An Adaptive Kalman Filter for ECG Signal Enhancement

Rik Vullings*, Bert de Vries, and Jan W. M. Bergmans

***Abstract*—The ongoing trend of ECG monitoring techniques to become more ambulatory and less obtrusive generally comes at the expense of decreased signal quality. To enhance this quality, consecutive ECG complexes can be averaged triggered on the heartbeat, exploiting the quasi-periodicity of the ECG. However, this averaging constitutes a tradeoff between improvement of the SNR and loss of clinically relevant physiological signal dynamics. Using a Bayesian framework, in this paper, a sequential averaging filter is developed that, in essence, adaptively varies the number of complexes included in the averaging based on the characteristics of the ECG signal. The filter has the form of an adaptive Kalman filter. The adaptive estimation of the process and measurement noise covariances is performed by maximizing the Bayesian evidence function of the sequential ECG estimation and by exploiting the spatial correlation between several simultaneously recorded ECG signals, respectively. The noise covariance estimates thus obtained render the filter capable of ascribing more weight to newly arriving data when these data contain morphological variability, and of reducing this weight in cases of no morphological variability. The filter is evaluated by applying it to a variety of ECG signals. To gauge the relevance of the adaptive noise-covariance estimation, the performance of the filter is compared to that of a Kalman filter with fixed, (*a posteriori*) optimized noise covariance. This comparison demonstrates that, without using *a priori* knowledge on signal characteristics, the filter with adaptive noise estimation performs similar to the filter with optimized fixed noise covariance, favoring the adaptive filter in cases where no *a priori* information is available or where signal characteristics are expected to fluctuate.**

***Index Terms*—Electrocardiography, Kalman filter, noise estimation.**

I. Introduction

MONITORING and analysis of the ECG has long been used in clinical practice. In recent years, the application field of ECG monitoring is expanding to areas outside the clinic. An example of such an area is at-home monitoring of patients with sleep apnea [1]. Also within the clinic, a transition in ECG monitoring applications is taking place. With developments in sensor technology (e.g., textile electrodes and capacitive electrodes), sensors that are incorporated in garments or the matrass of an incubator [2] have become available.

As a result of these new sensor technologies, the comfort of the patient is improving progressively. Whereas some years ago the patient had to reconcile himself or herself with the discomforts of the only available technology, nowadays patients prefer the more comfortable ways of recording the ECG. However, in most cases, this increased comfort comes at the expense of signal quality. Electrodes that are incorporated in garments generally provide signals with a lower SNR and more artifacts than contact electrodes that are glued to the body [3]. Another example of ECG signals with a typically low SNR is fetal ECG signals, either recorded invasively after membrane rupture [4] or noninvasively from the maternal abdomen [5].

Some of the SNR and artifact problems that arise during these recordings can be suppressed by simple, frequency-selective filtering [5]–[7]. However, due to the partial overlap of signal and noise bandwidths [8], [9], this frequency-selective filtering can only help to some extent. Further improvement of the ECG can be achieved by exploiting its (quasi-)periodicity. Consecutive ECG complexes resemble one another and are, moreover, in general uncorrelated with noise and artifacts. Hence, by averaging several consecutive ECG complexes, the SNR can be improved. For additive Gaussian noise, this improvement is directly related to the number of ECG complexes included in the average [10]. The drawback of averaging multiple consecutive ECG complexes is that, besides noise, also the physiological dynamics of the ECG are suppressed. That is, changes in the ECG that originate from physiological events—for instance, changes in the ST segment that might be associated to metabolic acidosis [11]—are suppressed in the averaging, complicating clinical diagnosis.

From this, it is clear that the averaging of ECG complexes entails a tradeoff between the pursued increase in SNR and the time scale at which physiologically relevant changes in ECG morphology are expected to occur. Hence, for each specific application, the number of complexes n included in the averaging needs to be reconsidered. If it were possible, however, to dynamically adapt the number of complexes in the average, based on newly arriving data, the problem of selecting a proper value for n could potentially be overcome. In this paper, we develop a filter that can do exactly this.

The filter is derived using a Bayesian framework and constitutes a Kalman filter in which the dynamic variations in the ECG are modeled by a covariance matrix that is adaptively estimated every time new data arrive. In contrast to filters that filter the ECG by modeling it by parametric functions [12], the presented filter uses the actual recorded ECG as basis and infers

Manuscript received August 26, 2010; revised November 2, 2010; accepted December 1, 2010. Date of publication December 13, 2010; date of current version March 18, 2011. This work was supported by the Dutch Technology Foundation STW. *Asterisk indicates corresponding author*.

*R. Vullings is with the Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, 5600 MB, The Netherlands (e-mail: r.vullings@tue.nl).

B. de Vries and J. W. M. Bergmans are with the Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, 5600 MB, The Netherlands (e-mail: b.de.vries@tue.n; j.w.m.bergmans@tue.nl).

Digital Object Identifier 10.1109/TBME.2010.2099229

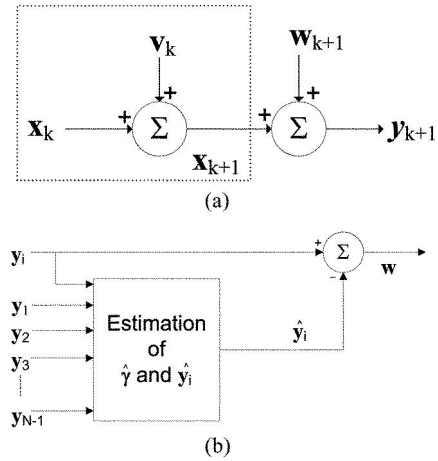

Fig. 1. (a) Illustration of the state-space model that describes the evolution of the ECG over time. The evolution of the state vectors is indicated by the dotted box. (b) Illustration of the measurement noise estimation.

whether this ECG is corrupted by noise or dynamic variations. As a result, unanticipated physiological anomalies in the ECG, which cannot be easily captured by simple parametric functions, can be accurately modeled. For parametric functions, to capture such physiological anomalies, large families of analytical functions or many function parameters need to be considered, both inherently slowing down the filter process.

The derivation of this filter is provided in Section II. The ECG dataset, on which the filter is evaluated, is discussed in Section III, and the results of this evaluation are provided in Section IV. Finally, discussion and conclusions are provided in Sections V and VI, respectively.

II. Derivation of Adaptive Kalman Filter

A. Bayesian Model

Typically, ECG complexes that originate from consecutive heartbeats are very similar but not identical. Moreover, when recording the ECG, the signals are corrupted to some extent by noise and artifacts. In a simplified form, both the relation between consecutive ECG complexes and the corruption of the recorded ECG can be described by means of a state-space model [see also Fig. 1(a)] as follows:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{v}_k \\ \mathbf{y}_{k+1} = \mathbf{x}_{k+1} + \mathbf{w}_{k+1} \end{cases} \tag{1}$$

where $\mathbf{x}_k$ denotes the $[T\times 1]$ ECG complex for heartbeat k and $\mathbf{y}_k$ denotes the $[T\times 1]$ recorded signal where T is the length of the ECG complex. The isolation of individual ECG complexes from the recorded signals is discussed in Section III-C. Also in this section, the choice for T and the implicit assumption of equal lengths for all ECG complexes is discussed. The evolution of the ECG complexes between heartbeats is modeled by the $[T\times 1]$ stochastic component $\mathbf{v}_k$ (also referred to as the process noise). The measurement noise, i.e., corrupting signals, such as electromyographic signals, movement artifacts, and interferences from the powerline grid, is represented by the $[T\times 1]$ vector $\mathbf{w}_k$.

