A Nonlinear Bayesian Filtering Framework for ECG Denoising

Reza Sameni*, *Student Member, IEEE*, Mohammad B. Shamsollahi, *Member, IEEE*, Christian Jutten, *Senior Member, IEEE*, and Gari D. Clifford, *Senior Member, IEEE*

Abstract—In this paper, a nonlinear Bayesian filtering framework is proposed for the filtering of single channel noisy electrocardiogram (ECG) recordings. The necessary dynamic models of the ECG are based on a modified nonlinear dynamic model, previously suggested for the generation of a highly realistic synthetic ECG. A modified version of this model is used in several Bayesian filters, including the Extended Kalman Filter, Extended Kalman Smoother, and Unscented Kalman Filter. An automatic parameter selection method is also introduced, to facilitate the adaptation of the model parameters to a vast variety of ECGs. This approach is evaluated on several normal ECGs, by artificially adding white and colored Gaussian noises to visually inspected clean ECG recordings, and studying the SNR and morphology of the filter outputs. The results of the study demonstrate superior results compared with conventional ECG denoising approaches such as bandpass filtering, adaptive filtering, and wavelet denoising, over a wide range of ECG SNRs. The method is also successfully evaluated on real nonstationary muscle artifact. This method may therefore serve as an effective framework for the model-based filtering of noisy ECG recordings.

Index Terms—ECG denoising, Kalman filtering, model-based filtering, nonlinear Bayesian filtering, adaptive filtering.

I. INTRODUCTION

THE extraction of high-resolution cardiac signals from a noisy electrocardiogram (ECG) remains an important problem for the biomedical engineering community. Despite of the rich literature in this field, there are still many clinical applications that lack reliable signal processing tools to extract the weak ECG components contaminated with background noise and permit the measurement of subtle features in the ECG. The numerous noncardiac ECG contaminants, such as electromyographic (EMG) noise, overlap with the cardiac components in the frequency domain, particularly in the 0.01 Hz to 100 Hz range. Bandpass filtering is therefore inadequate to suppress such contaminants [1], [2].

Ensemble Averaging (EA) is another common approach for the extraction of small cardiac components from the noise contaminated ECG. However, as EA requires the averaging of many beats, the subtle but important inter-beat variations in the cardiac cycle are lost in the averaging procedure [3]. As an improvement over EA, classical *Adaptive Filter* (AF) architectures have also been used for the noise cancellation of ECGs containing baseline wander, power line interference, EMG noise, and motion artifacts [4], [5].

For stationary signals, the *Wiener Filter* (WF) is the optimal linear filtering technique in the *Minimum Mean Square Error* (MMSE) sense, applied either in a causal sense in the time-domain, or as the noncausal WF applied in the frequency domain. However, the WF is not expected to (and does not) give good results for a noisy ECG, due to the nonstationary nature of the cardiac signal. In some related works, filtering approaches have been proposed based on time-frequency [6], [3], and time-scale [7], [8] WFs. The intuition behind the use of the time-frequency or wavelet transforms in these applications is to apply the WF in two domains, to facilitate the tracking of ECG nonstationarities.

Wavelet Denoising (WD) is now a common practice for denoising of signals having multiresolution characteristics such as the cardiac signal in the ECG. Donoho and Johnstone [9] proposed a soft thresholding method for the so-called *shrinkage* of the noise components in the wavelet domain. Their approach together with some *ad hoc* variants of it, have since been used for many applications, including high-resolution ECG denoising [7], [10], [11]. In these cases, the model of the ECG is essentially based on the frequency content of the ECG and to some degree, the localization of the ECG peaks in time.

Statistical techniques such as *Principal Component Analysis* (PCA) [12], *Independent Component Analysis* (ICA) [13], [14] and *Neural Networks* (NNs) [15] have also been used to extract a statistical-based model of the signal and noise, allowing the removal of in-band noise by discarding the dimensions corresponding to noise. Although these are powerful in-band noise filtering schemes, the model is fairly arbitrary and, unless the basis functions are *trained* on a global set of beat-types, can be extremely sensitive to small changes in either the signal or the noise. Nonlinear projective filtering [16], is another similar method of forming an *ad hoc* statistical model of the ECG and

*R. Sameni is with the Biomedical Signal and Image Processing Laboratory (BiSIPL), School of Electrical Engineering, Sharif University of Technology, Azadi Avenue, Tehran, Iran, and also with the Laboratory of Grenoble Image Parole Signal Automatique (GIPSA), Institute National Polytechnique de Grenoble (INPG), 46, avenue Felix Viallet, 38031 Grenoble, France (e-mail: r_sameni@mehr.sharif.edu).

M. B. Shamsollahi is with the Biomedical Signal and Image Processing Laboratory (BiSIPL), School of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mail: mbshams@sharif.ir).

C. Jutten is with the Laboratory of Grenoble Image Parole Signal Automatique (GIPSA), Institute National Polytechnique de Grenoble (INPG), 38031 Grenoble, France (e-mail: christian.jutten@inpg.fr).

G. D. Clifford is with the Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02142 USA (e-mail: gari@mit.edu).

although potentially extremely powerful, is extremely computationally intensive and highly sensitive to outliers and changes in data dimensionality through changes in background noise. These methods can be generally considered as members of the AF family, in which some model of the cardiac signal is constructed (either explicitly or implicitly), and used as a reference signal to constrain the filter to improve the signal-noise separation performance. In this paper, we demonstrate that by using a realistic model to describe the quasi-periodic behavior of the ECG, the idea of model-based filtering may be further extended to a general Bayesian filtering framework for online ECG denoising.

The framework presented in this paper is based on a dynamic model, previously developed for the generation of synthetic PQRST complexes with their relationship to the beat-to-beat RR-interval timing [17], [18]. Considering the simplicity and flexibility of this model it is reasonable to assume that it can describe the dynamics of a broad class of normal and abnormal ECGs. Moreover, as shown in [19], [20] the applications of this model are not limited to ECGs and may be extended to other quasi-periodic signals. In recent works [21], [22] the authors employed this model to develop an *Extended Kalman Filter* (EKF) for noisy ECG filtering. In this paper, this synthetic ECG model is further modified and used in conjunction with several nonlinear Bayesian filtering approaches such as the EKF, *Extended Kalman Smoother* (EKS), and *Unscented Kalman Filter* (UKF). Furthermore, the model parameter selection is automated in order to adapt the method to different ECG channels. In order to evaluate the proposed method, different portions of white and colored Gaussian noises have been added to visually inspected clean ECGs recorded from various ECG lead configurations with differing morphologies. The Signal-to-Noise Ratios (SNRs) of the filter outputs have then been compared with several conventional ECG denoising schemes. An example of the filtering performance, in presence of real nonstationary muscle artifact (MA) is also presented. The results demonstrate that the proposed filters can accurately track the ECG signal in very low SNR conditions, where the cardiac signal is almost lost in background noise. This method is believed to serve as a novel framework for the model-based extraction of high-resolution ECG signals from noisy measurements. In particular, since the presented method provides an accurate separation of nonlinear and nonstationary signal and in-band noise, it is hoped that this method will be suitable for applications such as the extraction of ECG late potentials from high-resolution ECGs [23], or the noninvasive extraction of fetal ECG from the signals recorded from an array of electrodes placed on the maternal abdomen [24], where due to the low SNR of these signals, conventional filtering approaches do not give satisfactory results.

The paper is organized as follows. In Sections II and III, the required theoretical background and the previously developed synthetic ECG generator are described. The details of the proposed tracking and filtering methods are presented in Section IV. In Section V some implementation issues and a description of standard ECG denoising methods that have been used as a benchmark for the proposed methods are presented. The results of the different filtering methods are presented and discussed in Section VI, and the final section is devoted to some concluding remarks concerning the presented approach.

II. REVIEW OF THE BAYESIAN FILTERING THEORY

A classical problem in estimation theory is the estimation of the hidden states, that are observable through a set of measurements, of a system with an underlying dynamic model. The well-known *Kalman filter* (KF) is one such method and under certain general constraints, it can be proved to be the optimal filter in the MMSE sense [25]. The conventional KF assumes a linear model for the system dynamics and observation equations. In practice however, most systems are nonlinear in nature and in order to extend the idea of the conventional KF to such systems, several variants of the original KF have been developed. In this section, the theoretical foundations of some of these extensions are briefly reviewed to facilitate the presentation of the proposed methods.

A. Extended Kalman Filter

The *Extended Kalman Filter* (EKF) is an extension of the standard KF to nonlinear systems. Consider a discrete-time nonlinear system with the unobserved underlying state vector $\mathbf{x}_k$ and observation vector $\mathbf{y}_k$ at time instant k. A dynamic model of this system may be represented as follows:

$$\begin{cases} \mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{w}_k, k) \\ \mathbf{y}_k = g(\mathbf{x}_k, \mathbf{v}_k, k) \end{cases} \tag{1}$$

where $f(\cdot)$ is the state evolution function and $g(\cdot)$ represents the relationship between the state vector and the observations. The process and measurement noise vectors are $\mathbf{w}_k$ and $\mathbf{v}_k$ respectively, with associated covariance matrices $Q_k = E\{\mathbf{w}_k \mathbf{w}_k^T\}$ and $R_k = E\{\mathbf{v}_k \mathbf{v}_k^T\}$. The initial estimate of the state vector is also assumed to be known and is given by: $\bar{\mathbf{x}}_0 = E\{\mathbf{x}_0\}$, with $P_0 = E\{(\mathbf{x}_0 - \bar{\mathbf{x}}_0)(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T\}$.

