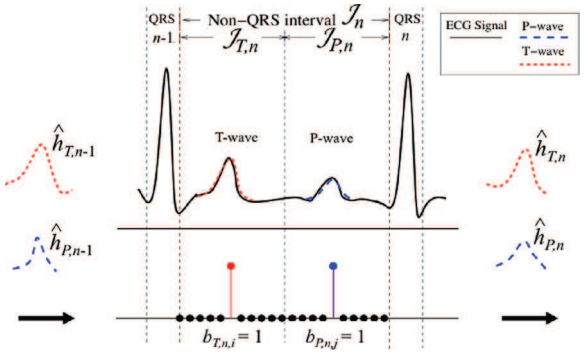

Fig. 1. Signal model for the beat-to-beat processing scheme.

and they potentially contain P and T waves. In this paper, we assume that the locations of the non-QRS intervals have been determined by a preliminary QRS detection step using, e.g., the Pan–Tompkins algorithm [15], and that baseline wanderings have been removed by, e.g., the median filtering technique proposed in [16]. As shown in Fig. 1, the non-QRS interval $\mathcal{J}_n$ associated with the nth beat consists of two complementary subintervals: a T search interval $\mathcal{J}_{T,n}$, which may contain a T wave, and a P search interval $\mathcal{J}_{P,n}$, which may contain a P wave. The temporal lengths of the intervals $\mathcal{J}_n$, $\mathcal{J}_{T,n}$, and $\mathcal{J}_{P,n}$ will be denoted by N_n, $N_{T,n}$, and $N_{P,n}$, respectively. Note that $N_{T,n}+N_{P,n}=N_n$. The lengths $N_{T,n}$ and $N_{P,n}$ can be determined by a cardiologist or simply as fixed percentages of N_n. In this work, we choose $N_{T,n}=N_{P,n}=N_n/2$ for simplicity. Our goal is to estimate the locations, amplitudes, and shapes of the P and T waves within their respective search intervals $\mathcal{J}_{T,n}$ and $\mathcal{J}_{P,n}$. Note that only the locations of the wave *peaks* are constrained to lie within their respective search intervals.

2.1.1. Convolution model

The baseline-free signal in the non-QRS interval $\mathcal{J}_n$ can be approximated by two pulses representing the P and T waves (see Fig. 1). Similar to the blind deconvolution problem in [17,18], the T wave is modeled by the convolution of an unknown binary "indicator sequence" $\boldsymbol{b}_{T,n}=(b_{T,n,1}\dots b_{T,n,N_{T,n}})^T$ indicating the wave locations ($b_{T,n,k}=1$ if there is a wave at the kth possible location, $b_{T,n,k}=0$ otherwise) with an unknown T waveform $\boldsymbol{h}_{T,n}=(h_{T,n,-L}\dots h_{T,n,L})^T$. Analogous definitions for the P wave yield $\boldsymbol{b}_{P,n}=(b_{P,n,1}\dots b_{P,n,N_{P,n}})^T$ and $\boldsymbol{h}_{P,n}=(h_{P,n,-L}\cdots h_{P,n,L})^T$. Here, the waveform length $2L+1$ is chosen as a fixed percentage of N_n that is large enough to accommodate the actual supports of the P and T waves. Within each indicator vector $\boldsymbol{b}_{T,n}$ and $\boldsymbol{b}_{P,n}$, at most one entry is nonzero because at most one wave may occur in any given search interval. According to this model, the nth non-QRS signal component can be expressed as follows:

$$x_{n,k}=\sum_{j=1}^{N_{T,n}} h_{T,n,k-j}b_{T,n,j}+\sum_{j=N_{T,n}+1}^{N_n} h_{P,n,k-j}b_{P,n,j-N_{T,n}}+e_{n,k} \qquad (1)$$

with $k\in\mathcal{J}_n=\{1,\dots,N_n\}$. Here, $e_{n,k}$ denotes white Gaussian noise with unknown variance $\sigma^2_{e,n}$. Furthermore, we have set $h_{T,n,k}=h_{P,n,k}=0$ for $k\notin\{-L,\dots,L\}$.

2.1.2. Waveform expansion

Following [19,20], we represent the P and T waveforms by a basis expansion using discrete-time versions of Hermite functions. Thus, the waveform vectors can be written as

$$\boldsymbol{h}_{T,n}=\boldsymbol{H}\boldsymbol{\alpha}_{T,n},\quad \boldsymbol{h}_{P,n}=\boldsymbol{H}\boldsymbol{\alpha}_{P,n} \qquad (2)$$

where $\boldsymbol{H}$ is a $(2L+1)\times G$ matrix whose columns are the first G Hermite functions (with $G\le 2L+1$), suitably sampled and truncated to length $2L+1$, and $\boldsymbol{\alpha}_{T,n}$ and $\boldsymbol{\alpha}_{P,n}$ are unknown coefficient vectors of length G. By using these expansions, the number of unknown parameters can be significantly reduced (from $2L+1$ to G for each waveform). More specifically, the ECG signals involved in our study were sampled with a sampling frequency of 250 Hz. Considering a heart rate of around 60 beats per minute, that makes on average 250 samples for each beat. We used 20 Hermite coefficients for each P and T wave plus two wave location parameters and one noise variation parameter. Thus, the ratio between the number of parameters to be estimated and the available data (used for the estimation) is approximately 0.2. Note that the amplitudes of the P and T waves are absorbed into the coefficient vectors $\boldsymbol{\alpha}_{T,n}$ and $\boldsymbol{\alpha}_{P,n}$. This is a difference from the model in [10,11], where the amplitudes were defined for each beat individually whereas the P and T waveforms were fixed for multiple beats.

2.1.3. Vector formulation

Using (2), we obtain the following vector representation of the non-QRS signal in (1):

$$\boldsymbol{x}_n=\boldsymbol{B}_{T,n}\boldsymbol{H}\boldsymbol{\alpha}_{T,n}+\boldsymbol{B}_{P,n}\boldsymbol{H}\boldsymbol{\alpha}_{P,n}+\boldsymbol{e}_n \qquad (3)$$

where $\boldsymbol{x}_n=(x_{n,1}\dots x_{n,N_n})^T$, $\boldsymbol{B}_{T,n}$ is the $N_n\times(2L+1)$ Toeplitz matrix with first row $(b_{T,n,L+1}\dots b_{T,n,1}\ 0\dots 0)$ and first column $(b_{T,n,L+1}\dots b_{T,n,N_{T,n}}\ 0\dots 0)^T$, $\boldsymbol{B}_{P,n}$ is the $N_n\times(2L+1)$ Toeplitz matrix with last row $(0\dots 0\ b_{P,n,N_{P,n}}\dots b_{P,n,N_{P,n}-L})$ and last column $(0\dots 0\ b_{P,n,1}\dots b_{P,n,N_{P,n}-L})^T$, and $\boldsymbol{e}_n=(e_{n,1}\dots e_{n,N_n})^T$ is a Gaussian vector with zero mean and covariance matrix $\sigma^2_{e,n}\mathbf{I}_{N_n}$, with $\mathbf{I}_{N_n}$ denoting the identity matrix of size $N_n\times N_n$.