When critically assessing (1) and Fig. 1, it is clear that based on the state-space model alone, no clear distinction between the process noise $\mathbf{v}_k$ and the measurement noise $\mathbf{w}_k$ can be made. Therefore, a separate model [illustrated in Fig. 1(b)] is used for estimating the measurement noise. In this model, the spatial correlation between ECG signals recorded simultaneously at different locations is exploited. This spatial correlation renders it possible to approximate a particular ECG signal by the combination of the other, simultaneously recorded ECG signals. The part of the ECG signal that cannot be approximated by the combination of the other signals is subsequently assumed to be measurement noise. The estimation of the measurement noise will be discussed in more detail in Section II-B. With regard to the process noise, $\mathbf{v}_k$ is assumed to be zero mean with adaptive covariance $\mathbf{\Lambda}_k$. Similarly, the measurement noise $\mathbf{w}_k$ is assumed to be zero mean with covariance $\mathbf{\Sigma}_k$. The assumption of zero mean for both $\mathbf{v}_k$ and $\mathbf{w}_k$ can be justified by high-pass filtering the ECG signals, as will be described in Section III-B.

In the state-space description of (1), the problem of enhancing the SNR of the ECG is reduced to the problem of sequentially estimating the model parameter vector $\mathbf{x}_k$ and the noise covariances $\mathbf{\Sigma}_k$ and $\mathbf{\Lambda}_k$. Here, sequential estimation refers to the estimation of the relevant parameters based on the earlier estimate and all newly arriving data.

B. Estimation of Measurement Noise

When recording several ECG signals simultaneously, these signals are spatially correlated to some extent. Specifically, the electrical activity of the heart can be modeled as a time-dependent dipole that is variable in both amplitude and (3-D) orientation. In this model, each ECG signal constitutes the projection of the electrical field generated by this dipole onto the vector that describes the position of the recording electrode. Hence, each ECG signal can be constructed from the linear combination of three independent ECG signals [13]. For M recorded ECG signals, this means that the ECG signal $\mathbf{x}_i$ can be modeled [see also Fig. 1(b)] as follows:

$$\mathbf{x}_i = \mathbf{X}_{-i}\boldsymbol{\gamma} \tag{2}$$

where $\mathbf{X}_{-i}$ is a $[T \times (M-1)]$ matrix, of which the M ECG signals $\mathbf{x}_j$ constitute the column vectors, and for which the ith column is missing. The $[(M-1)\times 1]$ vector $\boldsymbol{\gamma}$ comprises the coefficients of the linear combination. The index k that denotes the heartbeat in (1) is omitted from this description for clarity.

With the adopted dipole model of the heart's electrical activity, it can be argued that dynamical variations in the ECG morphology are reflected in all recorded ECG signals $\mathbf{y}$. Analogously, measurement noise $\mathbf{w}$ that does not exhibit the same spatial correlation as the ECG is suppressed in the linear combination of ECG signals. As a result, the measurement noise vector $\mathbf{w}_i$ for ECG signal i can be approximated by $\hat{\mathbf{w}}_i$ using the estimate $\hat{\mathbf{y}}_i = \mathbf{Y}_{-i}\boldsymbol{\gamma}$ as follows:

$$\hat{\mathbf{w}}_i = \mathbf{y}_i - \hat{\mathbf{y}}_i \tag{3}$$

also yielding an estimate for the measurement noise covariance $\mathbf{\Sigma}$.

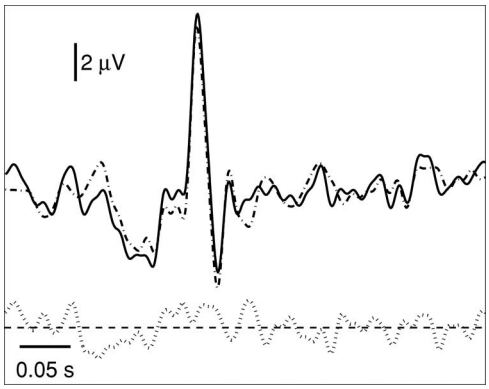

Fig. 2. Example of the estimation of the measurement noise in an ECG complex obtained from an eight-channel, noninvasive fetal ECG recording (see Section III). The solid line represents the recorded ECG complex and the "-." line represents the estimate of this ECG complex obtained by the linear combination of the seven simultaneously recorded signals. The differential signal, represented by the dotted line, constitutes an approximation of the measurement noise. Note that for clarity, the measurement noise signal is depicted with a vertical offset. The scalings of all signals are the same.

The estimates $\hat{\boldsymbol{\gamma}}$ that minimize the mean-squared error (MSE) between $\mathbf{y}_i$ and its estimate $\hat{\mathbf{y}}_i = \mathbf{Y}_{-i}\hat{\boldsymbol{\gamma}}$ can be determined by the following:

$$\hat{\boldsymbol{\gamma}} = \left(\mathbf{Y}_{-i}^T\mathbf{Y}_{-i}\right)^{-1}\mathbf{Y}_{-i}^T\mathbf{y}_i. \tag{4}$$

The matrix inversion in (4) exists in case the column vectors in $\mathbf{Y}_{-i}$ are linearly independent [14]. This condition of linear independence is satisfied in the case of ECG signals, for one due to the fact that each column vector is corrupted by independent, additive noise. The estimation of the measurement noise is illustrated in Fig. 2.

The main limitation of this method for estimating the measurement noise is that, at any time, at least four ECG signals have to be recorded: three independent ones to estimate the fourth. For most cases exemplified in Section I, however, the recording of multiple ECG signals is the standard procedure, and hence, the requirement for more than three signals does not impose a serious restriction to the applicability of the proposed SNR enhancement method.

C. Kalman Filter for Parameter Estimation

The uncertainty in the state-space model of (1) and in the associated noise parameters suggests the use of a probabilistic approach for solving the parameter estimation problem [15]. In addition, the sequential nature of the estimation problem motivates the use of a Bayesian framework in which the prior probability distribution assigned to the unknown parameters is updated every time new data arrive. Here, again, sequential refers to the estimation of model parameters based on earlier parameter estimates and on newly arriving data.

Using Bayes' rule, the solution to the parameter estimation problem can be described as follows:

$$p\left(\mathbf{x}_{k+1}\,|\mathbf{y}_{k+1},\boldsymbol{\Lambda}_k,\boldsymbol{\Sigma}_k\right) = \frac{p\left(\mathbf{y}_{k+1}\,|\mathbf{x}_{k+1},\boldsymbol{\Lambda}_k,\boldsymbol{\Sigma}_k\right)p\left(\mathbf{x}_{k+1}\,|\mathbf{y}_k,\boldsymbol{\Lambda}_k,\boldsymbol{\Sigma}_k\right)}{p\left(\mathbf{y}_{k+1}\,|\mathbf{y}_k,\boldsymbol{\Lambda}_k,\boldsymbol{\Sigma}_k\right)}. \tag{5}$$

The conditional probability density function $p(\mathbf{x}_{k+1}|\mathbf{y}_{k+1})$ is referred to as the *posterior*. Since it contains all statistical information about $\mathbf{x}_{k+1}$, this posterior constitutes the complete solution to the parameter estimation problem [16]. The probability density functions on the right-hand side of (5) are referred to as the *likelihood* and the *prior*, respectively, for the numerator and as the *evidence* for the denominator.