In order to use the KF formalism for this system, it is necessary to derive a linear approximation of (1) near a desired reference point $(\hat{\mathbf{x}}_k, \hat{\mathbf{w}}_k, \hat{\mathbf{v}}_k)$ [26], [27]. This leads to the following linear approximate model:

$$\begin{cases} \mathbf{x}_{k+1} \approx f(\hat{\mathbf{x}}_k, \hat{\mathbf{w}}_k, k) + A_k(\mathbf{x}_k - \hat{\mathbf{x}}_k) + F_k(\mathbf{w}_k - \hat{\mathbf{w}}_k) \\ \mathbf{y}_k \approx g(\hat{\mathbf{x}}_k, \hat{\mathbf{v}}_k, k) + C_k(\mathbf{x}_k - \hat{\mathbf{x}}_k) + G_k(\mathbf{v}_k - \hat{\mathbf{v}}_k) \end{cases} \tag{2}$$

where

$$A_k = \frac{\partial f(\mathbf{x}, \hat{\mathbf{w}}_k, k)}{\partial \mathbf{x}}\big|_{\mathbf{x}=\hat{\mathbf{x}}_k}$$
$$F_k = \frac{\partial f(\hat{\mathbf{x}}_k, \mathbf{w}, k)}{\partial \mathbf{w}}\big|_{\mathbf{w}=\hat{\mathbf{w}}_k}$$
$$C_k = \frac{\partial g(\mathbf{x}, \hat{\mathbf{v}}_k, k)}{\partial \mathbf{x}}\big|_{\mathbf{x}=\hat{\mathbf{x}}_k}$$
$$G_k = \frac{\partial g(\hat{\mathbf{x}}_k, \mathbf{v}, k)}{\partial \mathbf{v}}\big|_{\mathbf{v}=\hat{\mathbf{v}}_k}. \tag{3}$$

Moreover, to simplify the matrix notations, the F_k and G_k matrices are usually absorbed into the noise covariance matrices as follows:

$$F_k Q_k F_k^T \rightarrow Q_k, \quad G_k R_k G_k^T \rightarrow R_k.$$

With these notations, the EKF algorithm may be summarized as follows:

$$\hat{\mathbf{x}}_{k+1}^- = f(\hat{\mathbf{x}}_k^+, \mathbf{w}, k)|_{\mathbf{w}=\bar{\mathbf{w}}_k}$$
$$\mathbf{r}_k = \mathbf{y}_k - g(\hat{\mathbf{x}}_k^-, \mathbf{v}, k)|_{\mathbf{v}=\bar{\mathbf{v}}_k}$$
$$K_k = P_k^- C_k^T \left[C_k P_k^- C_k^T + R_k \right]^{-1}$$
$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + K_k \mathbf{r}_k$$
$$P_{k+1}^- = A_k P_k^+ A_k^T + Q_k$$
$$P_k^+ = P_k^- - K_k C_k P_k^- \tag{4}$$

where by definition $\mathbf{r}_k$ is the *innovation* signal, $\bar{\mathbf{w}}_k = E\{\mathbf{w}_k\}$, $\bar{\mathbf{v}}_k = E\{\mathbf{v}_k\}$, $\hat{\mathbf{x}}_k^- = \hat{E}\{\mathbf{x}_k|\mathbf{y}_{k-1}, \ldots, \mathbf{y}_1\}$ is the *a priori* estimate of the state vector in the kth stage using the observations $\mathbf{y}_1$ to $\mathbf{y}_{k-1}$, and $\hat{\mathbf{x}}_k^+ = \hat{E}\{\mathbf{x}_k|\mathbf{y}_k, \ldots, \mathbf{y}_1\}$ is the *a posteriori* estimate of this state vector after using the kth observation $\mathbf{y}_k$. P_k^- and P_k^+ are defined in the same manner to be the *a priori* and *a posteriori* estimates of the state vector covariance matrices before and after using the kth observation, respectively.

As it can be seen in (4), the key idea of the EKF is to linearize the nonlinear system model in the vicinity of the previous estimated point, and to recursively calculate the filter gain K_k, and the state covariance matrices P_k^- and P_k^+ from the linearized equations, while the KF time propagation is performed via the original nonlinear equation [26].

B. Extended Kalman Smoother

As with the *Kalman Smoother*, the *Extended Kalman Smoother* (EKS) uses the information of *future* observations to give better estimates of the *current* state. Due to this noncausal nature, the EKS is expected to have a better performance compared with the EKF. The EKS algorithm basically consists of a forward EKF stage followed by a backward recursive smoothing stage. Depending on the smoothing strategy, smoothing algorithms are usually classified into *fixed lag* or *fixed interval* smoothers [28]. In this paper, the fixed interval EKS is used, since the filtering procedure is carried out offline on the entirety of each ECG signal. For real-time applications of the proposed EKS methods, the fixed lag smoother is usually more appropriate.

C. Unscented Kalman Filter

For highly nonlinear systems, the linear estimate of the nonlinear model does not provide a good approximation of the model, and hence the EKF will not track the desired signal around sharp turning points (such as for the ECG). In recent years there has been great interest towards the extensions of the KF to highly nonlinear systems [29]. The *Unscented Kalman Filter* (UKF) is a filter based on the *Unscented Transform* (UT), a method for the estimation of the first and second order statistics of the outputs of highly nonlinear systems with Gaussian inputs [26]. In fact, for the UKF the linearization of the system model is no longer necessary since the prior estimate of the state covariance matrix, which is required for the Kalman gain calculations in (4), is directly estimated using the UT. The theory of the UKF and its implementation issues have already been discussed in the literature and the reader is referred to [26] for a detailed mathematical description. Note that the UKF is

TABLE I
PARAMETERS OF THE SYNTHETIC ECG MODEL IN (5)

Index(i)	P	Q	R	S	T
$\theta_i(rads.)$	$-\pi/3$	$-\pi/12$	0	$\pi/12$	$\pi/2$
a_i	1.2	-5.0	30	-7.5	0.75
b_i	0.25	0.1	0.1	0.1	0.4

numerically sensitive and the covariance matrices estimated by the UT may become semi-definite and therefore much effort has been made to achieve numerically stable versions of this algorithm. The UKF algorithm used in this paper is based on the *ReBEL* Matlab® library, previously developed for nonlinear Bayesian filtering [30] and is optimized to prevent the estimated covariance matrices from becoming semi-definite.

III. SYNTHETIC ECG GENERATOR

The dynamic equations used as the state model for the Bayesian filter variants in this paper are modifications of the synthetic ECG generator proposed by McSharry *et al.* [17]. This model has a variable number of free parameters that make it adaptable to many normal and abnormal ECGs. The dynamic model consists of a set of nonlinear dynamic state equations in the Cartesian coordinates:

$$\begin{cases} \dot{x} = \rho x - \omega y \\ \dot{y} = \rho y + \omega x \\ \dot{z} = -\displaystyle\sum_{i \in \{P,Q,R,S,T\}} a_i \Delta\theta_i \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (z - z_0) \end{cases} \tag{5}$$

where x, y, and z are the state variables, $\rho = 1 - \sqrt{x^2 + y^2}$, $\Delta\theta_i = (\theta - \theta_i)\mathrm{mod}(2\pi)$, $\theta = \mathrm{atan2}(y, x)$ is the four quadrant arctangent of the elements of x and y, with $-\pi \leq \mathrm{atan2}(y, x) \leq \pi$, and ω is the angular velocity of the trajectory as it moves around the limit cycle in the x–y plane. The a_i, b_i, and θ_i terms in (5) correspond to the amplitude, width, and center parameters of the Gaussian terms of this equation. Some typical values of these parameters taken from [17] are listed in Table I. In this model, the baseline wander of the ECG is modeled with the parameter z_0 that is assumed to be a relatively low amplitude sinusoidal component coupled with the respiratory frequency. As it is seen in (5), each of the P, Q, R, S, and T waves of the ECG waveform are modeled with a Gaussian function and are located at specific angular positions θ_i. In fact, the three dimensional trajectory generated by (5), consists of a circular limit cycle in the x–y plane that is pushed up and down as it approaches each of the θ_i. The z coordinate of this three dimensional trajectory, when plotted versus time gives the synthetic ECG. By introducing some random variations into the parameters of Table I, it is possible to generate quasi-periodic signals that appear to be realistic ECG. Clifford *et al.* [31] have recently developed methods for the estimation of the values of the model parameters for realistic ECGs based on nonlinear optimization of the parameters of (5) for a given ECG dataset. A rather similar approach is later used for the initial estimation of the proposed KF model parameters.

IV. METHODS

The KF theory and the previously developed dynamical ECG model were reviewed in preceding sections. With this general

overview, we intend to use the synthetic dynamical ECG model within a KF framework. In order to do so, the dynamic equations of (5) need to be modified for the problem of interest.

A. Modification of the Dynamic ECG Model

The dynamic equations proposed by McSharry *et al.* [17] are in Cartesian coordinates. As previously reported in [22], as a first modification, these equations can be transferred into polar coordinates. Moreover assuming the z state variable in (5) to be in millivolts, b_i's and θ_i's in radians, and time in seconds, it is clear that the a_i's are in $\text{mV}/(\text{rads.} \times \text{s})$. So in order to simplify the dimensions and later relate the model parameters with real ECG recordings, the a_i terms in (5) will be replaced with:

$$a_i = \frac{\alpha_i \omega}{b_i^2} \quad i \in \{P, Q, R, S, T\}$$

where the α_i are the peak amplitudes of the Gaussian functions used for modeling each of the ECG components, in millivolts. This definition may be verified from (5), by neglecting the baseline wander term $(z - z_0)$ and integrating the $\dot{z}$ equation with respect to t. With these changes, the new form of the dynamic equations in polar coordinates is as follows:

$$\begin{cases} \dot{r} = r(1 - r) \\ \dot{\theta} = \omega \\ \dot{z} = -\sum_{i \in \{P,Q,R,S,T\}} \frac{\alpha_i \omega}{b_i^2} \Delta\theta_i \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (z - z_0) \end{cases}$$

(6)

where r and θ are, respectively, the radial and angular state variables in polar coordinates. These new sets of equations have some benefits compared with the original equations proposed in [17]. First of all, the polar form is much simpler and its interpretation is straightforward. Accordingly, the first equation in (6) represents the radial behavior of the generated trajectory, and converges to the limit cycle of $r = 1$ for any initial value of $r \geq 1$. However, the second and third equations of (6) are independent from r, making the first differential equation redundant. Therefore, this first equation may be excluded as it does not affect the synthetic ECG (the z state variable). Another benefit of this representation is that the phase parameter θ, is an explicit state-variable that indicates the angular location of the P, Q, R, S, and T waves (Table I). This point is further used in the implementation of the filter. For the problem of interest, (6) may be further simplified by discarding the baseline wander term $(z - z_0)$. In this case, the simplified dynamic model of (6) in its discrete form, with the assumption of a small sampling period of δ is as follows:

$$\begin{cases} \theta_{k+1} = (\theta_k + \omega\delta)\text{mod}(2\pi) \\ z_{k+1} = -\sum_i \delta \frac{\alpha_i \omega}{b_i^2} \Delta\theta_i \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) + z_k + \eta \end{cases}$$

(7)

where $\Delta\theta_i = (\theta_k - \theta_i)\text{mod}(2\pi)$, η is a random additive noise that models the inaccuracies of the dynamic model (including the baseline wander), and the summation over i is taken over the number of Gaussian functions (or turning points) used for modeling the shape of the desired ECG channel. In fact, due to the flexibility of Gaussian mixtures, it is believed that by using a sufficient number of Gaussian functions, they can be fitted to

signals recorded from different ECG leads. However, in order to illustrate the general filtering framework, in this paper we only use five Gaussians to model the ECG channels containing the P, Q, R, S, and T waves.