2.2. Likelihood function, prior, and posterior

According to the parametrization introduced in Section 2.1, the unknown parameters for the nth non-QRS interval $\mathcal{J}_n$ are given by the random vector $\boldsymbol{\theta}_n\triangleq(\boldsymbol{b}^T_{T,n}\,\boldsymbol{b}^T_{P,n}\,\boldsymbol{\alpha}^T_{T,n}\,\boldsymbol{\alpha}^T_{P,n}\,\sigma^2_{e,n})^T$. Note, in particular, that the noise variance $\sigma^2_{e,n}$ may vary from one beat to another. Bayesian detection/estimation relies on the posterior distribution, $p(\boldsymbol{\theta}_n|\boldsymbol{x}_n)\propto p(\boldsymbol{x}_n|\boldsymbol{\theta}_n)p(\boldsymbol{\theta}_n)$, where $\propto$ means "equal up to a positive factor that does not depend on $\boldsymbol{\theta}_n$," $p(\boldsymbol{x}_n|\boldsymbol{\theta}_n)$ is the likelihood function, and $p(\boldsymbol{\theta}_n)$ is the prior distribution of $\boldsymbol{\theta}_n$. The next two subsections present the likelihood function and priors considered in this study.

2.2.1. Likelihood function

Using (3) and the fact that $e_{n,k}$ is white and Gaussian with variance $\sigma^2_{e,n}$, the likelihood function (viewed as

a function of $\boldsymbol{x}_n$) is obtained as

$$p(\boldsymbol{x}_n|\boldsymbol{\theta}_n) = \mathcal{N}(\boldsymbol{B}_{\mathrm{T},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n} + \boldsymbol{B}_{\mathrm{P},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P},n},\ \sigma_{e,n}^2\mathbf{I}_{N_n}) \tag{4}$$

where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$ denotes the multivariate Gaussian probability density function with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}$.

2.2.2. Prior distributions

Wave indicators: The indicators $b_{\mathrm{T},n,k}$ are subject to a *block constraint*: within $\mathcal{J}_{\mathrm{T},n}$, there is one T wave (thus[1], $\|\boldsymbol{b}_{\mathrm{T},n}\| = 1$) or none (thus, $\|\boldsymbol{b}_{\mathrm{T},n}\| = 0$), the latter case being very unlikely. Therefore, we define the prior of $\boldsymbol{b}_{\mathrm{T},n}$ as

$$p(\boldsymbol{b}_{\mathrm{T},n}) = \begin{cases} p_0 & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 0 \\ p_1 & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $p_1 = (1-p_0)/N_{\mathrm{T},n}$ and p_0 is chosen very small. Similarly, within $\mathcal{J}_{\mathrm{P},n}$, there is one P wave or none; therefore, the prior of $\boldsymbol{b}_{\mathrm{P},n}$ is defined as in (5), with $p_1 = (1-p_0)/N_{\mathrm{P},n}$. The wave indicator vectors $\boldsymbol{b}_{\mathrm{T},n}$ and $\boldsymbol{b}_{\mathrm{P},n}$ for different search intervals (i.e., different values of n) are assumed to be statistically independent.

Waveform coefficients: The waveform coefficient vectors $\boldsymbol{\alpha}_{\mathrm{T},n}$ and $\boldsymbol{\alpha}_{\mathrm{P},n}$ for the nth non-QRS interval $\mathcal{J}_n$ are supposed to depend on the respective coefficient vectors in the $(n-1)$th non-QRS interval $\mathcal{J}_{n-1}$. Consider the T wave as an example. The prior of $\boldsymbol{\alpha}_{\mathrm{T},n}$ is defined as

$$p(\boldsymbol{\alpha}_{\mathrm{T},n}|\boldsymbol{b}_{\mathrm{T},n}, \boldsymbol{\alpha}_{\mathrm{T},n-1}) = \begin{cases} \delta(\boldsymbol{\alpha}_{\mathrm{T},n} - \boldsymbol{\alpha}_{\mathrm{T},n-1}) & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 0 \\ \mathcal{N}(\boldsymbol{\alpha}_{\mathrm{T},n-1}, \sigma_\alpha^2\mathbf{I}_G) & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 1 \end{cases} \tag{6}$$

where $\delta(\cdot)$ is the Dirac delta function. For the variance σ_α^2, we choose a value that yields a reasonable variability of the waveform coefficients from one interval to another. Note that when there is no T wave in the search interval ($\|\boldsymbol{b}_{\mathrm{T},n}\| = 0$), the prior sets $\boldsymbol{\alpha}_{\mathrm{T},n}$ equal to $\boldsymbol{\alpha}_{\mathrm{T},n-1}$, i.e., the waveform coefficients are equal to those in the previous interval $\mathcal{J}_{\mathrm{T},n-1}$. The prior of the P waveform coefficient vector $\boldsymbol{\alpha}_{\mathrm{P},n}$ is defined in an analogous way, with $\boldsymbol{\alpha}_{\mathrm{T},n-1}$ replaced by $\boldsymbol{\alpha}_{\mathrm{P},n-1}$. These definitions of the priors of $\boldsymbol{\alpha}_{\mathrm{T},n}$ and $\boldsymbol{\alpha}_{\mathrm{P},n}$ introduce a memory in the statistical model for the P and T waveforms and, in turn, induce a sequential processing.

Noise variances: The noise variances $\sigma_{e,n}^2$ are modeled as independent random variables distributed according to an inverse gamma distribution $p(\sigma_{e,n}^2) = \mathcal{IG}(\xi, \eta)$, where ξ and η are fixed hyperparameters defining a vague prior (as in [21]).

We note at this point that the Gaussian priors of $\boldsymbol{\alpha}_{\mathrm{T},n}$ and $\boldsymbol{\alpha}_{\mathrm{P},n}$ are *conjugate priors* with respect to the Gaussian likelihood function (4), i.e., the resulting full conditional distributions (required in the Gibbs sampler) are also Gaussian [22, p. 97]. A similar remark applies to the inverse gamma prior of $\sigma_{e,n}^2$. The choice of conjugate priors yields a considerable simplification of our detection/estimation algorithm.

Joint prior: Since there are no known relations between $(\boldsymbol{b}_{\mathrm{T},n}, \boldsymbol{\alpha}_{\mathrm{T},n})$, $(\boldsymbol{b}_{\mathrm{P},n}, \boldsymbol{\alpha}_{\mathrm{P},n})$, and $\sigma_{e,n}^2$, all these sets of parameters are assumed to be *a priori* statistically independent. Therefore, the joint prior for the total parameter vector $\boldsymbol{\theta}_n = (\boldsymbol{b}_{\mathrm{T},n}^T\ \boldsymbol{b}_{\mathrm{P},n}^T\ \boldsymbol{\alpha}_{\mathrm{T},n}^T\ \boldsymbol{\alpha}_{\mathrm{P},n}^T\ \sigma_{e,n}^2)^T$ factors as

$$\begin{aligned} p(\boldsymbol{\theta}_n|\boldsymbol{\alpha}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P},n-1}) &= p(\boldsymbol{\alpha}_{\mathrm{T},n}|\boldsymbol{b}_{\mathrm{T},n}, \boldsymbol{\alpha}_{\mathrm{T},n-1})\ p(\boldsymbol{b}_{\mathrm{T},n}) \\ &\quad \times p(\boldsymbol{\alpha}_{\mathrm{P},n}|\boldsymbol{b}_{\mathrm{P},n}, \boldsymbol{\alpha}_{\mathrm{P},n-1})\ p(\boldsymbol{b}_{\mathrm{P},n})\ p(\sigma_{e,n}^2). \end{aligned} \tag{7}$$

