat{\boldsymbol{z}}_{k+1} = \hat{\boldsymbol{z}}_k + \boldsymbol{K}_{k+1}^{QRS}(\boldsymbol{y}_{k+1} - G(\boldsymbol{t}, \hat{\mu}_{k+1}, \hat{\boldsymbol{z}}_k)) \quad (10)$$

$$\boldsymbol{P}_{k+1} = \boldsymbol{P}_k + \boldsymbol{\Sigma}_k - \boldsymbol{K}_{k+1}^{QRS} J_{k+1}^T (\boldsymbol{P}_k + \boldsymbol{\Sigma}_k) \quad (11)$$

$$\boldsymbol{K}_{k+1}^{QRS} = (\boldsymbol{P}_k + \boldsymbol{\Sigma}_k)\boldsymbol{J}_{k+1} \cdot (\boldsymbol{J}_{k+1}^T(\boldsymbol{P}_k + \boldsymbol{\Sigma}_k)\boldsymbol{J}_{k+1} + \boldsymbol{\Lambda}_{k+1})^{-1}, \quad (12)$$

with $\boldsymbol{P}_k$ the $[4 \times 4]$ covariance in $\boldsymbol{z}_k$, $\boldsymbol{K}_{k+1}^{QRS}$ the $[4 \times T]$ Kalman gain of the QRS model, and $\boldsymbol{J}_{k+1}$ the $[4 \times T]$ Jacobian matrix $\boldsymbol{J}_{k+1} = \frac{\partial G}{\partial \boldsymbol{z}}\Big|_{(\boldsymbol{z}=\hat{\boldsymbol{z}}_k)}$.

2) Kalman filter to estimate HR model: Using the HR model that is explained in section II-B2, the $k+1$-th RR-interval can be described by the following state-space equations:

$$\begin{aligned} \boldsymbol{\theta}_{k+1} &= \boldsymbol{\theta}_k + \boldsymbol{d}_k & \boldsymbol{d}_k &\sim \mathcal{N}(0, \boldsymbol{Q}_k) \\ w_{k+1} &= \boldsymbol{w}_k^T \boldsymbol{\theta}_{k+1} + v_{k+1} & v_{k+1} &\sim \mathcal{N}(0, R_{k+1}) \end{aligned} \quad (13)$$

Following a similar derivation as in [31], it is possible to show that the MAP estimate $\hat{\theta}_{k+1}$ for $\boldsymbol{\theta}_{k+1}$ can be calculated using the linear Kalman filter update equations:

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \boldsymbol{K}_{k+1}^{HR}(w_{k+1} - \boldsymbol{w}_k^T \hat{\boldsymbol{\theta}}_k) \quad (14)$$

$$\boldsymbol{V}_{k+1} = \boldsymbol{V}_k + \boldsymbol{Q}_k - \boldsymbol{K}_{k+1}^{HR} \boldsymbol{w}_k^T (\boldsymbol{V}_k + \boldsymbol{Q}_k) \quad (15)$$

$$\boldsymbol{K}_{k+1}^{HR} = \frac{(\boldsymbol{V}_k + \boldsymbol{Q}_k)\boldsymbol{w}_k}{\boldsymbol{w}_k^T (\boldsymbol{V}_k + \boldsymbol{Q}_k)\boldsymbol{w}_k + R_{k+1}} \quad (16)$$

where $\boldsymbol{V}_k$ is the $[p \times p]$ covariance of $\boldsymbol{\theta}_k$ and $\boldsymbol{K}_{k+1}^{HR}$ is the $[p \times 1]$ Kalman gain for the HR model.

F. Level 3: Noise estimation

1) Process noise $\boldsymbol{\Sigma}_{k+1}$ of QRS model: Variations in the QRS waveform due to e.g. fetal movement are described by the process noise $\boldsymbol{\Sigma}_{k+1}$. We assume identical and uncorrelated process noise for all QRS model parameters $\boldsymbol{a} = [a_1, ..., a_3]^T$. The process noise variance σ_a^2 is assumed constant, and is defined as a fraction of the maximum of the squared average QRS complex in the initialization phase (see section II-H): $\sigma_a^2 = c_a \max |QRS|^2$, where fraction c_a is empirically determined. Furthermore, although the width of the fetal QRS complex can vary during pregnancy, fetal movement will have little effect on the width. Hence the width is assumed to remain constant throughout a recording. Therefore, we assume zero process noise for parameter b, which means that b is estimated recursively. The process noise $\boldsymbol{\Sigma}_{k+1}$ for $\boldsymbol{z}_{k+1}$ is thus described as

$$\boldsymbol{\Sigma}_{k+1} = \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_a^2 & & & 0 \\ & \sigma_a^2 & & \\ & & \sigma_a^2 & \\ 0 & & & 0 \end{bmatrix}. \quad (17)$$

2) Observation noise $\boldsymbol{\Lambda}_{k+1}$ of QRS model: To estimate the covariance $\boldsymbol{\Lambda}_{k+1}$, we will assume that the observation noise is uncorrelated and constant for observation $\boldsymbol{y}_{k+1}$. Note that since $\boldsymbol{\xi}_{k+1}$ also describes correlated physiological noise (i.e. muscle activity) this assumption is only approximately true. The covariance matrix can then be written as $\boldsymbol{\Lambda}_{k+1} = \lambda_{k+1}\boldsymbol{I}_T$. with $\boldsymbol{I}_T$ the $[T \times T]$ identity matrix.

The estimation of λ_{k+1} is based on the variance in $\boldsymbol{y}_{k+1}$. Typically, the duration of the fetal QRS complex is about $L = 40$ ms [32], which is about 10% of a fECG complex (assuming a heart rate of 150 bpm). In the state estimation in level 1 (where the location of μ_{k+1} is estimated), the contribution of the QRS complex to the variance of $\boldsymbol{y}_{k+1}$ is accounted for by omitting the top 10% of the values of $\boldsymbol{y}_{k+1}$. For the estimation of $\boldsymbol{z}_{k+1}$ in level 2 we update our estimate of λ_{k+1} by omitting samples of $\boldsymbol{y}_{k+1}$ within the region $\hat{\mu}_{k+1} \pm \frac{L}{2}$. The value of λ_{k+1} is determined as the variance in $\boldsymbol{y}_{k+1}$ over the remaining samples.