Henceforth, θ_k and z_k are assumed as the state variables, and $\omega, \alpha_i, \theta_i, b_i$, and η are assumed as i.i.d. Gaussian random variables considered to be process noises. Following the notation of (1), the system state and process noise vectors are defined as follows:

$$\mathbf{x}_k = [\theta_k, z_k]^T$$
$$\mathbf{w}_k = [\alpha_P, \dots \alpha_T, b_P, \dots, b_T, \theta_P, \dots, \theta_T, \omega, \eta]^T \quad (8)$$

and the process noise covariance matrix is given as $Q_k = E\{\mathbf{w}_k \mathbf{w}_k^T\}$.

B. Linearization of the Nonlinear Dynamic ECG Model

In order to set up an EKF model based on the nonlinear synthetic model of (7), it is necessary to have a linearized version of the model. Consequently, the state-equation of (7) requires linearization using (2) and (3). By defining

$$\begin{cases} \theta_{k+1} = F_1(\theta_k, \omega, k) \\ z_{k+1} = F_2(\theta_k, z_k, \omega, \alpha_i, \theta_i, b_i, \eta, k) \end{cases}$$

(9)

the following equations represent the linearized model with respect to the state variables θ_k and z_k:

$$\frac{\partial F_1}{\partial z_k} = 0 \qquad \frac{\partial F_1}{\partial \theta_k} = \frac{\partial F_2}{\partial z_k} = 1$$
$$\frac{\partial F_2}{\partial \theta_k} = -\sum_{i \in \{P,Q,R,S,T\}} \delta \frac{\alpha_i \omega}{b_i^2} \left[1 - \frac{\Delta\theta_i^2}{b_i^2}\right] \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right).$$

(10)

Similarly, the linearization of (9) with respect to the process noise components yields

$$\frac{\partial F_1}{\partial \omega} = \delta \qquad \frac{\partial F_2}{\partial \eta} = 1 \qquad i \in \{P, Q, R, S, T\}$$
$$\frac{\partial F_1}{\partial \alpha_i} = \frac{\partial F_1}{\partial b_i} = \frac{\partial F_1}{\partial \theta_i} = \frac{\partial F_1}{\partial \eta} = 0$$
$$\frac{\partial F_2}{\partial \alpha_i} = -\delta \frac{\omega \Delta\theta_i}{b_i^2} exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right)$$
$$\frac{\partial F_2}{\partial b_i} = 2\delta \frac{\alpha_i \omega \Delta\theta_i}{b_i^3} \left[1 - \frac{\Delta\theta_i^2}{2b_i^2}\right] \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right)$$
$$\frac{\partial F_2}{\partial \theta_i} = \delta \frac{\alpha_i \omega}{b_i^2} \left[1 - \frac{\Delta\theta_i^2}{b_i^2}\right] \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right)$$
$$\frac{\partial F_2}{\partial \omega} = -\sum_i \delta \frac{\alpha_i \Delta\theta_i}{b_i^2} \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right).$$

(11)

C. Observation Equations

The noisy ECG recordings are assumed to be observations for the KF. The relationship between the states and observations of the KF depends on the location of the electrodes and the origin of the measurement noise. For example, motion artifacts, environmental noise or bioelectrical artifacts such as EMG or electrogastric noise, may be assumed as the measurement noises. While the measurement noise can generally contaminate the
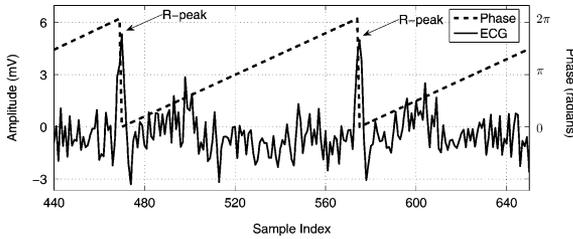
Fig. 1. Illustration of the phase assignment approach.

Fig. 2. Several cycles of the ECG phase-wrapped in the state space.

ECG in a nonlinear and non-Gaussian form, the results of this paper are based on the assumption of additive Gaussian noise.

In addition to the noisy ECG observations, the phase θ may also be added as a second observation. In fact, by studying the values of Table I, it is noticed that the R-peak is always assumed to be located at $\theta = 0$ and the ECG contents lying between two consecutive R-peaks are assumed to have a phase between 0 and 2π (or $-\pi$ and π). So by simply detecting the R-peaks an additional observation is achieved. While the R-wave detection is a rather simple and routine procedure when the signal is relatively clean, one may benefit from more sophisticated and robust approaches for very low SNR scenarios [32]. This additional phase information will also help to synchronize the dynamical KF trajectories with the reference noisy signals, without the need for manual synchronization. This RR-interval *phase warping* technique may be assumed as a generalization of the *external reference*, previously used for the synchronization of AFs for event-related signals [5], [33].

Hence the phase observations ϕ_k and the noisy ECG measurements s_k may be related to the state vector as follows:

$$\begin{bmatrix} \phi_k \\ s_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \theta_k \\ z_k \end{bmatrix} + \begin{bmatrix} u_k \\ v_k \end{bmatrix} \quad (12)$$

where $R_k = E\{[u_k, v_k]^T [u_k, v_k]\}$ is the observation noise covariance matrix.

In the context of estimation theory, the variance of the observation noise in (12) represents the degree of reliability of a single observation. In other words, when a rather precise measurement of the states of a system is valid the diagonal entries of R_k are small, and the KF gain is adapted so as to rely on that specific measurement. While for the epochs where the measurements are too noisy or there are no measurements available, the R_k entries are large and the KF tends to follow its internal dynamics rather than tracking the observations [21]. Recall that for the phase state variable, θ_k has a periodic value that starts from $\theta = 0$ at the R-peak and ends at $\theta = 2\pi$ with the next R-peak. Although the only valid phase observation is obtained from the R-peak locations, it is possible to linearly assign a phase value between 0 and 2π to the intermediate samples, as illustrated in Fig. 1. The later presented results are all based on this linear phase assignment. However, to indicate the increased uncertainty in the phases assigned to the intermediate samples, the first diagonal entry of R_k corresponding to the time varying variance of the measurement phase noise, may be increased. Another alternative, and rather sophisticated, approach for the estimation of the intermediate phase values is to directly detect the location of the
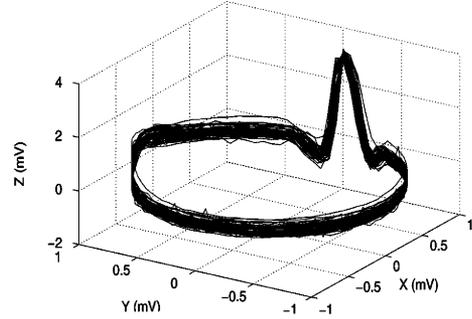
P, Q, S, and T waves from the original signal. However, the previous approach is preferred since R-peak detection is far more reliable in high-noise scenarios. Furthermore, except in the case of abnormal beats, the positions of the P, Q, S, and T waves relative to the R-peak change by only a few milliseconds.

D. Estimation of the Model Parameters

Prior to the implementation of the filter, it is necessary to select the values of the process and measurement noise covariance matrices. Generally, by using m Gaussian kernels in (7), the process noise vector defined in (8) has $3m + 2$ entries (here 17), leading to a $(3m + 2) \times (3m + 2)$ process noise covariance matrix, Q_k. However, if the noise sources are assumed to be uncorrelated with each other, a reasonable approximation adopted here, then the matrix is simplified to be diagonal. The measurement noise covariance matrix R_k is similarly considered to be diagonal.

In order to automate the parameter selection procedure for any given ECG, the parameters should be estimated from the signal itself. For this, as described in the previous subsection, any noisy ECG may be transformed to a three dimensional representation by plotting the noisy ECG versus the periodic phases that are assigned to each sample in polar coordinates on the unit circle ($r = 1$). A typical phase-wrapped ECG with additive noise may be seen in Fig. 2. It is now possible to estimate the dynamic model parameters for the given ECG. For this, the mean and variance of the *phase-wrapped* ECG is calculated for all phases between 0 and 2π. This gives the average of the ECG waveform. A typical signal produced by this approach is depicted in Fig. 3. The error bar in this figure corresponds with the standard deviation (SD) of different ECG cycles around the mean ECG cycle. Next, the problem is to find the optimal parameters of (7) that can best fit the mean ECG cycle. In this stage, many optimization methods may be used. For example as suggested in [31], by using a nonlinear least-squares approach, the best estimate of these parameters in the MMSE sense can be found. A practical means of solving this nonlinear least-squares problem is the *lsqnonlin* function of Matlab® that was used to estimate the initial parameters for the results presented later in this paper.

The next step is to find an estimate for the covariance values of Q_k. This may be done by using the error values as depicted in Fig. 3. In fact, in this step we are attempting to calculate the magnitude of the deviation of the parameters of the five Gaussian functions in (5) around the estimated mean, that best
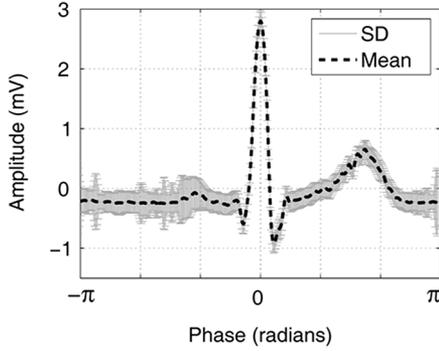
Fig. 3. Average and standard deviation-bar of 30 ECG cycles of a noisy ECG.

model the acceptable deviations of the ECG around the mean ECG $(\overline{\mathrm{ECG}}(\theta))$. This is again a nonlinear least-squares problem that is solved by finding the optimal parameters that generate the best fit of the mean ECG within the upper and lower ranges of $\overline{\mathrm{ECG}}(\theta) + \sigma_{\mathrm{ECG}}(\theta)$ and $\overline{\mathrm{ECG}}(\theta) - \sigma_{\mathrm{ECG}}(\theta)$.

It should be noted that the parameter estimation procedure detailed above is an offline approach that estimates the optimal parameter values for any dataset. It is also possible to develop an online extension of this algorithm that estimates the model and noise parameters from the most recent cycles of the ECG. However, for short ECG recordings we have found that the parameters of the model remain relatively constant and this online process is unnecessary for such signals. Furthermore, as it will be later noted, by monitoring the innovation signal of the KF, it is possible to fine-tune the estimated parameters throughout the filtering process, without the need for their re-estimation. Therefore, for any new dataset, it is possible to start with approximate values for the parameters of the model, which have been calculated from similar data, and to modify these values throughout the filtering process.

The angular frequency ω may be set to $\omega = 2\pi/T$; where T is the RR-interval period in each ECG cycle. For short signals with minor RR-interval deviations, a simpler approximation is to use a global ω using the average RR-interval of the whole signal. It should be noted however, that ω can also be considered to vary on an intra-beat basis too, since the PR- and QT-intervals are known to change with varying autonomic tone, heart rate and to some extent, with each changing RR-interval. The results presented in this paper are based on the more simple approximation above, using a global angular frequency.