2.2.3. Posterior distribution

The posterior distribution of the parameter vector $\boldsymbol{\theta}_n$ is obtained by using Bayes' rule, i.e.,

$$p(\boldsymbol{\theta}_n|\boldsymbol{x}_n, \boldsymbol{\alpha}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P},n-1}) \propto p(\boldsymbol{x}_n|\boldsymbol{\theta}_n) p(\boldsymbol{\theta}_n|\boldsymbol{\alpha}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P},n-1}) \tag{8}$$

where the right-hand term can be further expressed and factored using (4) and (7). Because the proposed method works sequentially and all estimates from the previous beat are available, we can substitute the estimates $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$ and $\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}$ for $\boldsymbol{\alpha}_{\mathrm{T},n-1}$ and $\boldsymbol{\alpha}_{\mathrm{P},n-1}$ in $p(\boldsymbol{\theta}_n|\boldsymbol{x}_n, \boldsymbol{\alpha}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P},n-1})$ when estimating $\boldsymbol{\theta}_n$ based on (8). Due to the complexity of the posterior distribution, we propose to use a Monte Carlo (sample-based) detection/estimation method. More specifically, we propose a BGS that generates samples asymptotically distributed according to $p(\boldsymbol{\theta}_n|\boldsymbol{x}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1})$ (see Section 2.3). From these samples, the discrete parameters $\boldsymbol{b}_{\mathrm{T},n}$ and $\boldsymbol{b}_{\mathrm{P},n}$ are then detected by means of the sample-based maximum a posteriori (MAP) detector, and the continuous parameters $\boldsymbol{\alpha}_{\mathrm{T},n}$, $\boldsymbol{\alpha}_{\mathrm{P},n}$, and $\sigma_{e,n}^2$ are estimated by means of the sample-based minimum mean square error (MMSE) estimator, as described in Section 4.

2.3. Block Gibbs sampler for beat-to-beat wave extraction

The proposed BGS for the nth non-QRS interval $\mathcal{J}_n$ is summarized in Algorithm 1. Note that the interval index n is omitted for all parameters to simplify the notation, while the index $n-1$ is kept to avoid any ambiguity. The term "block Gibbs sampler" is used to reflect the block constraints related to the wave indicator vectors $\boldsymbol{b}_{\mathrm{T}}$ and $\boldsymbol{b}_{\mathrm{P}}$, which are encompassed in the corresponding priors (see (5)). To see that Algorithm 1 is a valid Gibbs sampler, note that the sampling steps for $\boldsymbol{b}_{\mathrm{T}}$ and $\boldsymbol{\alpha}_{\mathrm{T}}$ are equivalent to jointly sampling $\boldsymbol{b}_{\mathrm{T}}$ and $\boldsymbol{\alpha}_{\mathrm{T}}$ from $p(\boldsymbol{b}_{\mathrm{T}}, \boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \sigma_e^2, \boldsymbol{x})$, and similarly for $\boldsymbol{b}_{\mathrm{P}}$ and $\boldsymbol{\alpha}_{\mathrm{P}}$. Closed-form expressions of the sampling distributions used in Algorithm 1 are presented and derived in the technical report [23].

Algorithm 1. Block Gibbs sampler.

Sample $\boldsymbol{b}_{\mathrm{T}}$ from $p(\boldsymbol{b}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \sigma_e^2, \boldsymbol{x})$
Sample $\boldsymbol{\alpha}_{\mathrm{T}}$ from $p(\boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}}, \boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \sigma_e^2, \boldsymbol{x})$
Sample $\boldsymbol{b}_{\mathrm{P}}$ from $p(\boldsymbol{b}_{\mathrm{P}}|\boldsymbol{b}_{\mathrm{T}}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}, \boldsymbol{\alpha}_{\mathrm{T}}, \sigma_e^2, \boldsymbol{x})$
Sample $\boldsymbol{\alpha}_{\mathrm{P}}$ from $p(\boldsymbol{\alpha}_{\mathrm{P}}|\boldsymbol{b}_{\mathrm{T}}, \boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}, \boldsymbol{\alpha}_{\mathrm{T}}, \sigma_e^2, \boldsymbol{x})$
Sample σ_e^2 from $p(\sigma_e^2|\boldsymbol{b}_{\mathrm{T}}, \boldsymbol{b}_{\mathrm{P}}, \boldsymbol{\alpha}_{\mathrm{T}}, \boldsymbol{\alpha}_{\mathrm{P}}, \boldsymbol{x})$

3. Dynamic beat-to-beat Bayesian model and marginalized particle filter

Section 2 presented a beat-to-beat Bayesian model that describes dependencies among waveform coefficients. A prior "with memory" (depending on the previous estimates of the P and T waveforms) was assigned to the current beat. In this section, elaborating on [24], an MPF method [25] is proposed to take into account *all the*

[1] $\|\cdot\|$ denotes the ℓ_2 norm, i.e., $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$.

information contained in the past of the current beat to be processed. First, we present a dynamic model as a basis for performing simultaneously P and T wave delineation and waveform estimation on a beat-to-beat basis. This dynamic model is similar to the Bayesian model introduced in Section 2. However, it adapts to the morphology variations across the ECG beats by using a random walk model for the waveform coefficients. Then, following the sequential Monte Carlo principle, an MPF is used to estimate the unknown parameters of the proposed model. The idea is to generate particles only for the states appearing nonlinearly in the dynamics and run one Kalman filter for each of these particles to estimate the "linear" parameters.

3.1. Dynamic signal model for non-QRS intervals

As in Section 2, we assume that the locations of the non-QRS intervals have been determined and baseline wanderings have been removed by a preprocessing stage. The signal model is the same as in Section 2.1, except for the following two differences.

First, the model (1) is split into its T and P parts:

$$x_{n,k} = \sum_{j=1}^{N_{\mathrm{T},n}} h_{\mathrm{T},n,k-j} b_{\mathrm{T},n,j} + e_{n,k}, \quad k \in \mathcal{J}_{\mathrm{T},n} = \{1, \ldots, N_{\mathrm{T},n}\} \tag{9}$$

$$x_{n,k} = \sum_{j=1}^{N_{\mathrm{P},n}} h_{\mathrm{P},n,k-j-N_{\mathrm{T},n}} b_{\mathrm{P},n,j} + e_{n,k}, \quad k \in \mathcal{J}_{\mathrm{P},n} = \{N_{\mathrm{T},n}+1, \ldots, N_n\}. \tag{10}$$

Using (2), we obtain the following representation of the signal vector $\boldsymbol{x}_{\mathrm{T},n} = (x_{n,1} \cdots x_{n,N_{\mathrm{T},n}})^T$ corresponding to the T wave interval in (9):

$$\boldsymbol{x}_{\mathrm{T},n} = \tilde{\boldsymbol{B}}_{\mathrm{T},n} \boldsymbol{H} \boldsymbol{\alpha}_{\mathrm{T},n} + \boldsymbol{e}_{\mathrm{T},n} \tag{11}$$

where $\tilde{\boldsymbol{B}}_{\mathrm{T},n}$ comprises the first $N_{\mathrm{T},n}$ rows of $\boldsymbol{B}_{\mathrm{T},n}$ defined in Section 2.1.3. A similar representation can be obtained for the signal vector $\boldsymbol{x}_{\mathrm{P},n} = (x_{n,N_{\mathrm{T},n}+1} \cdots x_{n,N_n})^T$ corresponding to the P wave interval in (10) using $\tilde{\boldsymbol{B}}_{\mathrm{P},n}$, which comprises the last $N_{\mathrm{P},n}$ rows of $\boldsymbol{B}_{\mathrm{P},n}$.