By assuming the prior and likelihood to be Gaussian, the posterior and evidence are necessarily Gaussian as well. The use of Gaussian approximations is dictated by the fact that they render the posterior describable by a limited number of parameters and, as such, enable the estimation of the ECG in a maximum *a posteriori* (MAP) fashion [15]. For applications in which the posterior is expected to be multimodal (i.e., a function with several peaks), a combination of Gaussians can be used, each describing a different mode of the posterior. The fact that, here, the posterior is assumed as a single Gaussian, implies that the parameter vector estimate $\hat{\mathbf{x}}_{k+1}$ and its associated covariance $\boldsymbol{\Psi}_{k+1}$ together completely describe the posterior probability density function and can be inferred analytically. Hence, using (5) and the assumptions in the state-space model, the posterior is given by [15] the following:

$$\mathcal{N}\left(\mathbf{x}_{k+1}\,|\hat{\mathbf{x}}_{k+1},\boldsymbol{\Psi}_{k+1}\right) = \frac{\mathcal{N}\left(\mathbf{y}_{k+1}\,|\mathbf{x}_{k+1},\boldsymbol{\Sigma}_{k+1}\right)\mathcal{N}\left(\mathbf{x}_{k+1}\,|\hat{\mathbf{x}}_k,\boldsymbol{\Psi}_k+\boldsymbol{\Lambda}_k\right)}{\mathcal{N}\left(\mathbf{y}_{k+1}\,|\hat{\mathbf{x}}_k,\boldsymbol{\Psi}_k+\boldsymbol{\Lambda}_k+\boldsymbol{\Sigma}_{k+1}\right)} \tag{6}$$

where $\mathcal{N}(x|y,z)$ denotes a Gaussian probability distribution for x with mean y and covariance z.

By rewriting (6), the optimal Bayes estimate $\hat{\mathbf{x}}_{k+1}$ and its variance $\boldsymbol{\Psi}_{k+1}$ can be sequentially updated according to

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_k + \mathbf{K}_{k+1}\left(\mathbf{y}_{k+1} - \hat{\mathbf{x}}_k\right) \tag{7}$$

$$\boldsymbol{\Psi}_{k+1} = \boldsymbol{\Psi}_k + \boldsymbol{\Lambda}_k - \mathbf{K}_{k+1}\left(\boldsymbol{\Psi}_k + \boldsymbol{\Lambda}_k\right) \tag{8}$$

where $\mathbf{K}_{k+1}$ is known as the Kalman gain [17]

$$\mathbf{K}_{k+1} = \frac{\boldsymbol{\Psi}_k + \boldsymbol{\Lambda}_k}{\boldsymbol{\Sigma}_{k+1} + \boldsymbol{\Psi}_k + \boldsymbol{\Lambda}_k}. \tag{9}$$

Together, (7)–(9) constitute the Kalman filter equations.

D. Adaptive Process Noise Covariance Estimation

A limitation of the derived Kalman filter is its implicit assumption of known *a priori* statistics for the measurement noise $\mathbf{w}_k$ and process noise $\mathbf{v}_k$. Moreover, in the ECG monitoring applications for which the filter is intended, the noise statistics are expected to be nonstationary and, hence, any choice for particular noise covariances potentially leads to large estimation errors [18]. These estimation errors can nonetheless be restricted by including a sequential estimation of the noise statistics in the Kalman filter equations.

The estimation of the measurement noise statistics has been discussed in Section II-B. The discussion in this section is hence limited to the estimation of the process noise covariance $\boldsymbol{\Lambda}_k$.

Again using Bayes' rule, the conditional probability density function for $\mathbf{\Lambda}_k$, given the recorded signal $\mathbf{y}_{k+1}$ is given by

$$p\left(\mathbf{\Lambda}_k \left|\mathbf{y}_{k+1}, \mathbf{\Sigma}_k\right.\right) = \frac{p\left(\mathbf{y}_{k+1} \left|\mathbf{y}_k, \mathbf{\Lambda}_k, \mathbf{\Sigma}_k\right.\right)}{p\left(\mathbf{y}_{k+1} \left|\mathbf{y}_k\right.\right)} p\left(\mathbf{\Lambda}_k \left|\mathbf{y}_k, \mathbf{\Sigma}_k\right.\right). \tag{10}$$

It can be noted here that the likelihood of the noise covariance $p\left(\mathbf{y}_{k+1} \left|\mathbf{y}_k, \mathbf{\Lambda}_k, \mathbf{\Sigma}_k\right.\right)$ is identical to the evidence function in the parameter estimation level of (5). Hence, maximizing the evidence function in this parameter estimation level is analogous to maximizing the likelihood of $\mathbf{\Lambda}_k$ for newly arriving data. Maximization of the evidence function, however, yields that the estimated noise covariance constitutes the maximum likelihood (ML) estimate instead of the MAP estimate, implying the assumption of no knowledge of the prior at the noise estimation level [15].

When defining the model residual to be

$$\begin{aligned}\boldsymbol{\rho}_{k+1} &\triangleq \mathbf{y}_{k+1} - \mathbf{E}\left[\mathbf{y}_{k+1} \left|\mathbf{y}_k, \mathbf{\Lambda}_k, \mathbf{\Sigma}_k\right.\right] \\ &= \mathbf{y}_{k+1} - \hat{\mathbf{x}}_k \end{aligned} \tag{11}$$

it can easily be calculated that $\mathbf{E}[\boldsymbol{\rho}_{k+1}|\mathbf{y}_k] = \mathbf{0}$ and $\mathbf{E}[\boldsymbol{\rho}_{k+1}\boldsymbol{\rho}_{k+1}^T|\mathbf{y}_k] = \mathbf{\Psi}_k + \mathbf{\Lambda}_k + \mathbf{\Sigma}_{k+1}$. Since in addition $\mathbf{E}[\boldsymbol{\rho}_k^T \boldsymbol{\rho}_l|\mathbf{y}_k] = 0$, it follows that

$$p\left(\boldsymbol{\rho}_{k+1}\right) = \frac{\exp\left[-\frac{1}{2}\boldsymbol{\rho}_{k+1}^T \left(\mathbf{\Psi}_k + \mathbf{\Lambda}_k + \mathbf{\Sigma}_{k+1}\right)^{-1} \boldsymbol{\rho}_{k+1}\right]}{\left(2\pi\right)^{T/2} \left|\mathbf{\Psi}_k + \mathbf{\Lambda}_k + \mathbf{\Sigma}_{k+1}\right|^{1/2}}$$

is equivalent to the evidence function at the parameter estimation level given in (6). Hence, by maximizing $p(\boldsymbol{\rho}_{k+1})$ with respect to the process noise covariance $\mathbf{\Lambda}_k$, the ML estimates for this covariance can be obtained.

The maximization of $p(\boldsymbol{\rho}_{k+1})$ can be simplified, if we return to the intended purpose of the Kalman filter, to adaptively vary the number of averages n used in the enhancement of the ECG complexes, depending on the dynamic variations in signal morphology. From (7), it can be inferred that this purpose means that the Kalman gain $\mathbf{K}_k$ can be simplified to a scalar matrix (i.e., a diagonal matrix with all entries equal), or even a scalar. Specifically, by varying the scalar value of $\mathbf{K}_k$ in (7), either more or less weight can be ascribed to the newly arriving ECG complex $\mathbf{y}_{k+1}$. In other words, the relative contribution of preceding ECG complexes to the estimate $\hat{\mathbf{x}}_{k+1}$ varies with the value of $\mathbf{K}_k$, essentially similar to adaptation of the number of averages used. The scalar value for $\mathbf{K}_k$ here ensures that all T samples in $\mathbf{y}_{k+1}$ and all T samples in $\hat{\mathbf{x}}_k$ are assigned the same weight ($\mathbf{K}_k$ for $\mathbf{y}_{k+1}$ and $(1 - \mathbf{K}_k)$ for $\hat{\mathbf{x}}_k$), preventing distortion of the ECG complexes. With the assumption of the Kalman gain being a scalar matrix, from (9), it then follows that also $\mathbf{\Psi}_k$, $\mathbf{\Lambda}_k$, and $\mathbf{\Sigma}_k$ can be assumed scalar matrices (i.e., $\psi_k^2\mathbf{I}$, $\lambda_k^2\mathbf{I}$, and $\sigma_k^2\mathbf{I}$, respectively, with $\mathbf{I}$ the $[T\times T]$ identity matrix $\mathbf{I}$), implicitly also assuming that both the measurement and process noise are spatially uncorrelated. The effect of the latter assumptions will be discussed in Section V. With the simplification of scalar matrices, not only can each of the scalar covariance matrices be regarded as the matrix representation of the variances of the vectors $\mathbf{x}_k$, $\mathbf{v}_k$, and $\mathbf{w}_k$, but also does the maximization of (the logarithm of) $p(\boldsymbol{\rho}_{k+1})$ reduce to the derivative of $\ln p(\boldsymbol{\rho}_{k+1})$ with respect to λ_k^2 equated to zero