3) Process noise $\boldsymbol{Q}_{k+1}$ of HR model: In a hierarchical Bayesian model, the likelihood function at a particular level is equal to the evidence at the next higher level [31]. This means that by maximizing the evidence of the HR model in level 2 we can obtain the maximum likelihood solution for $\boldsymbol{Q}$.

The evidence of the HR model is given by:

$$p(w_{k+1}|\boldsymbol{y}_{k+1}, \Phi_k) = \mathcal{N}(w_{k+1}|\boldsymbol{w}_k^T\hat{\boldsymbol{\theta}}_k, \boldsymbol{w}_k^T(\boldsymbol{Q}_k + \boldsymbol{V}_k)\boldsymbol{w}_k + R_{k+1}). \quad (18)$$

The value of $\hat{\boldsymbol{Q}}_{k+1}$ can be calculated by differentiating Eq. 18 with respect to $\boldsymbol{Q}$ and equating it to zero. If we assume that $\boldsymbol{Q}_{k+1} = q_{k+1}\boldsymbol{I}_p$ and define the residual error of the HR model as $r_{k+1} = w_{k+1} - \boldsymbol{w}_k^T\hat{\boldsymbol{\theta}}_k$, it can be shown that the maximum likelihood estimate for q_{k+1} is [31]:

$$\hat{q}_{k+1} = \begin{cases} \frac{r_{k+1}^2 - \mathbb{E}[r_{k+1}^2|q_{k+1}=0]}{\boldsymbol{w}_k^T\boldsymbol{w}_k} & \text{if } q_{k+1} \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

Here $\mathbb{E}[r_{k+1}^2|q_{k+1} = 0] = \boldsymbol{w}_k^T\boldsymbol{V}_k\boldsymbol{w}_k + R_{k+1}$, is the covariance of r_{k+1} if we assume that $q_{k+1} = 0$. Eq. 19 implies that, if r_{k+1} is larger than the predicted variance, the value of q_{k+1} increases. If q_{k+1} increases, the Kalman gain $\boldsymbol{K}^{HR}$ also increases and a new estimate of $\boldsymbol{\theta}$ will depend more on incoming data.

4) Observation noise R_{k+1} of HR model: Following discussions on HR variability, we know that variations in the HR often cannot be fully described by a linear model [33]. This means that even in case we have optimally inferred $\boldsymbol{\theta}$, we can only approximate a new RR-interval up to a certain degree. Any stochastic or non-linear variation in the HR is described by R_{k+1}.

Since the value of R_{k+1} defines how well we expect our linear model to describe the HR, we will assume that $R_{k+1} = R_0$ is constant. A low value of R_0 leads to relatively high values of $\boldsymbol{Q}$ through Eq. 19. This means that $\boldsymbol{\theta}$ will be adapted even for a small residual error. A large value of R_0 leads to relatively low values of $\boldsymbol{Q}$. In this case adaptation of $\boldsymbol{\theta}$ is reduced, which allows for larger discrepancies between the predicted HR and the observed HR. The value for R_0 was empirically determined.

G. R-peak prediction

We can use our estimates for $\boldsymbol{z}_k$ and $\boldsymbol{\theta}_k$ at the k-th iteration to predict the R-peak location at the $k+1$-th iteration. Accounting for the covariances in our estimation of $\boldsymbol{z}_k$ and $\boldsymbol{\theta}_k$ (denoted by $\boldsymbol{P}_k$ and $\boldsymbol{V}_k$, respectively), and making use of the state space model in Eq. 6, we can write:

$$p(\mu_{k+1}|\boldsymbol{y}_k, \Phi_k) = \mathcal{N}(\mu_{k+1}|\hat{\mu}_k + \hat{w}_{k+1}, \Gamma^{HR}), \quad (20)$$

$$p(\boldsymbol{y}_{k+1}|\mu_{k+1}, \Phi_k) = \mathcal{N}(\boldsymbol{y}_{k+1}|\hat{\boldsymbol{y}}_{k+1}, \boldsymbol{\Gamma}^{QRS}), \quad (21)$$

with

$$\hat{w}_{k+1} = \boldsymbol{w}_k^T\hat{\boldsymbol{\theta}}_k, \quad (22)$$

$$\Gamma^{HR} = \boldsymbol{w}_k^T(\boldsymbol{V}_k + \boldsymbol{Q}_k)\boldsymbol{w}_k + R_{k+1}, \quad (23)$$

$$\hat{\boldsymbol{y}}_{k+1} = G(\boldsymbol{t}, \mu_{k+1}, \hat{\boldsymbol{z}}_k), \text{ and} \quad (24)$$

$$\boldsymbol{\Gamma}^{QRS} = \boldsymbol{J}_{k+1}^T(\boldsymbol{P}_k + \boldsymbol{\Sigma}_k)\boldsymbol{J}_{k+1} + \boldsymbol{\Lambda}_{k+1}. \quad (25)$$

Hence, the log-posterior for μ_{k+1} in Eq. 8 can be written as:

$$\mathcal{L} = \text{const.} - \frac{(\mu_{k+1} - \hat{\mu}_k - \hat{w}_{k+1})^2}{\Gamma^{HR}} - (\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1})^T\boldsymbol{\Gamma}^{QRS^{-1}}(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1}). \quad (26)$$

Notice that $\hat{\boldsymbol{y}}_{k+1}$ depends on μ_{k+1} through the function $G(\boldsymbol{t}, \mu_{k+1}, \hat{\boldsymbol{z}}_k)$. Furthermore, $\mathcal{L}$ can be multimodal and is necessarily not Gaussian.

Regardless of the signal quality, it is always possible to find a $\hat{\mu}_{k+1}$ that maximizes $\mathcal{L}$. However, in case of poor signal quality $\hat{\mu}_{k+1}$ might be inaccurate, leading to mis-detections of the R-peak. To prevent mis-detection, a sanity check is performed during which we compare the probability that $\hat{\mu}_{k+1}$ is the location of an R-peak to the probability that $\hat{\mu}_{k+1}$ is not the location of an R-peak.