Second, the variance of the noise $e_{n,k}$ has not been included in the parameter vector (as in the proposed BGS) since it would increase significantly the computational complexity of the algorithm. In our simulations, the noise variance was estimated in a preprocessing step using the BGS (although other methods could be used as well). Thus, $\boldsymbol{e}_{\mathrm{T},n} = (e_{n,1} \ldots\ e_{n,N_{\mathrm{T},n}})^T$ and $\boldsymbol{e}_{\mathrm{P},n} = (e_{n,N_{\mathrm{T},n}+1} \ldots e_{n,N_n})^T$ are Gaussian vectors with zero mean and covariance matrix $\sigma_e^2 \mathbf{I}_{N_{\mathrm{T},n}}$ and $\sigma_e^2 \mathbf{I}_{N_{\mathrm{P},n}}$, respectively, σ_e^2 being the estimated noise variance.

3.2. Likelihood function, posterior, and prior

Using the modified signal model from Section 3.1, the likelihood function—now taking into account all beat indices up to n—factors as

$$p(\boldsymbol{x}_{1:n}|\boldsymbol{\theta}_{0:n}) = p(\boldsymbol{x}_{\mathrm{T},1:n}|\boldsymbol{b}_{\mathrm{T},0:n}, \boldsymbol{\alpha}_{\mathrm{T},0:n})\, p(\boldsymbol{x}_{\mathrm{P},1:n}|\boldsymbol{b}_{\mathrm{P},0:n}, \boldsymbol{\alpha}_{\mathrm{P},0:n}). \tag{12}$$

Here, e.g., $\boldsymbol{x}_{1:n} \triangleq (\boldsymbol{x}_1^T \ldots \boldsymbol{x}_n^T)^T$ and $\boldsymbol{\theta}_{0:n} \triangleq (\boldsymbol{\theta}_0^T \ldots \boldsymbol{\theta}_n^T)^T$. As before (cf. Section 2.2.2), we assume that the T wave parameters are independent of the P wave parameters. Therefore, using (12), the joint posterior distribution can be written

$$p(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n}) \propto p(\boldsymbol{b}_{\mathrm{T},0:n}, \boldsymbol{\alpha}_{\mathrm{T},0:n}|\boldsymbol{x}_{\mathrm{T},1:n}) p(\boldsymbol{b}_{\mathrm{P},0:n}, \boldsymbol{\alpha}_{\mathrm{P},0:n}|\boldsymbol{x}_{\mathrm{P},1:n}).$$

This allows us to split the estimation problem into two independent problems related to the P and T waves. In the following, only the T wave dynamic model and estimation problem are discussed, and the subscript T is omitted for notational convenience.

Due to the parametrization (11), the state vector for the nth T wave interval is given by $\boldsymbol{\theta}_n = (\boldsymbol{b}_n^T\ \boldsymbol{\alpha}_n^T)^T$. Note that $\boldsymbol{\theta}_n$ is now short for $\boldsymbol{\theta}_{\mathrm{T},n}$, and thus different from the $\boldsymbol{\theta}_n$ used, e.g., in Section 2.2. For the indicator vector $\boldsymbol{b}_n$, we use the prior in (5) with $p_0 = p_1 = 1/(N_{\mathrm{T},n}+1)$. This prior is a uniform distribution on the set of all possible $\boldsymbol{b}_n$ such that $\|\boldsymbol{b}_n\| = 1$ or $\|\boldsymbol{b}_n\| = 0$. Indicator vectors $\boldsymbol{b}_n$ for different beat indices n are assumed to be statistically independent. Since the ECG waveforms are usually similar for two consecutive beats, we propose to assign a random walk prior to the T waveform coefficient vector $\boldsymbol{\alpha}_n$, i.e.,

$$\boldsymbol{\alpha}_n = \boldsymbol{\alpha}_{n-1} + \boldsymbol{v}_{n-1} \tag{13}$$

where $\boldsymbol{\alpha}_{n-1}$ denotes the T waveform coefficient vector of the $(n-1)$th beat and the vectors $\boldsymbol{v}_n \sim \mathcal{N}(\boldsymbol{0}, \sigma_\alpha^2 \mathbf{I}_G)$ are statistically independent (of each other and of $\boldsymbol{\alpha}_{0:n}$) additive white Gaussian noise vectors. This leads to the conditional prior $p(\boldsymbol{\alpha}_n|\boldsymbol{\alpha}_{n-1}) = \mathcal{N}(\boldsymbol{\alpha}_{n-1}, \sigma_\alpha^2 \mathbf{I}_G)$, which is the same as in the second case of (6). Note that here, in contrast to (6), the coefficient vector changes even if $\|\boldsymbol{b}_n\| = 0$. The variance σ_α^2 depends on how fast the waveform coefficients are expected to change with time. Since the non-QRS components are normalized by dividing by the amplitude of the respective R peak, we have to account for possible significant variations of the waveforms with time. We therefore propose to use a large value of σ_α^2, which corresponds to a non-informative conditional prior of $\boldsymbol{\alpha}_n$. Note that the value of σ_α^2 can be further adjusted by an expert or by calculating the ECG waveform variance of an example ECG segment in an off-line parameter selection procedure as in [26]. Because of (13) and the independence of $\boldsymbol{v}_n$ for different n as well as of $\boldsymbol{\alpha}_{0:n}$, the waveform coefficient vector $\boldsymbol{\alpha}_n$ is conditionally independent, given $\boldsymbol{\alpha}_{n-1}$, of all previous coefficient vectors $\boldsymbol{\alpha}_{0:n-2}$, i.e., $p(\boldsymbol{\alpha}_n|\boldsymbol{\alpha}_{0:n-1}) = p(\boldsymbol{\alpha}_n|\boldsymbol{\alpha}_{n-1})$.

3.3. A marginalized particle filter for beat-to-beat wave analysis

Our goal is to estimate jointly the discrete-valued indicator vector $\boldsymbol{b}_n$ and the waveform vector $\boldsymbol{\alpha}_n$, i.e., to estimate the state vector $\boldsymbol{\theta}_n$. In a Bayesian framework, all inference is based on the posterior distribution of the unknown parameters given the set of available observations, expressed as $p(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n})$. Particle filters (PFs) are a class of methods well-suited to perform the estimation of the hybrid state vector $\boldsymbol{\theta}_{0:n}$. They approximate the target distribution by an empirical distribution

$$\hat{p}(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)} \delta(\boldsymbol{\theta}_{0:n} - \boldsymbol{\theta}_{0:n}^{(i)}), \quad \text{where} \quad \sum_{i=1}^{N_s} w_n^{(i)} = 1.$$

The weights $w_n^{(i)}$ and the particles $\boldsymbol{\theta}_{0:n}^{(i)}$ are classically obtained by sequential importance sampling and a selection (resampling) step to prevent degeneracy [25].