$$\begin{aligned}\frac{\partial}{\partial\lambda_k^2} \ln p\left(\boldsymbol{\rho}_{k+1}\right) = \frac{1}{2}\mathrm{tr}\left[\boldsymbol{\rho}_{k+1}^T \left(\psi_k^2\mathbf{I} + \lambda_k^2\mathbf{I} + \sigma_{k+1}^2\mathbf{I}\right)^{-2} \boldsymbol{\rho}_{k+1}\right] \\ - \frac{1}{2}\mathrm{tr}\left[\left(\psi_k^2\mathbf{I} + \lambda_k^2\mathbf{I} + \sigma_{k+1}^2\mathbf{I}\right)^{-1}\right] = 0 \end{aligned} \tag{12}$$

where $\mathrm{tr}[\cdot]$ denotes the trace of the matrix. The use of $\ln p(\boldsymbol{\rho}_{k+1})$ instead of the use of $p(\boldsymbol{\rho}_{k+1})$ is justified by the monotonic behavior of the logarithm function.

Solving (12) for λ_k^2 yields an estimate for the process noise covariance as follows:

$$\hat{\lambda}_k^2 = \frac{1}{T}\boldsymbol{\rho}_{k+1}^T\boldsymbol{\rho}_{k+1} - \psi_k^2 - \sigma_{k+1}^2. \tag{13}$$

By computing the second derivative of $p(\boldsymbol{\rho}_{k+1})$, it is straightforward to prove that this result indeed corresponds to a global maximum in $p(\boldsymbol{\rho}_{k+1})$. In case the model errors $\frac{1}{T}\boldsymbol{\rho}_{k+1}^T\boldsymbol{\rho}_{k+1}$ are smaller than what the theoretical value of the measurement noise σ_{k+1}^2 predicts, no additional process noise input is required. This leads to the estimator as follows:

$$\lambda_k^2 = \begin{cases} \frac{1}{T}\boldsymbol{\rho}_{k+1}^T\boldsymbol{\rho}_{k+1} - \psi_k^2 - \sigma_{k+1}^2, & \text{if positive,} \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

The operation of the filter can be explained as follows. In case the model error $\frac{1}{T}\boldsymbol{\rho}_{k+1}^T\boldsymbol{\rho}_{k+1}$ is larger than what its theoretical value σ_{k+1}^2 predicts, λ_k^2 increases and this in turn leads to an increase in the Kalman gain. Hence, more emphasis is put on newly arriving data [15]. To improve the robustness and statistical significance of the estimator of (15), instead of a single residual $\boldsymbol{\rho}_{k+1}$, the sample mean of N residuals will be used [18]. The effect of the chosen value for N will be evaluated in Section IV.

Implementation of the aforedescribed methods in an algorithm constitutes the sequential execution of (4) and (3) to estimate the measurement noise covariance and, subsequently, (9), (7), (8), (11), and (14) for estimation of the ECG signals and process noise covariances. The algorithm is illustrated schematically in Fig. 3.

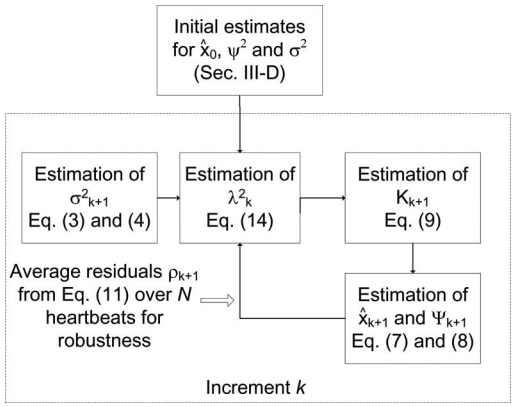

Fig. 3. Illustration of the algorithmic implementation of the developed adaptive Kalman filter.

III. Data Preparation and Initial Filter Settings

A. Data Acquisition

To evaluate the developed Kalman filter, a diversity of ECG signals is used. These signals comprise ECG signals of adult patients that suffer from *T*-wave alternans (TWA: a condition that renders the amplitude or shape of the ECG's *T*-wave to often vary significantly between heartbeats), fetal ECG signals recorded from the maternal abdomen, and neonatal ECG signals recorded with textile electrodes.

The three categories of ECG signals are mainly used to illustrate the performance of the filter. The first category (i.e., adult ECG signals with TWA), however, is also used for quantitative evaluation of the filter. In total, 12-lead ECG signals from 23 patients suffering from TWA and of 2 min length each are used in this evaluation. The signals are obtained from the Massachusetts Institute of Technology–Beth Israel Hospital TWA challenge database [19]. To map the filter's performance as a function of the SNR of the recorded signals as well, then to the performance's dependence on N, the ECG signals are corrupted with additive Gaussian noise of various amplitudes, yielding ECG signals with SNRs ranging from -3 to 24 dB.

The TWA signals comprise a rather ideal dataset for evaluating the performance of the developed filter. They exhibit relatively high SNR values that can be made smaller by additive Gaussian noise, and that moreover facilitate quantitative assessment of the filtered ECG signals (i.e., by comparing the filtered, with additive noise corrupted, ECG signals to the original ECG signals). In addition, the TWA signals exhibit morphological variability in the ECG that originates from underlying physiology.

B. Preprocessing

The acquired ECG data are preprocessed to remove noise, artifacts, and baseline wander that do not exhibit spectral overlap with the ECG. To this end, two frequency-selective fourth-order Butterworth filters [20] are used: one high-pass filter with cutoff frequency at 0.5 Hz and one low-pass filter with cutoff frequency at 90 Hz.

To suppress the interferences from the powerline grid, a notch filter centered around 50 Hz (or 60 Hz for recordings performance in U.S.) is used. Again, this filter is implemented as a fourth-order Butterworth filter. In contrast to the aforementioned high-pass and low-pass filters, this filter, however, also affects the ECG due to the fact that the frequency content of the ECG extends from frequencies of about 1 Hz to frequencies beyond 50 Hz. By choosing the width of the notch filter relatively small—but wide enough to account for fluctuations in the powerline frequency—the distortion of the ECG signals can be kept small.

For the transabdominally recorded fetal ECG recordings, it has to be noted that these are processed as described earlier, but with an additional processing step in which the interfering maternal ECG is suppressed. The details of this maternal ECG suppression are provided in [21].

C. Demarcation of Individual ECG Complexes

The consecutive individual ECG complexes that are used as input for the filter are defined based on their QRS complexes. Specifically, the QRS complexes are detected in the signals, and the ECG complexes are subsequently defined as the signal within a predefined time window around the QRS complex. After these ECG complexes have been filtered by the developed Kalman filter, the filtered counterparts of the original ECG signals are generated by placing the filtered ECG complexes back on their original positions. Because the Kalman filter is limited by the fact that the length of the input ECG complexes needs to be fixed (i.e., the length should be T s) and because the interval between consecutive ECG complexes varies over time, this ECG signal generation approach suffers from the drawback that in some cases there will be overlap between consecutive filtered ECG complexes, and in some cases there will be a gap. By choosing the time-window for the ECG complexes, such as to minimize the number of gaps, as much of the original ECG signals as possible are filtered. The gaps that inevitably remain are smoothed by interpolation of the data between successive complexes. The overlapping signal parts, in turn, are smoothed by gradually averaging the contributions of both overlapping complexes. Specifically, the contribution of the first ECG complex is gradually reduced and the contribution of the trailing complex is gradually increased. The value of T chosen here is 120% of the mean interval between consecutive heartbeats.