Focussing on the samples during a QRS complex $\boldsymbol{i} = \hat{\mu}_{k+1} \pm \frac{L}{2}$, we can compare the likelihood for $\boldsymbol{y}_i = y(\boldsymbol{i})$ in case $\hat{\mu}_{k+1}$ is an R-peak to the likelihood for $\boldsymbol{y}_i$ in case $\hat{\mu}_{k+1}$ is not an R-peak. If $\hat{\mu}_{k+1}$ is an R-peak, the likelihood for $\boldsymbol{y}_i$ is $p(\boldsymbol{y}_i|\hat{\mu}_{k+1} = \text{R-peak}) = \mathcal{N}(\boldsymbol{y}_i|\hat{\boldsymbol{y}}_i, \boldsymbol{\Gamma}^{QRS})$. If $\hat{\mu}_{k+1}$ is not the location of an R-peak, the likelihood for $\boldsymbol{y}_i$ is $p(\boldsymbol{y}_i|\hat{\mu}_{k+1} \neq \text{R-peak}) = \mathcal{N}(\boldsymbol{y}_i|0, \boldsymbol{\Lambda}_{k+1})$. This means that:

$$\hat{\mu}_{k+1} \begin{cases} = \text{R-peak, if } (\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i)^T\boldsymbol{\Gamma}_{k+1}^{QRS^{-1}}(\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i) < \boldsymbol{y}_i^T\boldsymbol{\Lambda}_{k+1}^{-1}\boldsymbol{y}_i \\ \neq \text{R-peak, otherwise} \end{cases} \quad (27)$$

If according to Eq. 27 $\hat{\mu}_{k+1}$ is the location of an R-peak, we use $\hat{\mu}_{k+1}$ in levels 2 and 3 to estimate the model parameters and noise covariances. In case $\hat{\mu}_{k+1}$ is not the location of an R-peak according to Eq. 27, the model parameters and noise covariances are not updated. Instead, we extrapolate an R-peak location based on the HR model and continue to predict μ_{k+2} given the prior information Φ_k (i.e. predict $p(\mu_{k+2}|\boldsymbol{y}_{k+2}, \Phi_k)$). In case more than five consecutive detected peaks were no R-peaks according to Eq. 27, the entire algorithm is re-initialized (see section II-H).

H. Parameter initialization

In the (re-)initialization phase, a segment of 10 seconds is used to obtain estimates for

$\Phi_0 = \{\mu_0, \boldsymbol{z}_0, \boldsymbol{\theta}_0, \boldsymbol{w}_0, \boldsymbol{\Lambda}_0, \boldsymbol{\Sigma}, \boldsymbol{Q}_0, R\}$. During initialization, the QRS model described by Eq. 1 is used as a frequency bandpass filter with a peak frequency of 42 Hz [21]. After filtering, a simple local maximum search is used for peak detection, where each search window is defined based on a previous maximum.

The detected peaks within the segment are used to calculate an average QRS complex, from which initial estimates $\boldsymbol{z}_0$ are obtained using a least-squares approach. Furthermore, RR-intervals are calculated from the detected peaks and RR-intervals outside a range of RR_{min} to RR_{max} are removed. On the remaining RR-intervals a Yule-Walker method is used to obtain initial estimates $\boldsymbol{\theta}_0$. The first R-peak is used as μ_0 and the first p RR-intervals were used as $\boldsymbol{w}_0$.

Initial covariances of the QRS model $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}_0$ are calculated as described in Section II-F1 and II-F2, and we initialized $\boldsymbol{P}_0 = 3\boldsymbol{\Sigma}$. Initial covariances of the HR model were set to $\boldsymbol{Q}_0 = 10^{-4} \max |\boldsymbol{\theta}_0|^2$, $R = R_0$, and $\boldsymbol{V}_0 = \boldsymbol{Q}_0$.

If forward prediction fails (i.e. more than five consecutive detected peaks were no R-peaks according to Eq. 27) and new parameters were found in re-initialization, we can use future information to obtain smoothed estimates of any missed (or mis-detected) R-peaks in the period before the start of the re-initialization. In essence, backwards prediction is similar to forward prediction, with the difference that the HR and QRS model are trained in the backwards direction. By combining the information of our forwards prediction and backwards prediction we obtain smoothed estimates for μ_{k+1}:

$$p(\mu_{k+1}|\boldsymbol{y}_{k+1}, \boldsymbol{y}_{k+2}, \Phi_k, \Phi_{k+2}) \propto p(\mu_{k+1}|\boldsymbol{y}_{k+1}, \Phi_k) \cdot p(\mu_{k+1}|\boldsymbol{y}_{k+2}, \Phi_{k+2}). \tag{28}$$

I. Multichannel extension

As described before, the abdominal ECG is commonly recorded from multiple channels. Because the ECG signals are correlated, recording the ECG on multiple channels can be exploited to improve the signal to noise conditions.

The model in Eq. 9 can be extended to M electrodes by training a QRS model to each channel. Assuming that the noise in each channel is uncorrelated, we can write

$$\begin{aligned} \boldsymbol{y}_{k+1}^{(1)} &= G(\boldsymbol{t}, \hat{\mu}_{k+1}, \boldsymbol{z}_{k+1}^{(1)}) + \boldsymbol{\xi}_{k+1}^{(1)} \qquad \boldsymbol{\xi}_{k+1}^{(1)} \sim \mathcal{N}(0, \boldsymbol{\Lambda}_{k+1}^{(1)}) \\ &\vdots = \\ \boldsymbol{y}_{k+1}^{(M)} &= G(\boldsymbol{t}, \hat{\mu}_{k+1}, \boldsymbol{z}_{k+1}^{(M)}) + \boldsymbol{\xi}_{k+1}^{(M)} \qquad \boldsymbol{\xi}_{k+1}^{(M)} \sim \mathcal{N}(0, \boldsymbol{\Lambda}_{k+1}^{(M)}), \end{aligned} \tag{29}$$

with $\boldsymbol{z}_{k+1}^{(m)} = [\boldsymbol{a}_{k+1}^{(m)}, b^{(m)}]$. Since the width of the QRS complex is similar for all channels, we only need to estimate one width parameter $b = b^{(1)}, ..., b^{(M)}$.