The variance of the process noise η should also be estimated. Noting that η is a parameter that represents the imprecision of the dynamic model, neglecting the other physiological sources that influence the ECG, a simple estimate for this parameter would be a zero mean Gaussian random variable with an appropriate variance. An intuitive value for this variance may be found from the deviations of the inactive segment of the ECG, between the end of the T-wave and the beginning of the next P-wave, which correspond to the ending segments of the ECG error-bar of Fig. 3, or the isoelectric segment between the end of the P-wave and the Q point, since no late potentials or baseline wander should manifest during this period.

From (12) we can observe that u_k is the phase measurement noise. As mentioned before, the phase for each beat is determined from the R-peaks of the signal. A possible noise source for u_k is the sampling error that occurs when the actual R-peak is located between two sample times. Another noise source arises from the additive noise spikes that can cause a misdetection of the R-peak location. The first of these may be easily modeled by assuming that the R-peak is uniformly distributed between two consecutive samples. By considering that each ECG cycle is equivalent to 2π in the phase domain, u_k would be uniformly distributed in the range of $\pm\omega\delta/2$, where ω is the angular frequency and δ is the sampling period. With this assumption we have: $E\{u_k^2\} = (\omega\delta)^2/12$. Although there have been rather robust R-peak detectors developed to overcome the misdetection of the R-peaks [32], a precise study of this issue requires the amplitude noise to be related to the phase error (or the so called *phase jitter*), a practice that has been well-studied in other contexts [34]. In this study, for the sake of simplicity, it is assumed that the R-peak detector is reliable and the only phase error is due to the imprecision of the sampling time. Moreover, as mentioned in the previous subsection, for the intermediate phases occurring between consecutive peaks, the variance of the phase noise can be increased to indicate the imprecision of the phase values.

There are also several ways to estimate the variance of the measurement noise, v_k. One method is to estimate the noise power from the deviations of the whole signal around the phase-wrapped ECG, or from the portions of the ECG between two successive T and P waves. A quantitative study of the accuracy of these estimates is presented later. There are also online approaches for noise power estimation, which have been previously suggested for similar applications [3], and apparently the selection of the method depends on the origin of the expected noise.

E. Stability and Convergence Issues

The stability and convergence issues of the KF and its extensions have been well-discussed in the literature. In order to ensure numerical stability of the KF equations, and to prevent the covariance matrices from becoming semi-definite, the *Joseph stabilized form* [35] of (4) are used for the *a posteriori* covariance estimation, to guarantee positive-definite estimates of the covariance matrix.

In practice due to the Gaussian assumption on the noise sources and the initial state vector values, the state estimate entries of $\hat{\mathbf{x}}_k^+$ should lie within the envelope of the square roots of their corresponding diagonal entries in P_k^+ for the majority of the time. Therefore, by monitoring the variance of the filter estimate, it is possible to detect the filter divergence. Moreover it is shown that it is possible to stabilize the KF online, by introducing a forgetting factor in the original filter equations [28].

Another approach known as *sequential measurement incorporation* [35], uses different observations one-by-one to aid the stability of the KF. This technique requires the observation noise covariance matrix to be diagonal. However, for nondiagonal observation noise covariance matrices, decorrelation methods have been developed to diagonalize this matrix.

It is also practically convenient to monitor the covariance matrix of the innovation signal throughout the filtering procedure and to compare it with the innovation covariance matrices estimated by the KF [36], [35]. This provides a means of monitoring the fidelity of the filter and updating the values of Q_k and R_k. Specifically, with a diagonal (or diagonalized) noise covariance matrix of R_k, the following term can be formed for the ith ECG measurement:

$$\gamma_i = \frac{1}{N} \sum_{k=i-N+1}^{i} \frac{(r_k^s)^2}{h_k} \qquad (13)$$

where r_k^s is the second entry of the zero-mean innovation vector of $\mathbf{r}_k$ defined in (4), corresponding with the kth ECG measurement, N is the length of the averaging window, and h_k is the KF estimated variance of r_k^s given by

$$h_k = E\{(r_k^s)^2\} = \mathbf{c}_k^T P_k^- \mathbf{c}_k + \sigma_{v_k}^2 \qquad (14)$$

where $\mathbf{c}_k$ is the second row of the C_k matrix defined in (3), and $\sigma_{v_k}^2 = E\{v_k^2\}$ is the second diagonal entry of R_k. An identical term can be defined for the phase observation ϕ_k.

In fact, γ_i is an average of the variances of the N recent ECG innovations, normalized by their KF estimated variances h_k. Therefore, as long as the KF is performing correctly, $\gamma_i \approx 1$. Values of γ_i much greater than unity indicate that the innovation signal variance is being underestimated by the KF, while values close to zero indicate that the innovation signal variance is being overestimated. Therefore, by monitoring γ_i it is possible to adaptively modify the KF noise parameters (such as Q_k and R_k), to ensure the filter stability and to achieve a better filtering performance. For example, by using the M most recent samples of the innovation signal, $\sigma_{v_k}^2$ can be adaptively modified as follows:

$$\sigma_{v_k}^2 = \lambda \sigma_{v_{k-1}}^2 + (1-\lambda) \frac{1}{M} \sum_{j=k-M}^{k-1} \left(r_j^s\right)^2 \qquad (15)$$

where $0 < \lambda < 1$ is the adaptation coefficient.[1] In [37], a similar means of online modification of the R_k and Q_k entries have been presented.

For the UKF, the algorithm presented in [26] has three parameters, $\tilde{\alpha}, \tilde{\beta}$, and $\tilde{\kappa}$. These control the stability of the filter and enable the algorithm to be fine-tuned for systems with different degrees of nonlinearity and non-Gaussian inputs. The parameter $\tilde{\alpha}$ is an indication of the spread of the state variables around their mean and is selected to be a small positive value in the range of $10^{-4} \leq \tilde{\alpha} \leq 1$. The parameter $\tilde{\beta}$ is used to incorporate prior knowledge about the distribution of the state vector, with $\tilde{\beta} = 2$ being optimal for Gaussian distributions [38]. The parameter $\tilde{\kappa}$ is a secondary *ad hoc* scaling parameter that is selected in accordance with the size of the state vector and the higher order statistics of the noise distributions [38], with $\tilde{\kappa} = 0$ being the optimal selection for a state vector of size two. A mathematical

study of the effect of these parameters on the UKF accuracy compared with the EKF, can be found in [38] and [39].

F. Practical Filtering Schema

Before presenting the experiments and results, the scope of the proposed filtering scheme needs to be further clarified. Following the discussions of this section, by using the Bayesian framework we are attempting to utilize *a priori* information about the underlying dynamics of ECG signals to extract the ECG components from background noise. Hence, compared with conventional filtering schemes that perform rather "blindly," Bayesian filters are naturally expected to give superior results as long as we provide them with valid *a priori* information concerning the signal and noise dynamics. This point becomes important when considering that abnormal ECGs can have high inter-beat variations in their wave timings or morphology, meaning that the underlying dynamics of the signals are not valid in pathological beats.

In the presented approach, due to the phase wrapping of the RR-interval to 2π, normal inter-beat variations of the RR-interval (between 10% to 20%), or consistent RR-interval abnormalities such as *bradycardia* or *tachycardia* do not considerably affect the filter performance. However, for abnormalities that only appear in some of the ECG cycles, the phase error of the model can lead to large errors in the Gaussian function locations. A similar case can occur when the R-peak is misdetected. In particular, for morphological abnormalities that appear in some of the ECG cycles, such as the *premature ventricular contraction* (PVC) [40], the filtering performance is not expected to be satisfactory for low input SNRs, since neither the model nor the measurements are reliable for the filter. For such occasional morphologic changes, even temporal adaptation of the filter parameters is not helpful, as the filter does not have sufficient time to adapt itself. However the benefit of the Gaussian mixture representation is that the effect of each Gaussian term vanishes very quickly (in less than the ECG period), meaning that the errors are not propagated to the following ECG cycles.[2] Moreover by monitoring the state estimates' covariance matrices and the variations of the innovation signals, it is possible to detect such unexpected abnormalities. Of course, it should be considered that the accurate denoising of abnormal ECGs with high morphologic changes remains an open problem even for conventional filtering methods.

Finally, we note that the later presented results have been implemented offline. However, the recursive KF equations are originally designed for online applications and even for the EKS, considering the quasi-periodic nature of the ECG, fixed lag smoothers with only one or two cycles of ECG lag can be used.

V. EXPERIMENTS

A. Dataset

The MIT-BIH Normal Sinus Rhythm Database [41], [42], was used to study the performance of the proposed methods.

---

[1]For $M = 1$ (single-step update), (15) reduces to the autoregressive model suggested in [37], and for $M > 1$ (15) represents a moving average filter with the λ parameter changing the slope of the filter's response. For ECG signals having sharp changes, the moving average model was found to be more robust to the peak changes.

[2]A formal justification of this statement requires mathematical proof of stability and convergence of the EKF, EKS, and UKF for the proposed model, and is beyond the scope of the current paper.

This database was recorded at a sampling rate of 128 Hz from 18 subjects with no significant arrhythmias. From this database 190 low-noise segments of 30 seconds without considerable artifacts were visually selected for the implementation of the proposed filters. These segments were taken from different subjects, recorded from clinical ECG leads V_1, V_2, and V_3. The heart rate of these ECG segments varied from 55 BPM to 90 BPM (74.1 ± 12.0 BPM on average), with RR-interval deviations between 5% to 25% ($16.7 \pm 6.2\%$ on average).

Moreover, to show the filtering performance in the presence of nonstationary noise, real muscle artifact (MA) was taken from the MIT-BIH Noise Stress Test Database [43], [44]. The MA originally had a sampling rate of 360 Hz and therefore they were anti-alias resampled to 128 Hz in order to match the sampling rate of the ECG. Note that the resampling did not considerably change the shape of the MA noise, since the frequency content of the MA noise was concentrated in frequencies below 64 Hz.

B. Noise Generation

Mathematically, white noise is defined to have a flat spectral density function over all frequencies. However, real noise sources have nonflat spectral densities that decrease in power at higher frequencies, making the spectrum *colored* and the noise samples correlated in time. There are different ways of generating colored noise [45], and realistic ECG artifacts [46]. For the current study, we will model the noise color by a single parameter representing the slope of a spectral density function that decreases monotonically with frequency:

$$S(f) \propto \frac{1}{f^\beta} \qquad (16)$$

where f is the frequency and β is a measure of noise color. *White noise* ($\beta = 0$), *pink noise* ($\beta = 1$) or *flicker noise*, and *brown noise* ($\beta = 2$) or the *random walk process*, are three of the most commonly referenced noises. The realization of colored noises with spectral densities described by (16) generally require nonlinear frequency domain filtering of white noise.[3] For random processes, the *expected value* of the squared magnitudes of their frequency transforms, or namely the *periodogram*, is known to be an estimate of the spectral density function of the original samples [47]. Therefore, in order to generate colored noise following (16), samples of white noise can be generated and transferred into the frequency domain using the *Discrete Fourier Transform* (DFT). By altering the frequency components of the DFT according to (16), and transferring the reshaped DFT back to the time-domain, typical samples of colored noise are realized. Note that this approach of frequency domain filtering causes transient behavior in the generated noise time series that should be discarded from the samples.