As mentioned earlier, before defining the individual ECG complexes, the QRS complexes need to be detected. To facilitate this detection, the SNR of the ECG signals is *a priori* enhanced by linearly combing the signals in such a way as to maximize the variance [principal component analysis (PCA)] [22]. The linear combination with maximum variance is referred to as the principal component. The QRS complexes are subsequently detected in the principal component as local extrema that exceed an adaptive threshold. This adaptive threshold is updated continuously by means of a simple Kalman filter and depends on the SNR of the ECG signals complexes in the principal component [21]; when the SNR changes, the threshold is adapted to prevent noise from exceeding it, in the mean time ensuring that the QRS complexes still exceed the threshold. For a more detailed description of the QRS complex detection, the reader is referred to [21] and [23].

D. Initializing the Filter

Before commencing the filtering of the ECG signals, all variables and parameters need to be initialized. For the initial estimate of the ECG complex $\hat{\mathbf{x}}_0$, the mean ECG complex over N heartbeats is used. Here, N is the same value as used for estimation of λ^2. The initial estimate for the measurement noise vector $\hat{\mathbf{w}}_1$ is determined according to its description in Section II-B. The initial estimates for the noise variances ψ_0^2 and σ_1^2 are determined as the variances of the respective vectors $\hat{\mathbf{x}}_0$ and $\mathbf{w}_1$.

For N, various values ranging between 1 and 25 are used and the performance of the filter for each particular value is evaluated in Section IV.

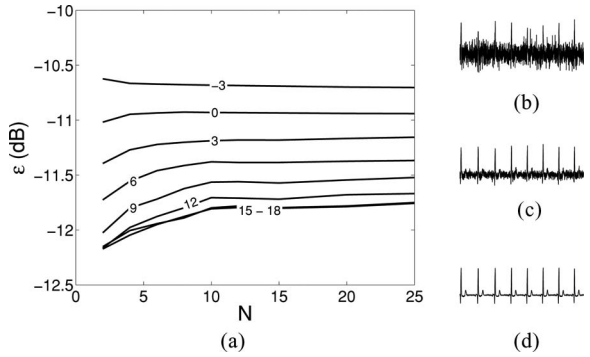

Fig. 4. In (a), the performance of the developed Kalman filter, expressed in terms of the normalized MSE ϵ, is plotted as a function of the SNR and the number of ECG complexes N used in the estimation of the noise covariances. The SNR values corresponding to each of the lines are provided in the graphs and expressed in decibel. In (b)–(d), examples of the TWA signals are plotted (each 4 s long). The SNRs of these signals are -3, 6, and 24 dB, respectively.

IV. Evaluation of Filter

A. TWA Signals

As mentioned earlier, the performance of the filter is assessed as a function of both N (i.e., the number of residuals $\boldsymbol{\rho}$ averaged for robust estimation of the process noise covariance) and the SNR, using the TWA signals of 23 different patients. The performance is quantified by calculating ϵ, the normalized MSE between the filtered ECG signals $\hat{\mathbf{x}}$ and the original ECG signals $\mathbf{x}$ used (i.e., the signals without additive noise) as follows:

$$\epsilon = \frac{\sum_k \left(\mathbf{x}_k - \hat{\mathbf{x}}_k\right)^T \left(\mathbf{x}_k - \hat{\mathbf{x}}_k\right)}{\sum_k \mathbf{x}_k^T \mathbf{x}_k} \tag{15}$$

where the summation indicates that ϵ is averaged over all heartbeats in the TWA signals. In addition, as the TWA signals comprise 12 individual ECG signals, ϵ is averaged over these signals as well (not indicated in (15) for clarity). Note that the original ECG signals $\mathbf{x}$ are not completely free of noise. However, as the amplitude of this noise is small compared to the amplitude of the ECG signals, the effect of this noise on the calculated ϵ values is small and disregarded in further discussions on the performance of the filter.

In Fig. 4, the normalized MSE ϵ is plotted as a function of both N and the SNR. The results in Fig. 4 are averaged over all the 23 TWA patients.

From Fig. 4(a) it can be seen that for the ECG signals with SNR of -3 dB and 0 dB, ϵ decreases with increasing N. That is, for the 0-dB signal, ϵ increases until $N = 5$, and decreases slightly for larger N. For the ECG signals with SNR larger than 0 dB, ϵ increases with N. These findings can be explained as follows. The fact that for high-SNR ECG signals ϵ is minimal for small N stems from the fact that, with almost no noise present, most variations in the ECG signals are of physiological origin. Since a large value for N causes slow adaptation of the process noise covariance, a large N would yield underestimation of the process noise covariance λ^2, and consequently, a too small weight ascribed to newly arriving data. Low-SNR signals, on the other hand, are mostly affected by measurement noise rather than by morphological variability. In this case, large values for N ensure that the process noise covariance is not overestimated, hence rendering the weight ascribed to newly arriving data relatively small.

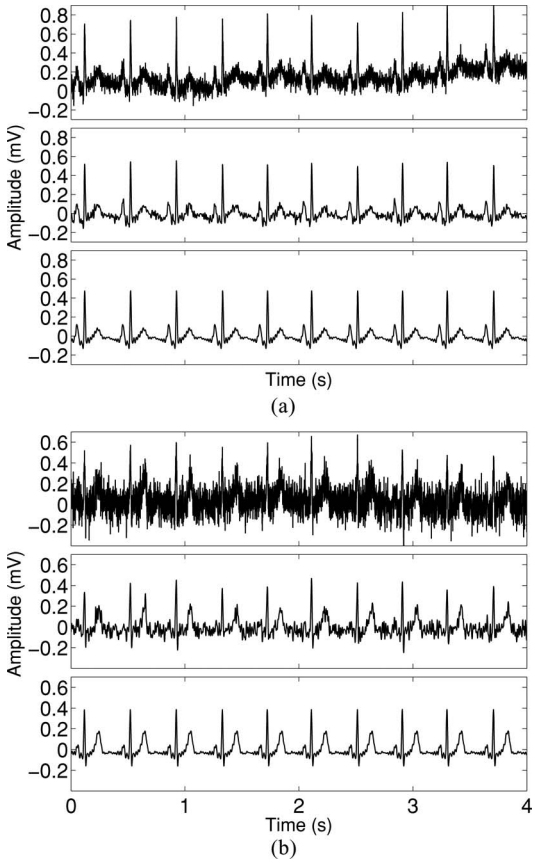

Fig. 5. Examples of TWA signals before filtering (top graph), after preprocessing (center graph), and after filtering with the developed adaptive Kalman filter (bottom graph) filtering, with $N = 5$. In (a), the SNR of the TWA signal is 8 dB, and in (b), this SNR = 0 dB.

From the aforementioned discussion, it is straightforward to state that the choice for N involves a tradeoff between robustness against measurement noise and flexibility of the process noise estimation. In Fig. 5, the performance of the developed filter (with N chosen as 5) is exemplified on two TWA signals.