We can write a similar state-space model as the single channel ECG in Eq. 9 for the multichannel ECG:

$$\begin{aligned} \boldsymbol{z}_{k+1}^a &= \boldsymbol{z}_k^a + \boldsymbol{\eta}_k^a & \boldsymbol{\eta}_k^a &\sim \mathcal{N}(0, \boldsymbol{\Sigma}_k^a) \\ \boldsymbol{y}_{k+1}^a &= G(\boldsymbol{t}, \hat{\mu}_{k+1}, \boldsymbol{z}_{k+1}^a) + \boldsymbol{\xi}_{k+1}^a & \boldsymbol{\xi}_{k+1}^a &\sim \mathcal{N}(0, \boldsymbol{\Lambda}_{k+1}^a). \end{aligned} \tag{30}$$

Here, we defined the multichannel state and observation vectors as:

$$\boldsymbol{z}_k^a = [\boldsymbol{a}_k^{(1)}, ..., \boldsymbol{a}_k^{(M)}, b]^T \text{ and} \tag{31}$$

$$\boldsymbol{y}_{k+1}^a = [\boldsymbol{y}_{k+1}^{(1)}, ..., \boldsymbol{y}_{k+1}^{(M)}]^T, \tag{32}$$

and the multichannel noise covariances as:

$$\boldsymbol{\Sigma}_k^a = \begin{bmatrix} \sigma_a^2 \boldsymbol{I}_3 & & 0 \\ & \ddots & \\ 0 & & \sigma_a^2 \boldsymbol{I}_3 \end{bmatrix} \text{ and} \tag{33}$$

$$\boldsymbol{\Lambda}_k^a = \begin{bmatrix} \boldsymbol{\Lambda}_k^{(1)} & & 0 \\ & \ddots & \\ 0 & & \boldsymbol{\Lambda}_k^{(M)} \end{bmatrix}. \tag{34}$$

A similar approach can be used as the one discussed in section II-E1 to calculate the MAP estimate $\hat{\boldsymbol{z}}_{k+1}^a$ for $\boldsymbol{z}_{k+1}^a$.

Note that our algorithm uses ICA as a pre-processing step. ICA has the limitation that the Independent Components (ICs) are arbitrarily ordered. Therefore, it is a priori unknown which of the ICs contain information about the fECG. Despite this limitation of ICA, our multichannel model uses information from all ICs and no separate channel selection step is required.

J. Algorithms from the literature

The performance of our algorithm is compared to the performance of the algorithms of Varanini et al. [17] and Behar et al. [18], both scored high in the Challenge and have source codes that were online available. Besides algorithms from the Challenge, we also implemented the algorithm of Biglari et al. [26], because they used similar (fixed) templates for peak detection as our QRS model presented in Eq. 1.

Our pre-processing for mECG suppression is equal to the one used by Varanini. After mECG suppression and ICA, Varanini performs fetal R-peak detection on all individual channels and the channel with the best fetal RR-series is selected. Varanini uses a two step fetal R-peak detection. In the first step, potential fetal R-peaks are detected using a derivative filter. In the second step a forward and backward TVAR model is trained on the RR-series from the first step, starting from an RR-interval that is closest to the mode RR-interval. The forward and backward TVAR model are then used in combination with the derivative signal to detect the fetal R-peaks. After R-peak detection, the channel with the best RR-series is selected based on some statistical features of the RR-series.

For the algorithm of Biglari we used the same mECG suppression as the one proposed by Varanini. After mECG suppression, Biglari uses predefined templates for fetal R-peak detection that are based on most common fetal QRS morphologies. Fetal R-peaks are detected in each channel separately. First, each channel is passed through the matched filters using normalized cross-correlation. From the cross-correlated signals, the vectorcardiogram amplitude (VA) [12] is calculated per channel, which is then used to detect the fetal R-peaks with a local search algorithm over a sliding window

(allowing for HR detection up to 210 bpm) [12]. Finally, channels were ranked according to the robust weighted average [34] and the R-peaks from the channel with the highest rank were stored.

The algorithm of Behar uses a combination of template subtraction techniques and BSS techniques to suppress the mECG and enhance the fECG after mECG suppression [18]. Then, fetal R-peak detection is performed on all generated signals using an adapted version of the Pan and Tomkins algorithm [35] and a quality index is used to select the best RR-series. Finally, a smoothing step is performed to remove extra detected fetal R-peaks and fix missed fetal R-peaks. Note that we unchecked the *cinc-match* option, since this reduced the performance of the algorithm of Behar.

III. Data acquisition and evaluation

We used simulations to evaluate the flexibility of our algorithm for changing conditions of the fECG morphology due to fetal movement. We assumed that the mECG was already suppressed and used the model presented in [12] to simulate the fECG. This model assumes that the electrical activity of the heart (at far-field) can be approximated by a single dipole field vector. The three-dimensional path that is described by this vector over time is called the vectorcardiogram (VCG). The ECG is a projection of the field vector onto electrode leads. We simulated a full rotation of the fetus within the abdomen by changing the position and orientation of the VCG with respect to a vertical and horizontal electrode lead (as shown in Fig. 5(a)). Similar to [12], non-stationary noise was generated based on real noise signals (muscle artifacts and electrode movements) that can be obtained in the MIT-BIH non-stress test database [36]. Note that we did not consider baseline wander, because this would be removed by high-pass frequency filtering in mECG suppression [17].

To evaluate our method we used the set-A of the 2013 Physionet/Computing in Cardiology Challenge [15]. The database consists of 75 abdominal ECG recordings, measured with four channels at a sampling rate of 1000 Hz. For set-A, the reference fetal R-peaks are provided. Recordings a33, a38, a52, a54, a71, and a74 were excluded due to partially missing annotations [17], [18], leaving 69 recordings for evaluation.