3.3.1. Development of the MPF

While the classical PFs are fairly easy to implement, a main drawback is that, in practice, the required number of particles increases quickly with the state dimension. The MPF can reduce the number of parameters estimated by the PF and therefore allows fewer particles to be used. More specifically, the MPF takes advantage of linear Gaussian substructures in the state parameters $\boldsymbol{\theta}_n$ to decrease the variance of the state estimates. The key idea is to split $\boldsymbol{\theta}_n$ into two parts $\boldsymbol{\theta}_n^{\mathrm{L}}$ and $\boldsymbol{\theta}_n^{\mathrm{NL}}$, where $\boldsymbol{\theta}_n^{\mathrm{L}}$ denotes the state parameters with conditionally linear dynamics and $\boldsymbol{\theta}_n^{\mathrm{NL}}$ denotes the nonlinear state parameters. We can then marginalize out $\boldsymbol{\theta}_n^{\mathrm{L}}$ and generate particles distributed according to $p(\boldsymbol{\theta}_n^{\mathrm{NL}}|\boldsymbol{x}_{1:n})$ using a PF. The particles are finally used to compute the MAP estimator of $\boldsymbol{\theta}_n^{\mathrm{NL}}$. In parallel, each particle is associated with a Kalman filter (KF) that computes recursively the mean and covariance matrix of the Gaussian distribution $p(\boldsymbol{\theta}_n^{\mathrm{L}}|\boldsymbol{\theta}_n^{\mathrm{NL}}, \boldsymbol{x}_{1:n})$.

It can be observed from (11) that both the discrete vector $\boldsymbol{b}_n$ and the continuous vector $\boldsymbol{\alpha}_n$ enter linearly in the observation $\boldsymbol{x}_n$, given the respective other parameter. Since only continuous parameters can be handled by the KF, we choose $\boldsymbol{\theta}_n^{\mathrm{L}} = \boldsymbol{\alpha}_n$ and $\boldsymbol{\theta}_n^{\mathrm{NL}} = \boldsymbol{b}_n$.The KF and the PF correspond to two factors of the joint posterior according to the following factorization:

$$p(\boldsymbol{b}_{0:n}, \boldsymbol{\alpha}_{0:n}|\boldsymbol{x}_{1:n}) = \underbrace{p(\boldsymbol{\alpha}_{0:n}|\boldsymbol{b}_{0:n}, \boldsymbol{x}_{1:n})}_{\mathrm{KF}} \underbrace{p(\boldsymbol{b}_{0:n}|\boldsymbol{x}_{1:n})}_{\mathrm{PF}}. \tag{14}$$

The marginal distribution of the discrete parameters is approximated by

$$\hat{p}(\boldsymbol{b}_{0:n}|\boldsymbol{x}_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)} \delta[\boldsymbol{b}_{0:n} - \boldsymbol{b}_{0:n}^{(i)}] \tag{15}$$

where N_s is the number of particles and $\delta[.]$ denotes the discrete-time unit sample. Then, by inserting (15) in (14) and summing out $\boldsymbol{b}_{0:n}$, the posterior distribution of the continuous parameters can be approximated by

$$\hat{p}(\boldsymbol{\alpha}_{0:n}|\boldsymbol{x}_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)} p(\boldsymbol{\alpha}_{0:n}|\boldsymbol{b}_{0:n}^{(i)}, \boldsymbol{x}_{1:n}). \tag{16}$$

Integrating out $\boldsymbol{\alpha}_{0:n-1}$ yields

$$\hat{p}(\boldsymbol{\alpha}_n|\boldsymbol{x}_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)} p(\boldsymbol{\alpha}_n|\boldsymbol{b}_{0:n}^{(i)}, \boldsymbol{x}_{1:n}). \tag{17}$$

It can be shown that $p(\boldsymbol{\alpha}_{0:n}|\boldsymbol{b}_{0:n}^{(i)}, \boldsymbol{x}_{1:n})$ in (16) and $p(\boldsymbol{\alpha}_n|\boldsymbol{b}_{0:n}^{(i)}, \boldsymbol{x}_{1:n})$ in (17) are Gaussian. Therefore, (16) and (17) represent mixtures of Gaussian distributions. Note that one KF is associated with each particle $\boldsymbol{b}_{0:n}^{(i)}$ with $i = 1, \ldots, N_s$. Furthermore, in practice, only the marginal distribution $\hat{p}(\boldsymbol{\alpha}_n|\boldsymbol{x}_{1:n})$ is updated (rather than $\hat{p}(\boldsymbol{\alpha}_{0:n}|\boldsymbol{x}_{1:n})$). The MPF recursions are summarized in Algorithm 2, presented for the T wave case. The different steps involved in this algorithm are detailed in the rest of this section.

Algorithm 2. Marginalized particle filter.

{Initialization}
for $i = 1, \ldots, N_s$ **do**
  Set $\boldsymbol{b}_0^{(i)} = \boldsymbol{0}_{N_{\mathrm{T},n} \times 1}$, $\boldsymbol{P}_0^{(i)} = \boldsymbol{0}_{G \times G}$, and $w_0^{(i)} = 1$, and choose a suitable initialization of the waveform coefficients $\hat{\boldsymbol{\alpha}}_0^{(i)}$ (see Section 5.1.1).
end for
{Time recursion}
for $n = 1, 2, \ldots$ **do**
  for $i = 1, \ldots, N_s$ **do**
    {KF and PF propagation}
    KF prediction for $\boldsymbol{\alpha}_n^{(i)}$ (see (18))
    Sample $\boldsymbol{b}_n^{(i)} \sim \Pr(\boldsymbol{b}_n = \boldsymbol{\beta}_k | \boldsymbol{b}_{0:n-1}^{(i)}, \boldsymbol{x}_{1:n})$ (see(19))
    KF correction for $\boldsymbol{\alpha}_n^{(i)}$ (see (20))
    Calculate weights (see (21) and (22))
    $\tilde{w}_n^{(i)} = w_{n-1}^{(i)} \sum_{k \in \mathcal{J}_{\mathrm{T},n}} p(\boldsymbol{x}_n | \boldsymbol{b}_n^{(i)} = \boldsymbol{\beta}_k, \boldsymbol{b}_{0:n-1}^{(i)}, \boldsymbol{x}_{1:n-1}) \Pr(\boldsymbol{b}_n^{(i)} = \boldsymbol{\beta}_k)$
  end for
  {Weight normalization}
  for $i = 1, \ldots, N_s$ **do**
    $w_n^{(i)} = \tilde{w}_n^{(i)} / \sum_{j=1}^{N_s} \tilde{w}_n^{(j)}$
  end for
  {State estimation}
  Estimate $\boldsymbol{b}_n$ and $\boldsymbol{\alpha}_n$ (see (23))
  {Particle resampling}
  Calculate $\hat{N}_{\mathrm{eff}} = 1 / \sum_{i=1}^{N_s} (w_n^{(i)})^2$
  if $\hat{N}_{\mathrm{eff}} \leq 0.7 \cdot N_s$ **then**
    Resample using systematic sampling scheme [25, p. 11]
  end if
end for

3.3.2. Kalman filter prediction

At time n, the previous MMSE state estimate is $\hat{\boldsymbol{\alpha}}_{n-1}^{(i)} = \mathrm{E}\{\boldsymbol{\alpha}_{n-1}|\boldsymbol{x}_{1:n-1}, \boldsymbol{b}_{0:n-1}^{(i)}\}$ and its covariance matrix is $\boldsymbol{P}_{n-1}^{(i)} = \mathrm{Cov}\{\boldsymbol{\alpha}_{n-1}|\boldsymbol{x}_{1:n-1}, \boldsymbol{b}_{0:n-1}^{(i)}\}$. We define the predicted state vector $\hat{\boldsymbol{\alpha}}_{n|n-1}^{(i)} \triangleq \mathrm{E}\{\boldsymbol{\alpha}_n|\boldsymbol{x}_{1:n-1}, \boldsymbol{b}_{0:n-1}^{(i)}\}$ and its covariance $\boldsymbol{P}_{n|n-1}^{(i)} \triangleq \mathrm{Cov}\{\boldsymbol{\alpha}_n|\boldsymbol{x}_{1:n-1}, \boldsymbol{b}_{0:n-1}^{(i)}\}$. Using (13), it can be shown that the prediction step of the KF can be written as

$$\hat{\boldsymbol{\alpha}}_{n|n-1}^{(i)} = \hat{\boldsymbol{\alpha}}_{n-1}^{(i)}, \quad \boldsymbol{P}_{n|n-1}^{(i)} = \boldsymbol{P}_{n-1}^{(i)} + \sigma_\alpha^2 \mathbf{I}_G. \tag{18}$$

Note that the predicted state vector and its covariance computed by the KF, $\hat{\boldsymbol{\alpha}}_{n|n-1}^{(i)}$ and $\boldsymbol{P}_{n|n-1}^{(i)}$, will be directly used to propagate the particles and compute their importance weights, as explained presently (see (20)).