To illustrate the tradeoff between robustness and flexibility of the filter, in Fig. 4(a), it can be seen that for all recordings, specifically the high-SNR recordings, the normalized MSE ϵ saturates at about -12 dB. This saturation is due to the persistent underestimation of the process noise covariance for large N. As a result, the output of the filter cannot fully keep track of morphological variations in the TWA signal, leading, in this particular case, to an estimation error of -12 dB. To ensure optimal performance of the developed filter, the choice for N should be based on expected signal behavior. For the TWA signals, N is chosen as 5. Due to the larger measurement noise amplitudes anticipated in the fetal and neonatal ECG signals that will be discussed shortly, for these signals N will be chosen as 10. From the relatively small gradients in Fig. 4(a), it can, nevertheless, be concluded that the specific choice for N does not strongly affect the performance of the filter. In other words,

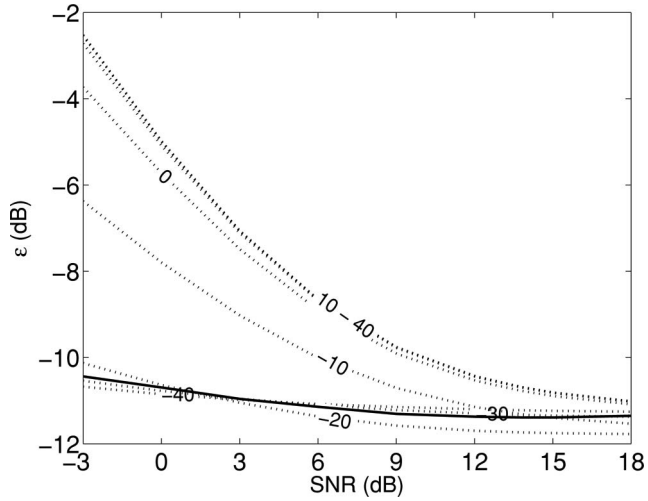

Fig. 6. Plot of the performance of the developed Kalman filter, expressed in terms of ϵ, as a function of the SNR of the ECG signals. The performance of the filter with adaptive noise-covariance estimation is indicated with the solid line (N is chosen as 5). The performances of the filters with fixed noise covariance are represented by the dotted lines. The selected noise covariances λ^2 are indicated in the graphs. Due to overlapping, the graphs for 10, 20, 30, and 40 dB are jointly labeled as 10–40.

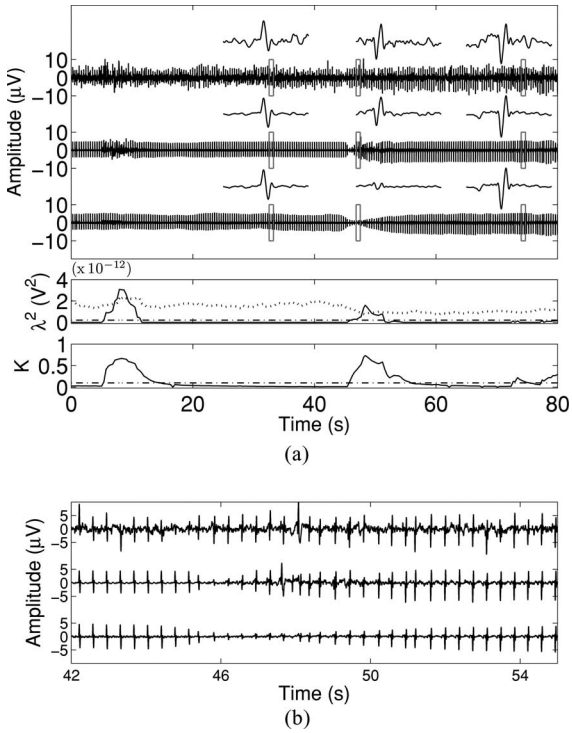

Fig. 7. (a) Top panel shows a fetal ECG signal recorded from the maternal abdomen, before filtering, but after preprocessing (top graph), after filtering with the adaptive Kalman filter (center graph), and after filtering with the fixed Kalman filter (bottom graph). The ECG complexes indicated in the rectangles are shown zoomed in in the accompanying graphs. The second panel shows the estimated process noise covariance λ^2 for the adaptive Kalman filter (solid line), the process noise covariance for the fixed Kalman filter (dash-dot line), and the estimated measurement noise covariance σ^2 (dotted line). The bottom panel shows the Kalman gain $\mathbf{K}$ for both the adaptive (solid line) and fixed (dash-dot line) filters. For the estimation of λ^2 and $\mathbf{K}$ in the adaptive Kalman filter, N is chosen equal to 10. Since the process noise covariance is often estimated as 0 [see (15)], λ^2 cannot be expressed in decibel as in Fig. 6. Hence, λ^2 is expressed here in an absolute sense (i.e., in V^2). In (b), a zoom of the top panel between 42 and 55 s is depicted.

irrespective of the value used for N, the performance of the developed filter will be about the same. Based on this conclusion, it can be argued that with the adaptive noise-covariance estimation, the problem of *a priori* selecting the noise covariances, mentioned in the beginning of Section II-D, is overcome and replaced by the much less critical problem of selecting N. To substantiate this remark, in Fig. 6, the performance of the developed filter is compared with the performance of the same Kalman filter, but now with the adaptive noise-covariance estimation replaced by a fixed *a priori* estimation. The estimation of the measurement noise covariance is kept the same for both filters. The values for λ^2 in this comparison range between -40 and 40 dB and are defined relative to the amplitude of the ECG signals.

From Fig. 6, it can be seen that for the Kalman filter with fixed process noise covariance, for simplicity, from here on referred to as the fixed Kalman filter, the performance improves with decreasing λ^2 until $\lambda^2 = -20$ dB; from here on, the performance slightly deteriorates. This behavior is consistent with the aforementioned discussion that the choice for λ^2 affects the performance of the filter more strongly than the choice for N. Underestimation of λ^2 will result in a relatively small Kalman gain and, hence, little flexibility of the filter output to account for morphological variations in the ECG. Overestimation of λ^2, on the other hand, yields a relatively large Kalman gain leading to an output of the filter that not only accounts for morphological variations but also for measurement noise. Straightforwardly, an optimal value exists for which the performance of the filter is rather good. In the case of the TWA signals used in Fig. 6, this optimal value is around -20 dB, but for other ECG signals, this value needs to be re-evaluated.

B. Fetal ECG Signals

When comparing the performance of the fixed filter with $\lambda^2 = -20$ dB in Fig. 6 with the performance of the adaptive Kalman filter, it can be concluded that this performance is about the same, with the fixed Kalman filter performing slightly better for high-SNR signals. It, however, needs to be emphasized once more that the fixed Kalman filter operates with a process noise covariance that is fixed at a rather optimal value. When employing both filters for long-term monitoring tasks, a proper *a priori* selection of λ^2 becomes virtually impossible. Moreover, even when it is possible to *a priori* assess for which value of λ^2 the filter will perform optimally, the long-term nature of the recording renders the ECG signals likely to exhibit dynamical variations. With these variations, the *a priori* assessed value for λ^2 needs to be re-evaluated and adapted. Since the adaptive Kalman filter is capable of making this adaptation, it is expected to outperform the fixed Kalman filter for long-term monitoring tasks. To study the behavior of both filters for a dynamical variation in the ECG signal, in Fig. 7, a transabdominally recorded fetal ECG signal

is presented, exhibiting significant morphological variation as a result of movement.

From Fig. 7, it can be seen that, as expected, the estimated adaptive process noise covariance λ^2 increases when variations in the ECG signal occur (e.g., around 5 and 45 s). Especially, the variation in the ECG around 45 s is of relevance because the fetus shows significant movement here, as was demonstrated by an echocardiographic analysis performed simultaneously with the ECG recording. After the movement epoch, the fetus has taken a slightly different orientation with respect to the electrodes on the maternal abdomen, affecting the morphology of the ECG signal. When the movement sets in, the increased process noise covariance causes an increase in the Kalman gain, hence ascribing more weight to the newly arriving data, as intended.