The performance of the R-peak detection is evaluated by comparing the detected fetal R-peaks to the annotated peaks. As evaluation metrics we used the accuracy (Ac):

$$\text{Ac} = \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP}}. \tag{35}$$

Here, TP is the number of correctly detected fetal R-peaks, FN is the number of missed R-peaks, and FP is the number of falsely detected R-peaks. A detected R-peak is considered a TP if it falls within a window of 50 ms from an annotated peak [16], [18].

To determine the performance for different SNR conditions, we calculated an average SNR for each recording and split the data-set into three groups: recordings with low SNR (lowest 10%), median SNR (middle group), and high SNR (highest 10%). We calculated the SNR in the signals after mECG suppression and ICA, because these are the signals that were used for fetal R-peak detection. For each annotated R-peak in each IC, the SNR was calculated as:

$$S = 10 \log_{10} \frac{P_s}{P_n}, \tag{36}$$

with P_s the power of the QRS complex and P_n the power of the noise. P_s was calculated as the variance in a window of ± 20 ms surrounding the annotated R-peak and P_n as the variance in a window of [-150:-20 20:150] ms surrounding the annotated R-peak. The average SNR was calculated for each IC and the SNR of a recording was defined as the maximum average SNR.

Note that we did not consider evaluation metrics for HR estimation (e.g. mean-squared-error (MSE) between detected and annotated HR). Post-processing techniques such as those proposed in [16] or [18] could be used to improve the fetal HR estimation, but this is outside the scope of this study. For the other sets of the Challenge (B and C), no evaluation of peak detection is possible, only on HR estimation.

TABLE II
Performance (Ac) for varying c_a and R_0.

		R_0 (ms^2)				
		1	4	25	100	400
c_a (a.u.)	10^{-1}	39.8	64.9	82.8	92.9	96.4
	10^{-3}	49.1	82.5	88.1	94.7	97.6
	10^{-5}	94.8	98.8	99.2	**99.6**	99.1
	10^{-7}	96.0	98.7	99.3	**99.6**	99.4
	10^{-9}	96.1	98.0	99.1	99.4	99.2

IV. Results

The process noise of the QRS model ($\boldsymbol{\Sigma}$, determined by c_a) and the observation noise of the HR model (R, determined by R_0) were empirically determined based on the optimization of the detection performance for set-A of the Challenge. Table II are shown for varying the value of c_a between 10^{-1} to 10^{-9} and R_0 between 1 to 400 ms^2. Based on these results, we set $c_a = 10^{-5}$ and $R_0 = 100$ ms.

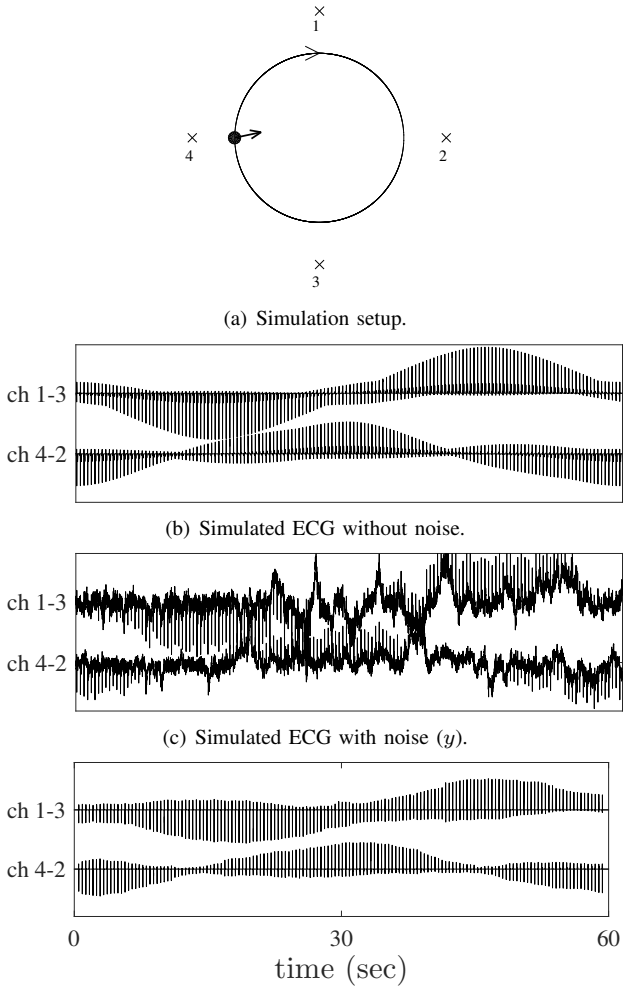
The simulated fECG signals of a verticale and horizontal lead for a full rotation of the fetus are shown in Fig. 5. The fECG with and without noise are shown in Fig. 5(b) and 5(c), respectively. The QRS complexes that were estimated by our algorithm ($\hat{y}$) are shown for both leads in Fig. 5(d).

In Fig. 6, examples for the log-posterior $\mathcal{L}$ are shown of a signal with good and a signal with poor quality. Note that the ICs after mECG suppression and ICA are presented. In Fig. 7, an example is shown with fECG visible in multip ICs and with SNR that varies over time. The detected fetal R-peaks are indicated by the vertical lines.

The detection performance for the set-A of the Physionet/Computing in Cardiology 2013 Challenge is shown in Table IV. Mean Ac is presented for different SNR regions. The mean Ac of the total dataset was 99.6% for our algorithm, and 98.6%, 92.9%, and 71.6% for the algorithms of Varanini, Behar, and Biglari, respectively.

V. Discussion

Fetal R-peak detection in abdominal fECG recordings is challenging due to the low SNR and the non-stationary nature



(a) Simulation setup.

(b) Simulated ECG without noise.

(c) Simulated ECG with noise (y).

(d) Estimated QRS complexes $\hat{y}$.