3.3.3. Importance distribution for the indicators

It is well known that the choice of the importance distribution is a critical issue in the design of efficient PF algorithms. To generate samples in relevant regions of the state space, i.e., corresponding to a high likelihood $p(\boldsymbol{x}_n|\boldsymbol{\theta}_n)$, a natural strategy consists of taking into account information from the most recent observations $\boldsymbol{x}_n$. The importance distribution that is optimal in the sense that it minimizes the variance of the importance weights is $q(\boldsymbol{b}_n|\boldsymbol{b}_{0:n-1}^{(i)}, \boldsymbol{x}_{1:n}) = p(\boldsymbol{b}_n|\boldsymbol{b}_{0:n-1}^{(i)}, \boldsymbol{x}_{1:n})$ [27]. Thus, the optimal importance distribution for $\boldsymbol{b}_n$ is obtained as

$$\Pr(\boldsymbol{b}_n = \boldsymbol{\beta}_k | \boldsymbol{b}_{0:n-1}^{(i)}, \boldsymbol{x}_{1:n}) \propto p(\boldsymbol{x}_n | \boldsymbol{b}_n = \boldsymbol{\beta}_k, \boldsymbol{b}_{0:n-1}^{(i)}, \boldsymbol{x}_{1:n-1}) \Pr(\boldsymbol{b}_n = \boldsymbol{\beta}_k) \tag{19}$$

where $\boldsymbol{\beta}_k$ for $k \in \mathcal{J}_{\mathrm{T},n} = \{1, \ldots, N_{\mathrm{T},n}\}$ is an $N_{\mathrm{T},n} \times 1$ vector whose kth entry is 1 and all remaining entries are zero. Note that $\boldsymbol{\beta}_0$ is the all-zero vector, which represents the case where there is no T wave. It can be shown that, for $\boldsymbol{b}_{0:n}$ given, $\boldsymbol{\alpha}_{0:n}$ and $\boldsymbol{x}_{1:n}$ are jointly Gaussian. It follows that the distribution $p(\boldsymbol{x}_n|\boldsymbol{b}_n = \boldsymbol{\beta}_k, \boldsymbol{b}_{0:n-1}^{(i)}, \boldsymbol{x}_{1:n-1})$ in (19) is a Gaussian one. According to (11), its mean $\hat{\boldsymbol{x}}_{n,k}^{(i)}$ and covariance matrix $\boldsymbol{S}_{n,k}^{(i)}$ can be computed from the KF prediction (18) as follows:

$$\hat{\boldsymbol{x}}_{n,k}^{(i)} = \tilde{\boldsymbol{B}}_{n,k} \boldsymbol{H} \hat{\boldsymbol{\alpha}}_{n|n-1}^{(i)}$$

$$\boldsymbol{S}_{n,k}^{(i)} = \tilde{\boldsymbol{B}}_{n,k} \boldsymbol{H} \boldsymbol{P}_{n|n-1}^{(i)} \boldsymbol{H}^T \tilde{\boldsymbol{B}}_{n,k}^T + \sigma_e^2 \mathbf{I}_{N_{\mathrm{T},n}}$$

where $\tilde{\boldsymbol{B}}_{n,k}$ is the matrix $\tilde{\boldsymbol{B}}_n$ that corresponds to $\boldsymbol{b}_n = \boldsymbol{\beta}_k$. Note that contrary to the standard PF, the importance distribution for the indicators no longer depends on the coefficient vector $\boldsymbol{\alpha}_{0:n}$, which has been marginalized out. On the other hand, it depends on the past sequence $\boldsymbol{b}_{0:n-1}$.

3.3.4. Kalman filter correction

After receiving the observation $\boldsymbol{x}_n$ for beat index n, the predicted waveform coefficients $\hat{\boldsymbol{\alpha}}^{(i)}_{n|n-1}$ can be updated for each generated wave indicator particle $\boldsymbol{b}^{(i)}_n$. The KF correction procedure can be written as

$$\boldsymbol{S}^{(i)}_n = \tilde{\boldsymbol{B}}^{(i)}_n \boldsymbol{H} \boldsymbol{P}^{(i)}_{n|n-1} \boldsymbol{H}^T (\tilde{\boldsymbol{B}}^{(i)}_n)^T + \sigma_e^2 \mathbf{I}_{N_{\mathrm{T},n}} \tag{20a}$$

$$\boldsymbol{K}^{(i)}_n = \boldsymbol{P}^{(i)}_{n|n-1} \boldsymbol{H}^T (\tilde{\boldsymbol{B}}^{(i)}_n)^T (\boldsymbol{S}^{(i)}_n)^{-1} \tag{20b}$$

$$\hat{\boldsymbol{\alpha}}^{(i)}_n = \hat{\boldsymbol{\alpha}}^{(i)}_{n|n-1} + \boldsymbol{K}^{(i)}_n (\boldsymbol{x}_n - \tilde{\boldsymbol{B}}^{(i)}_n \boldsymbol{H} \hat{\boldsymbol{\alpha}}^{(i)}_{n|n-1}) \tag{20c}$$

$$\boldsymbol{P}^{(i)}_n = (\mathbf{I}_G - \boldsymbol{K}^{(i)}_n \tilde{\boldsymbol{B}}^{(i)}_n \boldsymbol{H}) \boldsymbol{P}^{(i)}_{n|n-1} \tag{20d}$$

where $\tilde{\boldsymbol{B}}^{(i)}_n$ is the matrix $\tilde{\boldsymbol{B}}_n$ that corresponds to $\boldsymbol{b}_n = \boldsymbol{b}^{(i)}_n$.

3.3.5. PF weight computation

When the optimal importance distribution is used to propagate the particles, the weights satisfy the following recursion:

$$w^{(i)}_n \propto w^{(i)}_{n-1} p(\boldsymbol{x}_n | \boldsymbol{x}_{1:n-1}, \boldsymbol{b}^{(i)}_{0:n-1}). \tag{21}$$

Here, $p(\boldsymbol{x}_n | \boldsymbol{x}_{1:n-1}, \boldsymbol{b}^{(i)}_{0:n-1})$ is the normalization constant of (19), i.e.,

$$p(\boldsymbol{x}_n | \boldsymbol{x}_{1:n-1}, \boldsymbol{b}^{(i)}_{0:n-1}) = \sum_{k \in \mathcal{J}_{\mathrm{T},n}} p(\boldsymbol{x}_n | \boldsymbol{b}^{(i)}_n = \boldsymbol{\beta}_k, \boldsymbol{b}^{(i)}_{0:n-1}, \boldsymbol{x}_{1:n-1}) \Pr(\boldsymbol{b}^{(i)}_n = \boldsymbol{\beta}_k). \tag{22}$$

4. P and T wave detection, estimation, and delineation

In this section, we discuss sample-based wave detection, parameter estimation, and wave delineation for the two proposed methods.