C. Implementation

Having derived the state equations (7), the observation equations (12), the linearized state equations of the ECG dynamic model (10), (11), and the model parameters, the implementation

[3]Except for special cases such as $\beta = 2$ that can also be achieved through linear time-domain filtering of white noise.

$\theta_i (rads.)$	Gaussian kernel center $\pm 0.05\pi$
$\alpha_i (mV)$	Gaussian kernel peak $\pm 10\%$ of the peak amplitude
$b_i (rads.)$	Gaussian kernel width $\pm 0.05\pi$
$\eta (mV)$	1% of maximum ECG peak
$\omega (rads./sec)$	Mean beat-to-beat angular frequency ($\bar{\omega}$) $\pm$ SD
$u_k (rads.)$	$0.00 \pm (\bar{\omega}\delta)/\sqrt{12}$
$v_k (rads.)$	Ranges over different SNRs

of the EKF, EKS and the UKF are now possible. The procedure for calculating the parameters of the model and the noise covariance matrix entries were explained in Section IV. Using the explained methods, the mean ECG waveform was extracted for each ECG segment and the parameters of the five Gaussian kernels were calculated using the nonlinear least-squares method for each ECG segment. Due to the variety of the ECG leads and the wide range of studied SNR, the covariance matrices were calculated using a simple approach using the peak values of the Gaussian functions. The parameter selection approach is summarized in Table II. These parameters can be further customized for specific ECG leads and special ranges of the input SNR.

In Section IV-E, the control parameters of the UKF algorithm were introduced. Throughout the simulations, these parameters were set to $\tilde{\beta} = 2$ and $\tilde{\kappa} = 0$, which are the optimal selections for a system having two state variables with a Gaussian input noise. The parameter $\tilde{\alpha} = 1$ was also empirically derived as a compromise between performance and stability in different SNRs for the studied database.

D. Benchmark Methods

In order to have a comparison between the performance of the proposed methods and conventional ECG denoising schemes, wavelet denoising (WD), adaptive filtering (AF), and conventional finite impulse response (FIR) filtering were also tested on the database. Implementations of these standard methods are now described.

Conventional WD schemes are characterized by different parameters that allow the algorithms to be customized for different mixtures of signal and noise sources. The type of the *mother wavelet, shrinkage rule, hard-* versus *soft-thresholding, noise level rescaling approach*, and *number of decomposition levels* are among the different parameters of common WD algorithms [48]. There are of course, some general rules concerning the selection of these parameters. For example, the mother wavelet is usually selected from families that somehow resemble the shape of the desired signal, or the rescaling approach and shrinkage rules are selected according to the nature (white versus colored) and variance of the noise. However, the literature on the applications of WD for ECG is rather broad and diverse, making it difficult to judge what may be the best combination of these parameters. Consequently, a rather exhaustive search was carried out on the different combinations of the above mentioned parameters to find the best WD scheme for the database being studied. The different parameters that were tested on this database are listed in Table III. Among the different tested combinations, the *Stein's Unbiased Risk Estimate* (SURE) shrinkage rule, together with a single level rescaling and a soft thresholding strategy always gave superior results. Among the tested mother wavelets,

TABLE III
WD PARAMETER COMBINATIONS TESTED OVER THE DATABASE

Parameter	Values
Mother wavelet	Daubechies 1...8, Coiflets 1...5, Symlets 1...8
Shrinkage rule	SURE, Heuristic SURE, Universal, Minimax
Thresholding strategy	Hard, Soft
Rescaling approach	No scaling, Single level, Multiple level
Decomposition level	1...10

Coiflets2, Coiflets3, Symlets4, and *Symlets5* gave superior results. The best decomposition level ranged from 5 to 8, and showed little significant difference within this range. These results were achieved by the comparison of the SNR curves for input SNR ranging from 35 dB to −5 dB and averaged over 100 Monte Carlo runs with different random noise input vectors. The following reported results are based on the *Coiflets3* mother wavelet with 6 levels of decomposition.

The adaptive filtering approach suggested in [5], was the next filtering approach implemented on the dataset. The original results presented in [5] have been reported on a synthetic ECG, formed by taking a single QRS complex from an ECG and concatenating the same waveform several times to generate a *deterministic* ECG with no timing or morphology variation from beat to beat. An impulse signal, time-synchronized with the R-wave, was also used as the reference channel to enable single channel filtering of the ECG. Although their reported results are rather impressing for this simulated ECG, the method is not expected to give identical results on real ECGs. This is mainly due to the nonstationary behavior of the ECG that causes the ECG shape and RR-duration to change from beat to beat. However this AF scheme was also implemented on the dataset as another benchmark. In the original work [5], the number of the AF weights (L) is selected to be equal to the number of samples of the deterministic QRS complexes. For operating on real ECGs with variable RR-intervals we set L to the maximum sample period between the RR-waves of the input ECG. A convergence rate of $\mu = 0.1$, led to rapid adaptation and stable filter outputs for all the SNR range of this study (−5 dB–30 dB).

The last filtering approach applied to the dataset was a typical FIR filter, consisting of two cascaded high-pass and low-pass filters, with an overall passband of 0.4–40 Hz, a passband ripple of 1 dB, and a stop-band attenuation of about 60 dB. The main frequency components of typical ECGs lie within this frequency range [49], and the selection of a wider or narrower bandwidth would be a compromise between the attenuation of in-band ECG and noise components.

VI. RESULTS AND DISCUSSIONS

In order to investigate the performance of the different methods, artificial white and colored Gaussian noise with different variances were generated and added to the ECG segments, and the noisy signals were presented to the proposed filters. To ensure the consistency of the results, the whole procedure was repeated over the 190 ECG segments; each time using a different set of random noise at the input. The filter output SNR calculation was averaged over the whole 190 results. The SNRs were generally calculated over the second half of the filtered segments, to ensure that the transient effects of the filters would not influence the SNR calculations.
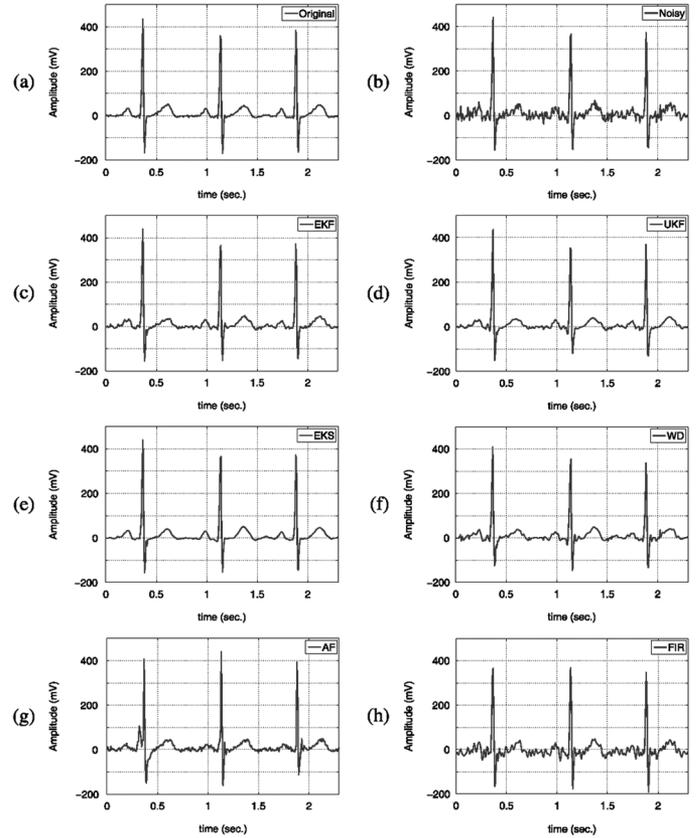

Fig. 4. Typical filtering results for an input signal of 6 dB. (a) Original. (b) Noisy. (c) EKF. (d) UKF. (e) EKS. (f) WD. (g) AF. (h) FIR.

In Fig. 4, typical results of the FIR, AF, WD, EKF, EKS, and the UKF are presented for an input SNR of 6 dB. Visually comparing these results, it can be seen that the proposed methods have admirably tracked the original signal in a rather low input SNR scenario. The EKS demonstrates the smoothest result, while the UKF outperforms the EKF, particularly around the sharp turning points of the signal. In fact, the main difference between the EKF and UKF results are in the QRS complex of the ECG, where the EKF performs slightly less well, since it tends to follow the noisy signal rather than the system dynamics. The reason for this may be seen by studying (2) to (4). According to (3), at low sampling rates and in the rapidly changing regions of the ECG, the approximated matrix A_k has large entries due to the differentiation, that in turn causes an increase in the values of the P_{k+1}^- matrix. This means that the KF tends to rely less on the dynamic model. This assumption was validated by limiting the maximum and minimum values of the derivatives calculated from (3). This change led to results that demonstrate increased flexibility around the QRS complex of the ECG and is reflective of the fact that derivative-free filters such as the UKF are more robust to severe nonlinearities of the input time series. This also suggests that the EKF and EKS can provide a better performance for signals having higher sampling rates.