Fig. 5. Simulation of a full fetal rotation. In (a), 'x' indicates electrodes and the black dot indicates the location of the fetal heart at the start of the simulation. The vector that originates from the fetal heart represents the direction of the main electrical activity at that position (third dimension is not displayed). The circle shows the path over which the fetal heart rotates.

TABLE III
AVERAGE AC FOR SET-A.

algorithm	All SNR (%)	Low SNR (%)	Median SNR (%)	High SNR (%)
this work	**99.6**	**99.4**	**99.6**	**99.9**
Varanini et al. [17]	98.6	93.5	99.1	99.8
Behar et al. [18]	92.9	59.7	96.3	99.8
Biglari et al. [26]*	71.6	62.3	71.2	84.3

* Biglari reported higher performance for similar, yet different, datasets.

of the fECG waveform in the abdominal recordings. Although important processing steps have been developed for suppression of the mECG, less attention has been paid to fetal R-peak detection. In this paper we present a hierarchical probabilistic framework for fetal R-peak detection that combines predictive models of the ECG waveform and fetal HR.

A. QRS and HR model

For the QRS model we made the assumption that the fetal orientation remains constant throughout the period of a QRS complex. In reality, fetal movement also occurs during the period of a QRS complex, but since the period of the QRS complex is relatively short (approximately 40 ms), we expect the effect on the QRS waveform to be limited. This assumption allows us to estimate the model parameters $\boldsymbol{z}$ over L samples instead of only one sample, making the estimation more robust against noise.

Although our QRS model requires a relatively small number of parameters, we expect that the model is able to describe the QRS complex for most normal abdominal fECG recordings [26]. The simulation in Fig. 5 confirm that the QRS model is flexible and that changes in fECG morphology due to fetal movement can be described by our model. However, if our model is unable to describe the QRS complex (e.g. in case of a congenital defect) the performance might be reduced due to the sanity check in Eq. 27. Moreover, since the model is limited to describing the QRS complex, it cannot be used for ECG waveform analysis.

For the HR model we used a linear autoregressive model. A more complete description of the HR could be obtained by a point-process model [37]. Besides, a linear model might be insufficient to fully describe the HR, as several studies in the literature have identified non-linear dynamics in the HR [33]. However, as noted in [37], a stochastic parameter estimation of a linear model can still lead to an accurate description of the HR under normal conditions. During labor it could be interesting to extend the HR model to a non-linear model (e.g. such as the one proposed in [38]) to describe complicated accelerations or decelerations.

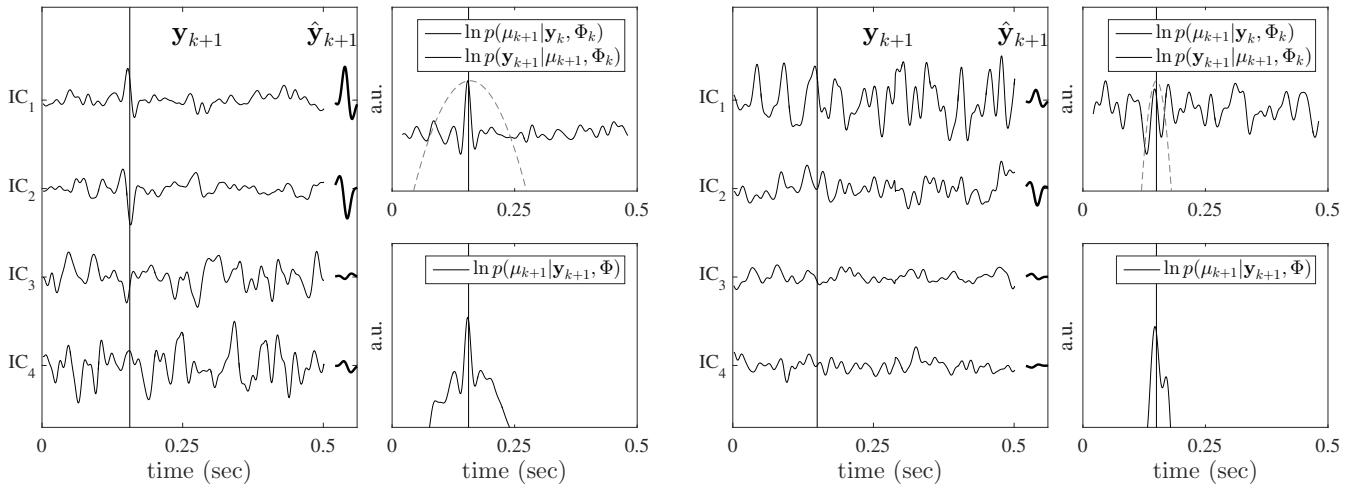
In the presence of cardiac arrhythmias we expect the performance of our algorithm to decrease. This can either be due to abnormal variations in the ECG waveform, abnormal variations in the HR or both. Since our model is developed to detect normal sinus rhythms, for future work it could be interesting to extend our model for arrythmia detection and classification at periods where R-peak detection fails.

B. Noise models

To allow for an analytically tractable solution, we have made assumptions that the observation and process noise in the QRS and HR model have a zero-mean Gaussian distribution.

The process noise covariance of the QRS model ($\boldsymbol{\Sigma}$) is assumed constant, meaning that the QRS waveform is allowed to vary similarly throughout a recording. In reality, variations in the QRS waveform are mainly caused by fetal movements or arrhythmias, which do not occur continuously over time. Hence, adaptive estimation of $\boldsymbol{\Sigma}$ seems more appropriate. For adaptive estimation of $\boldsymbol{\Sigma}$, the residual error between the estimated QRS waveform and the observed signal could be used (similar to the method used for estimating the process noise of the HR model). However, because the estimated QRS waveform is also used to identify mis-detections through Eq. 27, this approach reduces the performance. Despite having a fixed process noise covariance, the QRS model is still capable to adapt in case of fetal movement, as shown in Fig. 5.