4.1. Block Gibbs sampler

We will denote by $\mathcal{S} \triangleq \{\boldsymbol{b}^{(i)}_{\mathrm{T}}, \boldsymbol{b}^{(i)}_{\mathrm{P}}, \boldsymbol{\alpha}^{(i)}_{\mathrm{T}}, \boldsymbol{\alpha}^{(i)}_{\mathrm{P}}, \sigma_e^{2(i)}\}_{i=1}^{N_s}$ the set of samples produced by our BGS after a burn-in period. (The burn-in period is the initial period of sampler iterations during which the sampler converges; the samples produced by the sampler during the burn-in period are not used for detection/estimation [28, p. 5].)

For detecting and locating P and T waves, we use the following sample-based blockwise MAP detector for the wave indicators $\boldsymbol{b}_{\mathrm{T}}$ and $\boldsymbol{b}_{\mathrm{P}}$:

$$\hat{\boldsymbol{b}}_{\mathrm{T}} = \underset{i \in \{1,\ldots,N_s\}}{\arg\max}\, p_{\mathcal{S}}(\boldsymbol{b}^{(i)}_{\mathrm{T}}), \quad \hat{\boldsymbol{b}}_{\mathrm{P}} = \underset{i \in \{1,\ldots,N_s\}}{\arg\max}\, p_{\mathcal{S}}(\boldsymbol{b}^{(i)}_{\mathrm{P}}).$$

Here, $p_{\mathcal{S}}(\boldsymbol{b}_{\mathrm{T}})$ is a sample-based approximation of the posterior probability $p(\boldsymbol{b}_{\mathrm{T}} | \boldsymbol{x}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1})$. More specifically, $p_{\mathcal{S}}(\boldsymbol{b}_{\mathrm{T}})$ is defined as the number of samples $\boldsymbol{b}^{(i)}_{\mathrm{T}}$ in $\mathcal{S}$ that equal the respective value of $\boldsymbol{b}_{\mathrm{T}}$, normalized by the total number of samples, N_s. Analogous considerations apply to $p_{\mathcal{S}}(\boldsymbol{b}_{\mathrm{P}})$.

The detection step described above is followed by sample-based estimation of the waveform coefficients $\boldsymbol{\alpha}_{\mathrm{T}}$ and $\boldsymbol{\alpha}_{\mathrm{P}}$ and of the noise variance σ_e^2. Let us combine these parameters into the parameter vector $\boldsymbol{\theta}_{\sim \boldsymbol{b}} \triangleq (\boldsymbol{\alpha}_{\mathrm{T}}^T \; \boldsymbol{\alpha}_{\mathrm{P}}^T \; \sigma_e^2)^T$. Furthermore, we define the set $\mathcal{I}$ as the set of sample indices $i \in \{1, \ldots, N_s\}$ such that $\boldsymbol{b}^{(i)}_{\mathrm{T}} = \hat{\boldsymbol{b}}_{\mathrm{T}}$ and $\boldsymbol{b}^{(i)}_{\mathrm{P}} = \hat{\boldsymbol{b}}_{\mathrm{P}}$. To estimate $\boldsymbol{\theta}_{\sim \boldsymbol{b}}$, we use the sample mean

$$\hat{\boldsymbol{\theta}}_{\sim \boldsymbol{b}} = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \boldsymbol{\theta}^{(i)}_{\sim \boldsymbol{b}}$$

where $\boldsymbol{\theta}^{(i)}_{\sim \boldsymbol{b}} \triangleq (\boldsymbol{\alpha}_{\mathrm{T}}^{(i)T} \; \boldsymbol{\alpha}_{\mathrm{P}}^{(i)T} \; \sigma_e^{2(i)})^T$ and $|\mathcal{I}|$ denotes the number of elements in $\mathcal{I}$. This can be interpreted as a sample-based approximation of the MMSE estimator (note that the MMSE estimator is given by the posterior mean $\mathrm{E}\{\boldsymbol{\theta}_{\sim \boldsymbol{b}} | \boldsymbol{x}, \boldsymbol{b}_{\mathrm{T}}, \boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}\}$). Thus, $\hat{\boldsymbol{\theta}}_{\sim \boldsymbol{b}}$ depends on $\hat{\boldsymbol{b}}_{\mathrm{T}}$, $\hat{\boldsymbol{b}}_{\mathrm{P}}$, $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$, and $\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}$.

The final step is wave delineation (localization of the peaks and boundaries of the P and T waves). Because of the convolution model (1), our detection/estimation problem is affected by a time-shift ambiguity [17,29]. Following [17], we resolve this ambiguity by performing an appropriate time shift after generating the waveform samples in the block Gibbs sampler. This time shift ensures that the maximum of the waveform is located at the center $k = 0$ of the waveform support interval $\{-L, \ldots, L\}$ and, thus, the location of a nonzero detected indicator $\hat{b}_{\mathrm{T},k} = 1$ or $\hat{b}_{\mathrm{P},k} = 1$ directly indicates the peak of the respective T or P wave. A detailed description of an algorithm for resolving the time-shift ambiguity is also provided in [29].

It is broadly accepted that the turning points defined by the largest local maximum of the *curvature* of the estimated waveform on each side of the detected wave peak are good estimates of the wave boundaries [6,30]. The curvature of the estimated T waveform $\hat{h}_{\mathrm{T},k}$ is defined as [6]

$$\kappa_{\mathrm{T},k} \triangleq \frac{\hat{h}''_{\mathrm{T},k}}{[1 + (\hat{h}'_{\mathrm{T},k})^2]^{3/2}}, \quad k \in \{-L, \ldots, L\}$$

where $\hat{h}'_{\mathrm{T},k}$ and $\hat{h}''_{\mathrm{T},k}$ are discrete-time counterparts of the first and second derivatives (e.g., $\hat{h}'_{\mathrm{T},k}$ is defined as the difference $\hat{h}_{\mathrm{T},k} - \hat{h}_{\mathrm{T},k-1}$). Using the turning points for delineation avoids the use of rigid detection and delineation thresholds. Fig. 2 illustrates the method by showing the delineation results obtained for three different T wave morphologies. Simulation results for the proposed BGS will be presented in Section 5.

4.2. Marginalized particle filter

In the MPF, the sample-based blockwise MAP estimator is used for estimating the binary sequence $\boldsymbol{b}_n$, while the sample-based MMSE estimator is used for estimating the waveform coefficients $\boldsymbol{\alpha}_n$:

$$\hat{\boldsymbol{b}}_n = \underset{i \in \{1,\ldots,N_s\}}{\arg\max}\, \hat{p}(\boldsymbol{b}^{(i)}_n | \boldsymbol{x}_{1:n}), \quad \hat{\boldsymbol{\alpha}}_n = \sum_{i=1}^{N_s} w^{(i)}_n \hat{\boldsymbol{\alpha}}^{(i)}_n. \tag{23}$$

Here, $\hat{p}(\boldsymbol{b}^{(i)}_n | \boldsymbol{x}_{1:n})$ is obtained by marginalizing (15) and the estimate $\hat{\boldsymbol{\alpha}}_n$ is the mean of the Gaussian mixture (17) with $\hat{\boldsymbol{\alpha}}^{(i)}_n$ computed recursively using (20c).