Among the conventional filtering approaches, the WD outperforms the AF and the FIR; but contains some considerable distortions that do not correspond to the true ECG.
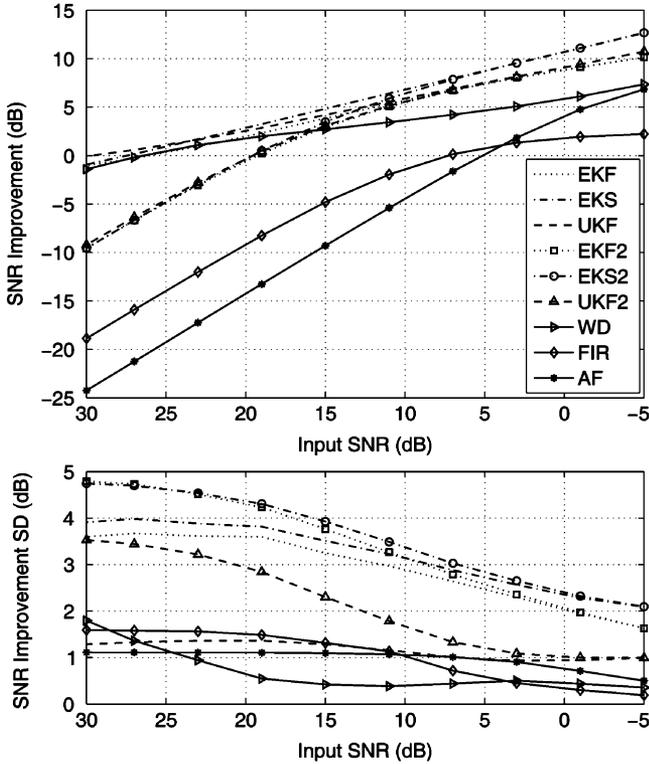
Fig. 5. The mean (top) and Standard Deviation (bottom) of the filter output SNR improvements versus different input SNRs. In these curves EKF, UKF, and EKS correspond to the results without $\sigma_{v_k}^2$ adaptation, and EKF2, UKF2, and EKS2 refer to the ones with $\sigma_{v_k}^2$ adaptation. Refer to text for further details.
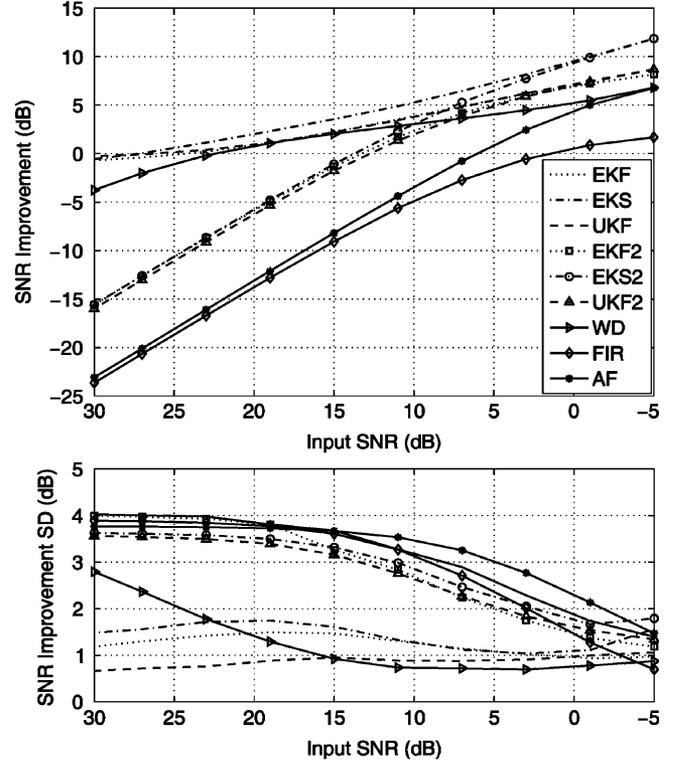

Fig. 6. The mean (top) and Standard Deviation (bottom) of the filter output SNR improvements over the ST-segment of the ECG, versus different input SNRs. In these curves EKF, UKF, and EKS correspond to the results without $\sigma_{v_k}^2$ adaptation, and EKF2, UKF2, and EKS2 refer to the ones with $\sigma_{v_k}^2$ adaptation. Refer to text for further details.

For a quantitative comparison, the mean and SD of the SNR improvements[4] versus different input SNRs, achieved over 190 ECG segments are plotted in Fig. 5. The results of the SNR improvements calculated over the ST-segment, which is extremely sensitive to noise and of great clinical significance, can also be seen in Fig. 6. In the presented results of Figs. 5 and 6, two approaches have been used for the EKF, UKF, and the EKS; the first without the online adaptation of $\sigma_{v_k}^2$, and the second with adaptation. For the first case, $\sigma_{v_k}^2$ was fixed to the *a priori* known variance of the additive noise, and for the second case, this initial value was adaptively modified by the filter. For the latter, the adaptation window length was selected to be $M = 13$ and is approximately equivalent to a 100 ms window for the sampling rate of 128 Hz. This window length is wider than the normal QRS complex at moderate and high heart rates, and therefore should be sufficient to prevent sharp variations in $\sigma_{v_k}^2$. The adaptation coefficient in (15) was set to $\lambda = 0.6$, a value that was empirically found to provide a compromise between adaptation time and stability of $\sigma_{v_k}^2$.

Figs. 5 and 6 illustrate how the results achieved for a constant $\sigma_{v_k}^2$ (i.e., without its online adaptation), are almost linearly related to the input SNR, and are generally better than the results from adaptively changing $\sigma_{v_k}^2$. In fact for input SNRs below 18 dB, the EKF, UKF, and the EKS degrade the input SNR when $\sigma_{v_k}^2$ is adapted; but for SNRs below approximately 10 dB, the results with and without adaptation are asymptotically similar.

[4]The SNR improvement is defined as the output SNR of the filter (in decibels) minus the input SNR (in decibels). Negative SNR improvements indicate a degradation of the input SNR caused by the filtering procedure.

In either case the EKS demonstrates the best average performance, and the UKF performs marginally better than the EKF. Among the conventional filtering methods the WD outperforms both the FIR filter and the AF, both of which are inferior to the other techniques. Furthermore, for input SNRs above 18 dB WD outperforms the proposed methods with $\sigma_{v_k}^2$ adaptation, but still underperforms the constant $\sigma_{v_k}^2$ approach.

The reason for the asymptotic behavior of the results with and without $\sigma_{v_k}^2$ adaptation can be explained by revisiting (14), where we see that the variance of the innovation signal estimated by the KF consists of two parts; the uncertainty of the model parameters (the first term), and the uncertainty of the observations $(\sigma_{v_k}^2)$. In high-input SNR scenarios, the first term dominates the second. Therefore, when adaptively changing $\sigma_{v_k}^2$ in high-input SNRs, we are in fact miscounting the model uncertainties as measurement errors, leading to the overestimation of the measurement noise. Conversely, in low SNR scenarios $\sigma_{v_k}^2$ dominates the first term, and the online adaptation is performed correctly. In different applications with specific ranges of input SNR, this information can help to make online corrections to the estimated noise parameters. These results also suggest that for stationary noise processes, it is preferable to keep the filter noise parameters constant, or alternatively, increase the innovation variance averaging length of M.

From the SD plots in Figs. 5 and 6, it is seen that WD has the least deviation over the different ECG segments, and among the proposed methods the UKF has the least deviation; meaning that the UKF's results are more robust to dataset variations.
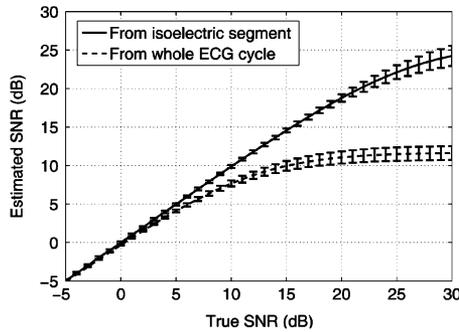
Fig. 7. Estimated signal SNR versus true SNR, from the isoelectric line segment of the ECG between the T-waves and the P-waves (solid), and from the whole ECG cycle (dashed).
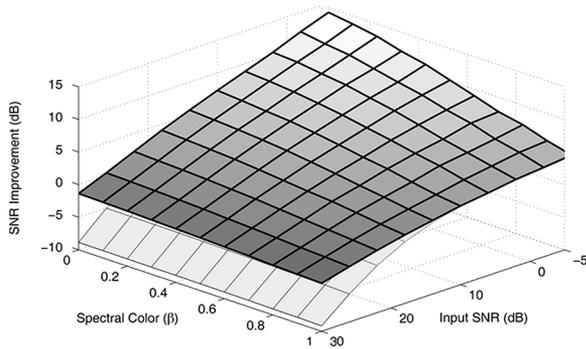

Fig. 8. Effect of noise color on the EKS results. (top) Without $\sigma_{v_k}^2$ adaptation, and (bottom) with $\sigma_{v_k}^2$ adaptation. Both curves merge for input SNRs below 10 dB.

In Section IV-D, it was noted that the initial values of the measurement noise variance may be estimated from the SD of the whole phased-wrapped ECG, or just using the SD of the isoelectric segment of the ECGs between two consecutive T and P waves. To show the accuracy of these estimates, the SNRs estimated from both methods were compared with the true SNR of the signal in presence of additive white noise. The mean and SD of these estimates versus the true input SNR, calculated over the whole database are depicted in Fig. 7. As can be seen from this figure, both methods have underestimated the true SNR (especially in high-input SNRs); but the SNR estimated from the isoelectric segment is very close to the true values for input SNRs below 20 dB. In practice, as we have to estimate the measurement noise variance from the noisy signals, this information may be used as a correction curve for finding accurate estimates of the noise variance.

The result of the noise color study is depicted in Fig. 8 for the SNR improvement of the EKS as a function of the input SNR and the input noise color β. This result was achieved using the average of the 190 ECG segments, both with and without the online adaptation of the measurement noise variance $\sigma_{v_k}^2$. As can be seen in Fig. 8, the EKS performance decreases almost linearly as the noise color ranges from white ($\beta = 0$) to pink ($\beta = 1$), while the slope of decrease is larger for lower input SNRs. As with the previous results, it is seen that the constant $\sigma_{v_k}^2$ results (top curve) outperforms the results with $\sigma_{v_k}^2$ adaptation (bottom curve), and both curves merge in input SNRs below approximately 10 dB.

TABLE IV
SNR RESULTS WITH REAL MUSCLE ARTIFACTS

SNR(dB)	without $\sigma_{v_k}^2$ adaptation			with $\sigma_{v_k}^2$ adaptation		
Input	6.0	12.0	18.0	6.0	12.0	18.0
EKF	10.0	14.1	18.8	14.5	17.8	20.1
UKF	9.5	13.8	18.7	13.8	16.5	18.2
EKS	12.0	15.5	19.5	15.8	18.5	20.1
WD	6.9	12.9	18.9	–	–	–
FIR	5.9	9.7	11.7	–	–	–
AF	5.0	5.4	5.5	–	–	–

The study of the noise color effect was not extended beyond pink noise, since as the noise becomes more colored in spectrum (β increases), the time-domain samples of the noise will have longer-term correlations. This means that for a valid statistical study of the noise effect longer ECG signals are required. Moreover, colored noise as defined in (16) is not generally guaranteed to be a *Wide Sense Stationary* (WSS) process, meaning that the colored noise samples do not have the same variance.[5] When adding such noise to the ECG, each sample receives a different amount of noise, making the overall SNR criteria rather meaningless.

It should be further noted that although the monotonic shape of the SNR surfaces of Fig. 8 proves the consistent behavior of the filtering approaches in different noise colors, the Bayesian framework is originally based on the assumption of white noise sources. In fact, for colored noise with a known spectral behavior, the systematic approach is to use parametric spectral estimation methods to model the colored noise as the output of a system driven by white noise. Subsequently, the state-space model of this system can be augmented with the original system model. In this way, the dynamics of the noise sources are also considered in the filtering procedure.