When comparing the Kalman gains of both the adaptive and fixed Kalman filter—for the latter, λ^2 is empirically set at 2×10^{-13} V^2 —it strikes that the gain of the fixed filter shows significantly fewer fluctuations than the gain of the adaptive Kalman filter. Consistent with the aforementioned discussion on long-term monitoring, this rigidity of the fixed Kalman filter renders the fixed Kalman filter less capable of accounting for fast morphological changes in the ECG. This statement is substantiated by the filtered ECG signals in the top panel of Fig. 7. After the movement of the fetus, the fixed Kalman filter needs about 10 s to completely adapt its output to the new ECG morphology [see Fig. 7(b)], whereas adaptation by the Kalman filter with adaptive noise covariance is more than twice faster.

Regarding the measurement noise covariance σ^2, it can be seen in Fig. 7 that it decreases as a result of the fetal movement. This decrease can be explained by the fact that, whereas the noise amplitude remains about the same, the signal amplitude increases as a result of the movement. This rise in the ECG amplitude can be explained by, e.g., a movement-inflicted reduction in the heart–electrode distance.

C. Neonatal ECG

For the (maternal) movement artifact occurring around 5 s in Fig. 7, it can be argued that the fixed Kalman has an advantage over the adaptive Kalman filter in that it is less affected by the movement artifact. However, when examining the original, unfiltered ECG signal, this signal appears so corrupted that it hardly contains any ECG information. The output generated by the fixed Kalman filter, therefore, mostly comprises information from earlier heartbeats. Naturally, in case of a local artifact that only affects a single ECG signal also the adaptive Kalman filter does not update the estimated ECG. That is, when the artifact occurs locally, the estimated measurement noise covariance will be relatively large, significantly decreasing the Kalman gain. Conversely, in cases where the artifact occurs simultaneously at more than one location, the estimated measurement noise will decrease and both the estimated process noise covariance and Kalman gain will increase. This process is illustrated in Fig. 8 in which a neonatal ECG, recorded in an incubator with textile electrodes, is depicted.

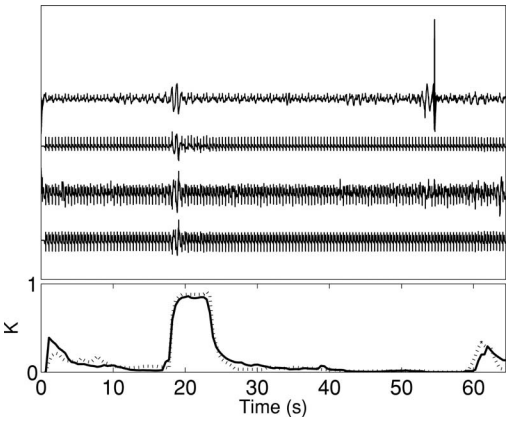

Fig. 8. In the top panel, two neonatal ECG signals recorded in the incubator using textile electrodes are depicted. The first and third signals from above are the preprocessed ECG signals, and the second and fourth signals are the corresponding signals after filtering with the adaptive Kalman filter. In the bottom panel, the Kalman gains for both filtered signals are plotted. The solid line corresponds to the upper ECG signal, and the dotted line corresponds to the lower ECG signal.

The neonatal ECG signals shown in Fig. 8 illustrate that in case of an artifact that occurs in more than one ECG signal at the same time, the newly arriving data are ascribed more weight (i.e., the Kalman gain is increased) to ensure rapid updating of the ECG estimate. For local artifacts, such as in the upper neonatal ECG signal occurs around 55 s, the Kalman gain changes only little, ensuring that the artifact is no longer present in the filtered ECG signal.

V. Discussion

A. Limitations of Bayesian Model

In the derivation of the adaptive Kalman filter, several assumptions are made for mathematical simplicity, but that might limit the applicability of the filter. For one, the ECG is assumed to be normally distributed, or equivalently, both the process and measurement noise are assumed to be Gaussian. In addition, the measurement and process noise are assumed to be uncorrelated. The latter assumption might limit the performance of the filter to some extent as, in particular, the process noise generally exhibits spatial correlation across the individual ECG signals. This limitation seems, however, small, as evidenced by the results. The same thing holds for the assumption of Gaussian noise. Although this assumption might impose a rather severe limitation to the filter's applicability, the evaluation of the filter on the fetal and neonatal ECG signals—that are corrupted by physiological noise, rather than by Gaussian noise—demonstrates that the filter also performs relatively well for non-Gaussian ECG signals. Similarly, the implicit assumption that the (prior) probability distributions in (5) and (10) sum to 1 might not be fully satisfied, but is demonstrated by our results to have only marginal effect (if any) on the performance of the filter. Besides providing a rather elegant solution to the filter problem, the mentioned mathematical simplification also relaxes the computational complexity of the filter, rendering an implementation of

the filter in MATLAB (The Mathworks, Inc.) capable of filtering at least 12 ECG signals simultaneously in real time.

Another decision that might limit the applicability of the developed filter is the fixed choice for the length of the ECG complexes (i.e., 120% of the mean interval between consecutive heartbeats). In cases of significant heart rate variability, it can occur that a single ECG complex contains information that originates from two consecutive heartbeats. Problems associated with this fixed ECG complex length can, however, be circumvented by only including those parts of the ECG complex that correspond to the same heartbeat in the calculation of the measurement noise, process noise, and Kalman gain. Upon reassembling the filtered ECG complexes into a filtered ECG signal that is composed of a multitude of heartbeats, the redundant parts of the filtered ECG complexes can be omitted.

B. Inaccuracies in Estimation of Measurement Noise

The accurate estimation of the measurement noise covariance is rather critical for the performance of the adaptive Kalman filter. When this covariance is overestimated, all ECG signal variations will be ascribed to measurement noise, and hence, the process noise covariance will be underestimated, rendering the filter less capable of quickly adapting to dynamical signal variations. Conversely, underestimation of the measurement noise covariance leads to overestimation of the process noise covariance, causing the filter to also ascribe more weight to ECG complexes that are corrupted by measurement noise. The estimation of the measurement noise covariance is performed by exploiting the spatial correlation of simultaneously recorded multichannel ECG signals. In a simplified model, all ECG signals can be assumed to originate from the same 3-D source, and hence, three independent ECG signals should be enough to predict the morphology of a fourth ECG signal. Those parts of the fourth ECG signal that cannot be accounted for by the three other ECG signals can, therefore, be marked as noise contributions. One of the main drawbacks of this approach is that for enhancing the ECG of one signal, at least three other ECG signals need to be recorded. However, for most clinical applications, several ECG signals are recorded, and even if there is no clinical relevance of recording several channels, it hardly constitutes a technological challenge to record more than three ECG signals simultaneously. Another drawback lies in the adopted model for the spatial correlation between the ECG signals. In this model, all ECG signal components that cannot be accounted for by the 3-D source are taken to be measurement noise, leading to underestimation of the process noise. By using more than four ECG signals, e.g., a standard configuration of 12 electrodes, this problem can be largely overcome.

C. Data Used in Filter Evaluation

As mentioned earlier, the TWA signals comprise a rather ideal set for quantitative evaluation of the developed filter. Here, however, distinction needs to be made between TWA signals that exhibit morphological variability at the microvolt level and macroscopic TWA signals that exhibit substantially more morphological variability. In our evaluation, only the latter TWA signals were used. With regard to the preprocessing of the TWA (and fetal and neonatal ECG) signals, the high-pass filter is expected to slightly distort the susceptible ST segment. Preprocessing is nevertheless performed to ensure that the measurement noise has indeed zero mean, as assumed in Section II. The effect of omitting the preprocessing, in order to yield as little distortion of the filtered ECG signals as possible, on the performance of the filter is a subject of future research. Also subject of future research is the evaluation of the filter's performance on ECG signals that are even more ideal than the TWA signals, such as ECG signals with isolated or slow pattern changes.