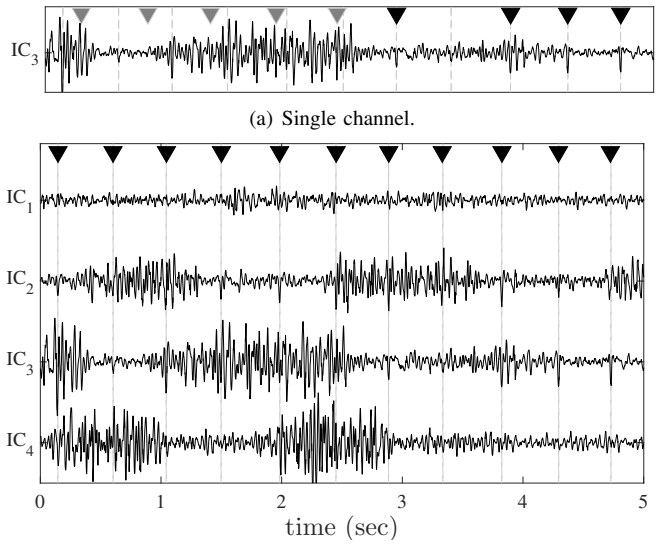
For the estimation of the observation noise covariance of the QRS model ($\boldsymbol{\Lambda}$), we assumed the observation noise to



(a) Good signal quality.

(b) Poor signal quality.

Fig. 6. Example of good (a) and poor (b) quality signal. The verticale line show the location of the annotated R-peak. In case of good quality, the QRS model dominates the posterior, while in case of poor quality the HR model dominates the posterior.



(a) Single channel.

(b) Multiple channels.

Fig. 7. Example of a signal with varying SNR in the ICs. The fECG is visible in IC$_2$ and IC$_3$. The annotated R-peaks are indicated by the verticale lines, and the detected peaks by the downward triangles. Correctly detected R-peaks are indicated by the black triangles and mis-detections by gray triangles. Detected peaks using a single IC are shown in (a) and detected peaks using all ICs are shown in (b).

be uncorrelated. However, also correlated physiological noise (e.g. muscle activity) is described by the observation noise of the QRS model. Although a whitening filter could be implemented as a pre-processing step, it is not straightforward to develop such a whitening filter for ECG recordings since the correlation in the signal varies over time. In [39], it was shown that even in case the observation noise is not white the Kalman filter provides the optimal mean squared error solution.

For estimation of the process noise of the HR model ($\boldsymbol{Q}$), the observation noise R plays an important role. If the value of R is overestimated, the process noise will be underestimated and the HR model becomes less capable at adapting its model parameters to dynamical HR variations that are caused by changes in autonomic regulation. On the other hand, underestimating R leads to overestimation of the process noise and overfitting of the parameter estimation. For future work it might be interesting to estimate R dynamically.

C. R-peak detection

New R-peaks are detected as the MAP estimate for the state-space model in Eq. 6. The log-posterior is the sum of the estimation of the QRS and HR model, weighted by the uncertainties in the respective models. If the quality of the recorded signal is high, the QRS model dominates the posterior distribution, as is shown in Fig. 6(a). Conversely, if the signal quality is poor, the HR model dominates the posterior distribution, as is shown in Fig. 6(b).

An MAP solution for R-peak detection was also proposed in [40]. However, the model that is used in [40] is designed for fiber-optic signals, from which it is not possible to detect the actual R-peaks locations due to differences in timing between audible and electrical activity of the heart. In [40], different detection methods are combined in an MAP estimate of the (ECG) R-peak location, that accounts for the delays and uncertainties in each of the methods with respect to true R-peak locations. This multimethod approach is not required in our case, because we measure the ECG and can detect actual R-peak locations.

In Fig. 7 an example is shown where ICA is unable to separate the fECG information in a single IC. In this example the SNR of the ICs varies over time, which reduces performance for algorithms that detect R-peaks on individual channels, as demonstrated in Fig. 7(a). In the multichannel extension of our model all channels are used for R-peak detection, which improves the performance in case of varying SNR conditions (Fig. 7(b)).

The performance of our algorithm has been evaluated for set-A of the Challenge. We compared performance of our algorithm to the performance of the algorithms of Varanini et

al. [17], Behar et al. [18], and Biglari et al. [26]. Varanini and Behar obtained the best scores for the Challenge. It should be noted that for set-B of the Challenge Andreotti et al. [16] and Lipponen et al. [19] achieved slightly better scores compared to Varanini and Behar in terms of fetal HR estimation. Since these algorithms were not publicly available we did not include them in our comparison.

Highest overall Ac was achieved by our algorithm (99.6%). The lowest performance was obtained by the algorithm of Biglari (71.6% overall Ac). Reasons for the relatively low performance of the algorithm of Biglari could be that Biglari uses fixed templates for R-peak detection and that R-peaks are detected on the individual channels. The algorithms of both Varanini and Behar achieved good overall Ac (98.6% and 92.9%, respectively). However, from Tabel IV it becomes clear that the performance of these algorithms is reduced for low SNR conditions (93.5% for Varanini and 59.7% for Behar). In contrast, our algorithm achieved 99.4% for the recordings with low SNR, indicating that our algorithm works well even for low SNR conditions.

VI. Conclusion

In this study, a hierarchical probabilistic framework was developed for fetal R-peak detection. The developed method combines predictive models of the ECG waveform and heart rate, and can be used for multichannel recordings. The developed method outperforms other methods that have been proposed in the literature in terms of detection accuracy.

VII. Acknowledgements

This research was performed within the framework of the IMPULS perinatology program.

References

[1] R.G. Kennedy. Electronic fetal heart rate monitoring: retrospective reflections on a twentieth-century technology. *J. R. Soc. Med.*, 91(5):244, 1998.

[2] Z. Alfirevic et al. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst. Rev.*, (3):CD006066, 2006.

[3] I. Amer-Wåhlin et al. Cardiotocography only versus cardiotocography plus st analysis of fetal electrocardiogram for intrapartum fetal monitoring: a swedish randomised controlled trial. *The Lancet*, 358(9281):534–538, 2001.

[4] J. Jezewski et al. Comparison of doppler ultrasound and direct electrocardiography acquisition techniques for quantification of fetal heart rate variability. *IEEE Transactions on Biomedical Engineering*, 53(5):855–864, 2006.