A more practical example is now given, using real nonstationary MA. For this study normalized values of real MAs, as explained in Section V, were added to the ECG signals. Due to the nonstationarity of the MAs, the variance of the measurement noise varies in time; so using a constant R_k for the whole signal is no longer an optimal choice. Here we can update the R_k matrix by monitoring the variance of the filters' innovation signals using (13) and (15). For this study, the monitoring and adaptation window lengths were respectively selected to be $N = 26$ and $M = 13$, values that are approximately equivalent to 200 ms and 100 ms windows respectively for a sampling rate of 128 Hz. The adaptation coefficient in (15), was again set to $\lambda = 0.6$. The parameter $\sigma_{v_k}^2$ was initially set to a value close to the variance of the beginning segments of the nonstationary noise, and the algorithm was allowed to modify this value according to the variance of the different noisy segments. The results of this study for different input SNRs, are presented in Table IV and compared with the case where $\sigma_{v_k}^2$ is not changed. As it can be seen, the variance adaptation generally improves the filter performance. The noisy input signals and the results of the EKS with and without $\sigma_{v_k}^2$ adaptation are depicted in Fig. 9 for an input SNR of 6.0 dB. The calculated γ_i terms with and without the variance adaptation are also depicted in the last two panels. As can be seen in this figure, when the measurement noise variance is kept constant, as the noise power increases at

[5]Specifically for the *random walk process* ($\beta = 2$), the variance of the noise samples increases linearly with time [47].
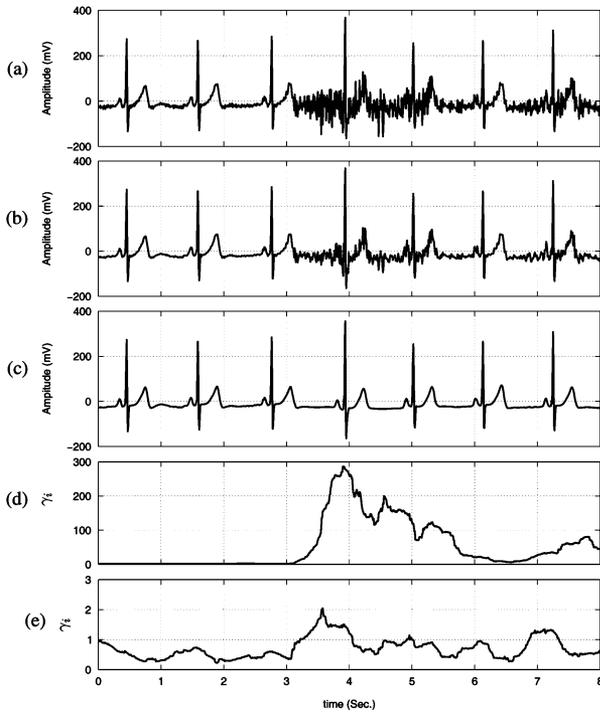
Fig. 9. Muscle artifact removal. (a) Noisy ECG with $\mathrm{SNR} = 6$ dB. (b) EKS result without measurement noise variance adaptation. (c) EKS result using measurement noise variance adaptation. (d) γ_i without adaptation. (e) γ_i with adaptation.

$t = 3$s, γ_i increases, indicating that the measurement noise is being underestimated. By allowing the measurement noise to be adaptively updated, γ_i remains around unity. It should be noted that in practice, $\sigma_{v_k}^2$ should not be manipulated based on very short windows of innovation signal estimates, and as long as γ_i is in the range of 0.5–2.0, the modification of $\sigma_{v_k}^2$ is not required.

VII. CONCLUSION

In this paper, a mathematical framework was proposed for the model-based Bayesian filtering of single channel noisy ECG recordings. Within this framework several sub-optimal filtering schemes were developed and the results were compared with conventional filtering methods. The results demonstrate that the proposed approach can serve as a novel framework for achieving high-resolution ECG, which is the state of the art in applications such as the detection of late potential ECG signals [23], or the noninvasive extraction of fetal cardiac signals from abdominal sensors [24].

The filtering schemes were based on a modified version of a previously proposed dynamic ECG model. However, the generality and modularity of the proposed methods allow for the improvement of the selected dynamic model. As we have shown, the architecture of the Bayesian filtering methods is such that the filters can work with a rather coarse dynamic model as long as the covariance matrices of the system noise vectors are well estimated. Nevertheless, it is also possible to improve the results by using more than five Gaussian functions to be able to model a broader range of normal and abnormal ECGs. Moreover, the derivative-free characteristics of the UKF also removes the need for an analytical form of the system dynamic model and the statistics of the signals are directly estimated from a finite

number of samples, rather than the linearized models required in the EKF or the EKS. This suggests that other morphological models of the ECG can be used instead of mathematical dynamic models. In other words, the nonlinear dynamical model of the ECG may be replaced with any synthetic ECG generator that is capable of generating normal or abnormal ECGs.

The presented results were based on the assumption of additive Gaussian noise sources. With recent developments in Bayesian filtering approaches such as the *Particle Filter* (PF), other noise distributions may also be considered [29].

Another extension of the proposed method would be to use the recently developed three-dimensional models of the *vector-cardiogram* (VCG) [46], to develop a multichannel extension of the hereby proposed methods. In fact, every ECG channel could be assumed as the projection of the VCG in the direction of the recording electrodes. Hence, multiple recordings from different electrodes could be used as additional observations for the filtering procedure and therefore further improve the performance of the filters.

As discussed in Section IV-F, in future works the problems related with abnormal ECGs should also be addressed in the filtering models. The extension of these methods to many of the common ECG abnormalities is rather straightforward, since the model parameters may be simply recalculated and used in the filter model. However, for some heart defects such as the PVC, where the abnormal wave only appears in certain cycles of the ECG, some revisions are necessary in the filtering process to be able to simultaneously filter the normal and abnormal segments of the ECG. An intuitive approach to this problem would be to pre-calculate the filter model parameters for different abnormalities and switch between these values by predicting the existence of an abnormal beat, using a standard ECG classifier or directly using the derived parameters of the model to make a classification itself, as suggested in [31].

Another related issue is the study of the appropriateness of the filtering procedure. In fact, while the Bayesian filtering framework can serve as a powerful tool for the removal of noise from the ECG, one should always be aware of the over-filtering of the signals that can lead to the removal of clinically important information from the ECG. However, results of this study suggest that by selecting a flexible dynamic model for the ECG and an adaptive update of the model parameters, it should be possible to prevent such issues. In future works, problems concerning the convergence-time, stability, estimation bias, and preciseness of the filter results can also be studied within a mathematical framework and by using the *Cramer Rao* lower estimation error bounds [25].

While the results of this paper were achieved with artificially generated noise, without any specific assumption on the noise origin, the generality of the approach allows the model to be customized for specific applications. In fact, as it was mentioned in previous sections, for situations where the nature of the contaminating noises are more well known, the noise model may be selected according to the particular situation. For example, any spectral information about the noise sources, such as the baseline wander caused by the respiratory system, may be transferred into state-space form by using conventional spectral factorization methods, and be augmented with the dynamic model of the system.

Although not addressed in this paper, the proposed methods could also be linked with *blind source separation* (BSS) techniques. Even though the idea behind conventional BSS techniques are somewhat different to KFs, it has been shown in previous works that it is possible to combine state-space models with BSS, to simultaneously benefit from the dynamical filtering abilities of KFs and the spatial filtering abilities of BSS [50]. In other words, the proposed framework together with recent BSS developments may be used to construct a joint spatio-temporal dynamic filter for noisy ECG recordings. This framework will possibly be of interest in applications such as fetal ECG extraction, where conventional filtering schemes and BSS have (separately) failed to extract the complete fetal ECG waveform due to the spatial, temporal, and spectral overlap of the fetal ECG with contaminating noises (such as the maternal ECG). Specifically, if an efficient three-dimensional extension of the proposed methods is constructed, it may be further possible to identify the maternal ECG signal that is projected onto the abdominal electrodes, and to remove it from the fetal recordings.

REFERENCES

[1] I. I. Christov and I. K. Daskalov, "Filtering of electromyogram artefacts from the electrocardiogram," *Med. Eng. Phys.*, vol. 21, pp. 731–736, 1999.

[2] A. Gotchev, N. Nikolaev, and K. Egiazarian, "Improving the transform domain ECG denoising performance by applying interbeat and intrabeat decorrelating transforms," in *Proc. 2001 IEEE Int. Symp. Circuits Syst. (ISCAS)*, Sydney, Australia, 2001, pp. 17–20.

[3] P. Lander and E. J. Berbari, "Time-frequency plane wiener filtering of the high-resolution ECG: Development and application," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 4, pp. 256–265, Apr. 1997.

[4] N. V. Thakor and Y. S. Zhu, "Applications of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 8, pp. 785–794, Aug. 1991.

[5] P. Laguna, R. Jane, O. Meste, P. W. Poon, P. Caminal, H. Rix, and N. V. Thakor, "Adaptive filter for event-related bioelectric signals using an impulse correlated reference input: Comparison with signal averaging techniques," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 10, pp. 1032–1044, Oct. 1992.

[6] P. Lander and E. J. Berbari, "Time-frequency plane wiener filtering of the high-resolution ECG: Background and time-frequency representations," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 4, pp. 247–255, Apr. 1997.

[7] H. A. Kestler, M. Haschka, W. Kratz, F. Schwenker, G. Palm, V. Hombach, and M. Höher, "Denoising of high-resolution ecg-signals by combining the discrete wavelet transform with the wiener filter," in *Proc. IEEE Conf. Comput. Cardiology*, 1998, pp. 233–236.

[8] N. Nikolaev and A. Gotchev, "ECG signal denoising using wavelet domain wiener filtering," in *Proc. Eur. Signal Process. Conf. EUSIPCO-2000*, Tampere, Finland, Sep. 2000, pp. 51–54.

[9] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[10] M. Popescu, P. Cristea, and A. Bezerianos, "High resolution ECG filtering using adaptive Bayesian wavelet shrinkage," in *Proc. Comput. Cardiology 1998*, Cleveland, OH, Sep. 13–16, 1998, pp. 401–404.

[11] P. M. Agante and J. P. M. de Sá, "ECG noise filtering using wavelets with soft-thresholding methods," in *Proc. Comput. Cardiology 1999*, 1999, pp. 535–542.

[12] G. Moody and R. Mark, "QRS morphology representation and noise estimation using the Karhunen-Loève transform," *Comput. Cardiology*, vol. 16, pp. 269–272, 1989.

[13] A. Barros, A. Mansour, and N. Ohnishi, "Removing artifacts from ECG signals using independent components analysis," *Neurocomputing*, vol. 22, pp. 173–186, 1998.