For noninvasive fetal ECG recordings performed on the maternal abdomen, the aforementioned requirement of at least four ECG signals, of which three are linearly independent, could be troublesome. Approximately between the 28th and 34th week of gestation, the fetus is covered by a waxy layer (i.e., vernix caseosa) that electrically isolates the fetus, apart from a few places that are hypothesized to be over the oronasal area and the umbilical cord [24]. As a result of this layer, preferred conduction paths for the ECG signals arise, potentially making the number of independent ECG signals drop below three. Additional research to determine whether the presented method indeed fails in the presence of vernix caseosa is, however, required.

Additional research is also required to assess whether the application field of the filter can be extended. For ECG enhancement, the filter basically operates by averaging consecutive ECG complexes and varying the number of complexes included in the average, based on the amount of variation in the data. This approach can, however, also be applied in enhancement of other quasi-periodical signals that may vary due to either changes in the process or changes in the measurement noise. An example of such an application is the SNR enhancement of event-related potentials in electroencephalography studies [25].

VI. Conclusion

In this paper, a Kalman filter with adaptive noise-covariance estimation has been developed and evaluated on a variety of ECG signals to assess whether the filter is capable of enhancing the SNR of these signals, while at the same time preserving clinically relevant morphological variations in the ECG. The filter operates by sequentially estimating the measurement and process noise covariances and uses these covariances to estimate the Kalman gain and update the estimated ECG complexes. In cases where the variations between consecutive ECG complexes can no longer be explained as measurement noise, the variations are taken to be morphological variations and the process noise covariance is increased. This, in turn, leads to an increase of the Kalman gain, and consequently, more weight is ascribed to the newly arriving ECG complex.

The performance of the filter is compared with the performance of a similar Kalman filter with fixed process noise covariance. For this fixed Kalman filter, the process noise covariance needs to be *a priori* estimated, and hence, to ensure adequate performance of the filter, requires rather detailed information on the ECG signal dynamics. The comparison between the fixed and adaptive Kalman filters demonstrates that the adaptive filter

performs almost as a fixed Kalman filter with optimally chosen process noise covariance. In addition, for long-term monitoring tasks in which the ECG signal characteristics change, the adaptive Kalman filter is capable of quickly adapting its noise estimation to match the filter's output to the new input. The fixed Kalman filter, on the other hand, needs about 10 s to adjust its output due to its less flexible estimation of the Kalman gain.

ACKNOWLEDGMENT

The authors would like to thank S. Bouwstra at the Eindhoven University of Technology for providing the textile-recorded neonatal ECG signals.

REFERENCES

[1] K. Chang, "Portable obstructive sleep apnea screening system using overnight ECG and a PDA-based wireless transmission system," *Telemed. J. E. Health*, vol. 15, no. 4, pp. 353–61, May 2009.

[2] W. Chen, C. Sonntag, F. Boesten, S. Bambang Oetomo, and L. Feijs, "A design of power supply for neonatal monitoring with wearable sensors," *J. Ambient Intell. Smart Env., Spec. Issue Wearable Sens.*, vol. 1, no. 2, pp. 185–196, 2009.

[3] A. Gruetzmann, S. Hansen, and J. Müller, "Novel dry electrodes for ECG monitoring," *Physiol. Meas.*, vol. 28, pp. 1375–1390, 2007.

[4] E. Hon, "Instrumentation of fetal heart rate and fetal electrocardiography. II. A vaginal electrode," *Amer. J. Obstet. Gynecol.*, vol. 86, pp. 772–84, 1963.

[5] J. H. V. Bemmel and H. van der Weide, "Detection procedure to represent the foetal heart rate and electrocardiogram," *IEEE Trans. Biomed. Eng.*, vol. 13, no. 4, pp. 175–82, Oct. 1966.

[6] M. Oehler, V. Ling, K. Melhorn, and M. Schilling, "A multichannel portable ECG system with capacitive sensors," *Physiol. Meas.*, vol. 29, pp. 783–793, 2008.

[7] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*, 3rd ed. ed. Englewood Cliffs, NJ: Prentice Hall, 2009.

[8] S. Abboud and D. Sadeh, "Spectral analysis of the fetal electrocardiogram," *Comput. Biol. Med.*, vol. 19, no. 6, pp. 409–415, 1989.

[9] J. Bailey, A. Berson, A. Garson Jr., L. Horan, P. Macfarlane, D. Mortara, and C. Zywietz, "Recommendations for standardization and specifications in automated electrocardiography: Bandwidth and digital signal processing. A report for health professionals by an ad hoc writing group of the Committee on Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiology, American Heart Association," *Circulation*, vol. 81, pp. 730–739, 1990.

[10] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1986.

[11] K. Greene, G. Dawes, H. Lilja, and K. Rosén, "Changes in the ST waveform of the fetal lamb electrocardiogram with hypoxemia," *Amer. J. Obstet. Gynecol.*, vol. 144, pp. 950–958, Dec. 1982.

[12] R. Sameni, M. Shamsollahi, C. Jutten, and G. Clifford, "A nonlinear bayesian filtering framework for ecg denoising," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2172–2185, Dec. 2007.

[13] E. Frank, "An accurate, clinically practical system for spatial vectorcardiography," *Circulation*, vol. 13, pp. 737–749, 1956.

[14] A. Ben-Israel and T. Greville, *Generalized Inverses*, 2nd ed. Berlin, Germany: Springer, 2003.

[15] J. De Freitas, M. Niranjan, and A. Gee, "Hierarchical Bayesian-Kalman models for regularisation and ARD in sequential learning," Dept. Eng., Cambridge Univ., Cambridge, U.K., Tech. Rep., 1998.

[16] Y. Bar-Shalom and X. Li, *Estimation and Tracking: Principles, Techniques and Software*. Norwood, MA: Artech House Inc., 1993.

[17] R. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[18] A. Jazwinski, "Adaptive filtering," *Automatica*, vol. 5, pp. 475–485, 1969.

[19] 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, and H. E. Stanley. (2000, Jun. 13). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. *101(23)*, pp. e215–e220. Available: http://circ.ahajournals.org/cgi/content/full/101/23/e215.

[20] S. Butterworth, "On the theory of filter amplifiers," *Exp. Wireless Wireless Eng.*, vol. 7, pp. 536–541, Oct. 1930.

[21] R. Vullings, C. Peters, R. Sluijter, M. Mischi, S. Oei, and J. Bergmans, "Dynamic segmentation and linear prediction for maternal ECG removal in antenatal abdominal recordings," *Physiol. Meas.*, vol. 30, pp. 291–307, 2009.

[22] S. Martens, R. Sluijter, S. Oei, and J. Bergmans, "Improving QRS detection in multi-channel electrocardiography by principal component analysis," presented at the Int. Federation for Medical and Biological Engineering, Prague, Czech Republic, Nov. 2005.

[23] F. Gritzali, G. Frangakis, and G. Papakonstantinou, "Detection of the P and T waves in an ECG," *Comput. Biomed. Res.*, vol. 22, no. 1, pp. 83–91, Feb. 1989.

[24] J. Oldenburg and M. Macklin, "Changes in the conduction of the fetal electrocardiogram to the maternal abdominal surface during gestation," *Am. J. Obstet. Gynecol.*, vol. 129, no. 4, pp. 425–433, Oct. 1977.

[25] M. Ford, S. Sands, and L. Lew H, "Overview of artifact reduction and removal in evoked potential and event-related potential recordings," *Phys. Med. Rehabil. Clin. N. Amer.*, vol. 15, no. 1, pp. 1–17, Feb. 2004.

Authors' photographs and biographies not available at the time of publication.