[5] A.K. Sundström et al. Fetal surveillance. *Gothenburg: Neoventa Medical AB*, 2000.

[6] S. Cerutti et al. Variability analysis of fetal heart rate signals as obtained from abdominal electrocardiographic recordings. *J. Perinat. Med.*, 14(6):445–452, 1986.

[7] R. Sameni and G.D. Clifford. A review of fetal ecg signal processing; issues and promising directions. *Open Pacing Electrophysiol Ther J.*, 3:4, 2010.

[8] T. F. Oostendorp et al. The effect of changes in the conductive medium on the fetal ecg throughout gestation. *Clin Phys Physiol Meas*, 10 Suppl B:11–20, 1989.

[9] M. Ungureanu et al. Fetal ecg extraction during labor using an adaptive maternal beat subtraction technique. *Biomed. Tech.*, 52(1):56–60, 2007.

[10] R. Vullings et al. Dynamic segmentation and linear prediction for maternal ECG removal in antenatal abdominal recordings. *Physiol. Meas.*, 30(3):291–307, Mar 2009.

[11] B. Widrow et al. Adaptive noise cancelling: Principles and applications. *Proceedings of the IEEE*, 63(12):1692–1716, 1975.

[12] R. Sameni. *Extraction of fetal cardiac signals from an array of maternal abdominal recordings*. PhD thesis, Citeseer, 2008.

[13] P.P. Kanjilal et al. Fetal ecg extraction from single-channel maternal ecg using singular value decomposition. *IEEE Trans. Biomed. Eng.*, 44(1):51–59, 1997.

[14] L. De Lathauwer et al. Fetal electrocardiogram extraction by blind source subspace separation. *IEEE Trans. Biomed. Eng.*, 47(5):567–572, 2000.

[15] I. Silva et al. Noninvasive fetal ecg: the physionet/computing in cardiology challenge 2013. In *Computing in Cardiology Conference (CinC), 2013*, pages 149–152. IEEE, 2013.

[16] F. Andreotti et al. Robust fetal ecg extraction and detection from abdominal leads. *Physiol. Meas.*, 35(8):1551, 2014.

[17] M. Varanini et al. An efficient unsupervised fetal qrs complex detection from abdominal maternal ecg. *Physiol. Meas.*, 35(8):1607, 2014.

[18] J. Behar et al. Combining and benchmarking methods of foetal ecg extraction without maternal or scalp electrode data. *Physiol. Meas.*, 35(8):1569, 2014.

[19] J.A. Lipponen and M.P. Tarvainen. Principal component model for maternal ecg extraction in fetal qrs detection. *Physiol. Meas.*, 35(8):1637, 2014.

[20] J. Behar et al. A practical guide to non-invasive foetal electrocardiogram extraction and analysis. *Physiol. Meas.*, 37(5):R1, 2016.

[21] M.J. Rooijakkers et al. Low-complexity R-peak detection for ambulatory fetal monitoring. *Physiol. Meas.*, 33(7):1135–1150, Jul 2012.

[22] M.J. Rooijakkers et al. Influence of electrode placement on signal quality for ambulatory pregnancy monitoring. *Comput. Math. Methods. Med.*, 2014, 2014.

[23] R. Vullings et al. Novel bayesian vectorcardiographic loop alignment for improved monitoring of ecg and fetal movement. *IEEE Trans. Biomed. Eng.*, 60(6):1580–1588, 2013.

[24] P.E. McSharry et al. A dynamical model for generating synthetic electrocardiogram signals. *IEEE Trans. Biomed. Eng.*, 50(3):289–294, 2003.

[25] R. Sameni et al. A nonlinear bayesian filtering framework for ecg denoising. *IEEE Trans. Biomed. Eng.*, 54(12):2172–2185, 2007.

[26] H. Biglari and R. Sameni. Fetal motion estimation from noninvasive cardiac signal recordings. *Physiol. Meas.*, 37(11):2003, 2016.

[27] H. Akaike. Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics*, 21(1):243–247, 1969.

[28] Matthias Arnold, XHR Milner, Herbert Witte, Reinhard Bauer, and Christoph Braun. Adaptive ar modeling of nonstationary time series by means of kalman filtering. *IEEE Transactions on Biomedical Engineering*, 45(5):553–562, 1998.

[29] J.G. Nijhuis et al. Are there behavioural states in the human fetus? *Early Hum. Dev.*, 6(2):177–195, 1982.

[30] A. Gelb. *Applied optimal estimation*. MIT press, 1974.

[31] J.F.G. de Freitas et al. Hierarchical bayesian-kalman models for regularisation and ARD in sequential learning. Technical report, Cambridge University, 1998.

[32] J. Stinstra et al. Multicentre study of fetal cardiac time intervals using magnetocardiography. *BJOG*, 109:1235–1243, Nov. 2002.

[33] A. Voss et al. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos. Trans. A. Math. Phys. Eng. Sci.*, 367(1887):277–296, 2009.

[34] Jacek M Leski. Robust weighted averaging [of biomedical signals]. *IEEE Transactions on Biomedical Engineering*, 49(8):796–804, 2002.

[35] J. Pan and W. J. Tompkins. A real-time qrs detection algorithm. *IEEE Trans. Biomed. Eng.*, 32(3):230–236, Mar 1985.

[36] et al. Goldberger. Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23):e215–e220, 2000.

[37] R. Barbieri et al. A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability. *Am. J. Physiol. Heart. Circ. Physiol.*, 288(1):H424–H435, 2005.

[38] G. Valenza et al. Point-process nonlinear models with laguerre and volterra expansions: Instantaneous assessment of heartbeat dynamics. *IEEE Trans. Signal Process.*, 61(11):2914–2926, 2013.

[39] B.D.O. Anderson and J.B. Moore. Optimal filtering., 1979.

[40] S. Šprager and D. Zazula. Optimization of heartbeat detection in fiber-optic unobtrusive measurements by using maximum a posteriori probability estimation. *IEEE journal of biomedical and health informatics*, 18(4):1161–1168, 2014.