[14] T. He, G. D. Clifford, and L. Tarassenko, "Application of ICA in removing artefacts from the ECG," *Neural Comput. Applic.*, vol. 15, no. 2, pp. 105–116, 2006.

[15] G. D. Clifford and L. Tarassenko, "One-pass training of optimal architecture auto-associative neural network for detecting ectopic beats," *Electron. Lett.*, vol. 37, no. 18, pp. 1126–1127, Aug. 2001.

[16] T. Schreiber and D. T. Kaplan, "Nonlinear noise reduction for electrocardiograms," *Chaos Interdiciplinary J. Nonlinear Sci.*, vol. 6, no. 1, pp. 87–92, Mar. 1996.

[17] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamic model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, Mar. 2003.

[18] P. E. McSharry and G. D. Clifford, ECGSYN—A realistic ECG waveform generator. PhysioNet, Cambridge, MA [Online]. Available: http://www.physionet.org/physiotools/ecgsyn/

[19] G. D. Clifford and P. E. McSharry, "A realistic coupled nonlinear artificial ECG, BP, and respiratory signal generator for assessing noise performance of biomedical signal processing algorithms," in *Proc. SPIE Int. Symp. Fluctuations Noise*, 2004, vol. 5467, no. 34, pp. 290–301.

[20] G. D. Clifford, "A novel framework for signal representation and source separation," *J. Biol. Syst.*, vol. 14, no. 2, pp. 169–183, Jun. 2006.

[21] R. Sameni, M. B. Shamsollahi, and C. Jutten, "Filtering electrocardiogram signals using the extended Kalman filter," in *Proc. 27th Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc. (EMBS)*, Shanghai, China, Sep. 1–4, 2005, pp. 5639–5642.

[22] R. Sameni, M. B. Shamsollahi, C. Jutten, and M. Babaie-Zadeh, "Filtering noisy ECG signals using the extended Kalman filter based on a modified dynamic ECG model," in *Proc. 32nd Annu. Int. Conf. Comput. Cardiology*, Lyon, France, Sep. 25–28, 2005, pp. 1017–1020.

[23] J. Bronzino, Ed., *The Biomedical Engineering Handbook*, 2nd ed. Boca Raton, FL: CRC Press, 2000.

[24] F. Vrins, V. Vigneron, C. Jutten, and M. Verleysen, "Abdominal electrodes analysis by statistical processing for fetal electrocardiogram extraction," in *Proc. 2nd IASTED Int. Conf. Biomed. Eng. (BioMed 2004)*, Innsbruck, Austria, Feb. 16–18, 2004, pp. 244–249.

[25] S. M. Kay, Ed., *Fundamentals of Statistical Signal Processing: Estimation Theory.*. Englewood Cliffs, NJ: Prentice Hall PTR, 1993.

[26] S. Haykin, Ed., *Kalman Filtering and Neural Networks*. New York: Wiley, 2001.

[27] M. Nørgaard, The Kalman Filter Toolbox. [Online]. Available: http://www.iau.dtu.dk/research/control/kalmtool.html

[28] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.

[29] Z. Chen, "Bayesian Filtering: From Kalman filters to particle filters, and beyond," Adaptive Systems Lab., McMaster Univ., Hamilton, ON, Canada [Online]. Available: http://soma.crl.mcmaster.ca/zhechen/download/ieee_bayesian.ps

[30] E. Wan *et al.*, ReBEL: Recursive Bayesian Estimation Library. [Online]. Available: http://choosh.ece.ogi.edu/rebel/

[31] G. D. Clifford, A. Shoeb, P. E. McSharry, and B. A. Janz, "Model-based filtering, compression and classification of the ECG," *Int. J. Bioelectromagnetism*, vol. 7, no. 1, pp. 158–161, 2005.

[32] B.-U. Köhler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *IEEE Eng. Med. Biol. Mag.*, vol. 21, no. 1, pp. 42–57, Jan./Feb. 2002.

[33] S. Olmos, L. Sörnmo, and P. Laguna, "Block adaptive filters with deterministic reference inputs for event related signals; BLMS and BRLS," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1102–1112, May 2002.

[34] Y. Surrel, "Additive noise effect in digital phase detection," *Appl. Opt.*, vol. 36, pp. 271–276, 1997.

[35] M. S. Grewal, L. R. Weill, and A. P. Andrews, Eds., *Global Positioning Systems, Inertial Navigation, and Integration*. New York: Wiley, 2001.

[36] M. A. Maasoumnia, "Estimation Theory and Optimal Filtering, Lecture Notes," Sharif Univ. of Technol., Tehran, Iran, 2003.

[37] M. P. Tarvainen, S. D. Georgiadis, P. O. Ranta-aho, and P. A. Karjalainen, "Time-varying analysis of heart rate variability signals with a Kalman smoother algorithm," *Physiol. Meas.*, vol. 27, pp. 225–239, Mar. 2006.

[38] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proc. 1995 Amer. Control Conf. (ACC'95)*, Seattle, WA, Jun. 21–23, 1995, vol. 3, pp. 1628–1632.

[39] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Contr.*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

[40] J. A. Malmivuo and R. Plonsey, Eds., *Bioelectromagnetism, Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford, U.K.: Oxford Univ. Press, 1995 [Online]. Available: http://butler.cc.tut.fi/malmivuo/bem/bembook

[41] 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, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation* vol. 101, no. 23, pp. e215–e220, Jun. 13, 2000 [Online]. Available: http://circ.ahajournals.org/cgi/content/full/101/23/e215

[42] The MIT-BIH Normal Sinus Rhythm Database. PhysioNet, Cambridge, MA [Online]. Available: http://www.physionet.org/physiobank/database/nsrdb/

[43] G. B. Moody, W. Muldrow, and R. G. Mark, "A noise stress test for arrhythmia detectors," in *Proc. Comput. Cardiology*, 1984, pp. 381–384.

[44] The MIT-BIH Noise Stress Test Database. PhysioNet, Cambridge, MA [Online]. Available: http://www.physionet.org/physiobank/database/nstdb/

[45] S. M. Kay, "Efficient generation of colored noise," *Proc. IEEE*, vol. 69, no. 4, pp. 480–481, Apr. 1981.

[46] R. Sameni, G. D. Clifford, C. Jutten, and M. B. Shamsollahi, "Multichannel ECG and noise modeling: Application to maternal and fetal ECG signals," *EURASIP J. Adv. Signal Process.* 2007, Article ID 43407. 14 pp. [Online]. Available: http://www.hindawi.com/GetArticle.aspx?doi=10.1155/2007/43407.

[47] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[48] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, Matlab® Wavelet Toolbox User's Guide Version 3..

[49] J. J. Bailey, A. S. Berson, and A. J. Garson, "Recommendations for standardization and specifications in automated electrocardiography: Bandwidth and digital signal processing.," *Circulation*, vol. 81, pp. 730–739, 1990.

[50] A. Cichocki and S. Amari, Eds., *Adaptive Blind Signal and Image Processing*. New York: Wiley, 2003.

Reza Sameni (S'05) was born in Shiraz, Iran, in 1977. He received a B.Sc. in electronics engineering from Shiraz University, Iran, and an M.Sc. in bioelectrical engineering from Sharif University of Technology, Iran, in 2000 and 2003, respectively. He is currently a joint Ph.D. student of electrical engineering in Sharif University of Technology and the Institut National Polytechnique de Grenoble (INPG), France. His research interests include statistical signal processing and time-frequency analysis of biomedical recordings, and he is working on the modeling, filtering, and analysis of fetal cardiac signals in his Ph.D. thesis. He has also worked in industry on the design and implementation of digital electronics and software defined radio systems.

Mohammad B. Shamsollahi (M'02) was born in Qom, Iran, in 1965. He received the B.Sc. degree in electrical engineering from Tehran University, Tehran, Iran, in 1988, and the M.Sc. degree in electrical engineering, Telecommunications, from Sharif University of Technology, Tehran, Iran, in 1991. He received the Ph.D. degree in electrical engineering, Biomedical Signal Processing, from the University of Rennes 1, Rennes, France, in 1997. Currently, he is an Associate Professor with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran. His research interests include biomedical signal processing, brain computer interface, time-scale and time-frequency signal processing.

Christian Jutten (A'92–M'03–SM'06) received the Ph.D. degree in 1981 and the Docteur és Sciences degree in 1987 from the Institut National Polytechnique of Grenoble (France). He taught as associate professor in the Electrical Engineering Department from 1982 to 1989. He was a visiting professor in the Swiss Federal Polytechnic Institute in Lausanne in 1989, before becoming a full professor in University Joseph Fourier of Grenoble, more precisely in the sciences and technologies department: Polytech'Grenoble. He is currently the associate director of the Grenoble images, speech, signal and control laboratory (GIPSA, 300 peoples) and head of the Department Images-Signal (DIS) of this laboratory. For 25 years, his research interests are blind source separation, independent component analysis and learning in neural networks, including theoretical aspects (separability, source separation in nonlinear mixtures), applications in signal processing (biomedical, seismic, speech) and data analysis. He is author or co-author of more than 40 papers in international journals, 16 invited papers and 100 communications in international conferences. He has been associate editor of IEEE Trans. on Circuits and Systems (1994–95), and co-organizer of the 1st International Conference on Blind Signal Separation and Independent Component Analysis (Aussois, France, January 1999). He is a reviewer of many international journals (IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, IEEE TRANSACTIONS ON NEURAL NETWORKS, *Signal Processing, Neural Computation, Neurocomputing,* etc.) and conferences in signal processing and neural networks (ICASSP, ISCASS, EUSIPCO, IJCNN, ICA, ESANN, IWANN, etc.).

Gari D. Clifford (S'01–M'03–SM'06) received a B.Sc. in physics and electronics from Exeter University, U.K., an M.Sc. in mathematics and theoretical physics from Southampton University, U.K, and a Ph.D. in neural networks and biomedical engineering from Oxford University, U.K., in 1992, 1995, and 2003, respectively. He has worked in industry on the design and production of several C∈and FDA-approved medical devices. Dr. Clifford is currently a research scientist in the Harvard-MIT Division of Health Sciences where he is the engineering manager of an R01 NIH-funded research program, "Integrating Data, Models, and Reasoning in Critical Care," and a major contributor to the well-known PhysioNet Research Resource. He has taught at Oxford, MIT, and Harvard and is currently an instructor in biomedical engineering at MIT. Dr. Clifford, a senior member of the IEEE, has authored and coauthored more than 40 publications in the field of biomedical engineering, including a recent book on ECG analysis. Dr. Clifford is on the editorial boards of BioMedical Engineering OnLine and the Journal of Biological Systems. His research interests include multidimensional biomedical signal processing, linear and nonlinear time series analysis, relational database mining, decision support, and mathematical modeling of the ECG and the cardiovascular system.