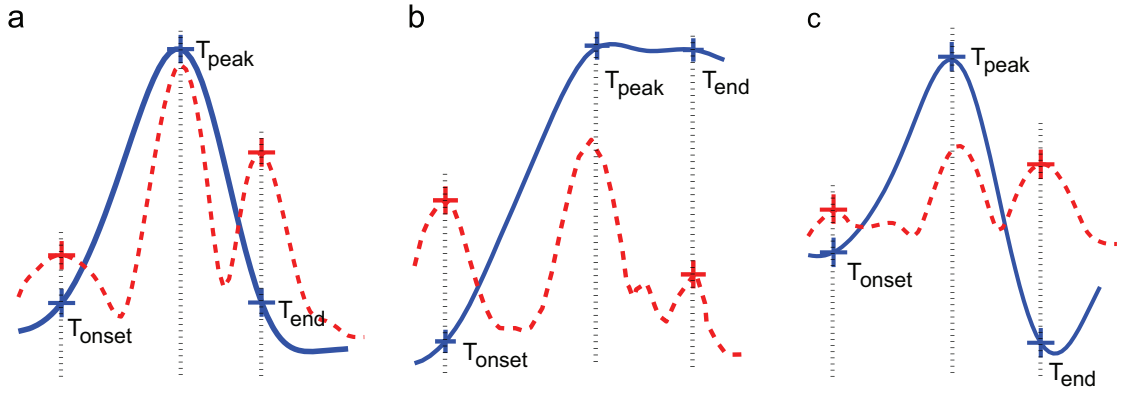

Fig. 2. Delineation results obtained for three different T wave morphologies. Solid blue line: estimated T waveform, dotted red line: corresponding curvature. The crosses indicate the estimated peak and boundary locations. (a) Normal sinus T wave from QT database (QTDB) [14] dataset sel17453 channel 1; (b) ascending T wave from QTDB sele0203 channel 1; (c) biphasic T wave from QTDB sele0603 channel 1. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

The wave delineation consists of determining the peaks and boundaries of the detected P and T waves. As mentioned in Section 4.1, because the time-shift ambiguity is removed, the nonzero wave indicator estimated by the MPF directly indicates the center of the corresponding waveform time window. Thus, the peak of the respective T or P wave is indicated by the location of the maximum of the estimated waveform. Furthermore, the wave boundaries can be located by applying the delineation criterion described in Section 4.1 to the estimated waveforms.

5. Simulation results

5.1. Simulation setup

Both of the proposed Bayesian wave detection/estimation/delineation methods were evaluated on the QT database (QTDB), which was previously used in several other studies [14]. The QTDB provides a reference for validating automatic wave-boundary estimation methods. It is a two-channel database containing cardiologist annotations for at least 30 beats per dataset for both channels. It includes 105 datasets from the widely used MIT-BIH arrhythmia database, the European ST-T database, and some other well-known databases. The cardiologist annotations of the QTDB were performed using two leads, whereas the proposed delineation methods work on a single-channel basis. To compare the single-channel delineation results produced by our methods with the manual annotations of the QTDB, we chose for each T or P wave the channel where the detected wave peak location was closer to the annotated one (as suggested in [4,30]). In a preprocessing step, the QRS complexes were detected and the borders of the non-QRS intervals $\mathcal{J}_n$ were determined using the Pan–Tompkins algorithm [15]. (The same preprocessing step was performed in [10,11].) In another preprocessing step, baseline wanderings were removed. P and T search intervals $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$ were then defined as the first and second half of $\mathcal{J}_n$. Both of the proposed methods sequentially process one non-QRS interval $\mathcal{J}_n$ after another.

5.1.1. BGS setup

For each non-QRS interval $\mathcal{J}_n$, the BGS generated 100 samples according to the conditional distributions specified in Algorithm 1. The first 40 samples constituted the burn-in period, and the remaining 60 were used for detection/estimation (thus, $N_s = 60$). The fixed hyperparameters involved in the prior distributions were chosen as $p_0 = 0.01$, $\sigma_\alpha^2 = 0.01$, $\xi = 11$, and $\eta = 0.5$; these values allow for an appropriate waveform variability from one beat to another and provide a noninformative prior for the noise variance $\sigma_{e,n}^2$. Note that the non-QRS components were normalized using the corresponding R peak values to handle different amplitude resolutions. For the first non-QRS interval ($n = 1$), the previous waveform coefficient estimates $\hat{\boldsymbol{\alpha}}_{\mathrm{T},0}$ and $\hat{\boldsymbol{\alpha}}_{\mathrm{P},0}$ were initialized with the coefficient vector $\boldsymbol{\alpha}$ for which $\boldsymbol{h}$ is closest to the $2L+1$ Hann window [31], with an amplitude equal to half the R peak amplitude. The waveform length was chosen as $2L+1 = N_n/3$, which is large enough to accommodate the actual support of the T or P wave.

Because the proposed beat-to-beat BGS method processes only one non-QRS interval at any given time, both its memory requirements and its computational complexity are smaller than those of the window-based method of [10]. For instance, for the proposed method using 100 sampler iterations, the processing time per beat is approximately 0.3 s using a nonoptimized MATLAB implementation running on a 3.0-GHz Pentium IV computer, compared to about 2 s for the method of [10]. Note that this computation time could be further reduced by developing implementations on digital signal processors.

5.1.2. MPF setup

In the MPF method, the fixed hyperparameters involved in the prior distributions were chosen as $\sigma_\alpha^2 = 0.01$ and $\sigma_e^2 = 0.1$. The chosen value of σ_α^2 allows for an appropriate waveform variability from one beat to another. The chosen value of σ_e^2 was obtained from a previous estimation of the noise level, using the mean value estimated by the BGS method, but could be taken from any other noise estimator. The non-QRS components were again normalized using the corresponding R peak values to handle different amplitude

resolutions. The waveform vector $\hat{\boldsymbol{h}}_0 = \boldsymbol{H}\hat{\boldsymbol{\alpha}}_0$ was initialized as in the BGS method.

An important issue with PF methods is the number of particles. Using the estimated parameters, we reconstructed the non-QRS part of the signal (based on the noiseless parts of models (9) and (10)) and compared it to the original signal non-QRS part. This allowed us to compute a normalized mean square error (NMSE) to assess the quality of the estimation. Table 1 shows the NMSE versus the number of particles N_s. As can be seen, benefiting from the optimal importance distribution derived in Section 3.3.3, good estimation performance can be obtained with a moderate number of particles. We chose $N_s = 200$ particles for all the following simulations in order to guarantee an NMSE close to -40 dB. For the MPF method using 200 particles, the processing time per beat is approximately 0.5 s using a nonoptimized MATLAB implementation running on a 3.0-GHz Pentium IV computer.

Table 1
Normalized mean square error (NMSE) versus number of particles used in the MPF method.

N_s	10	50	100	200	300
NMSE (dB)	−25	−31	−34	−40	−42

5.2. Qualitative analysis

In this section, we first show the posterior distributions as well as estimation and delineation results obtained by the proposed beat-to-beat BGS method on a typical example. Then, we present a qualitative comparison of the proposed BGS and MPF methods with state-of-the-art methods on several representative ECG segments.

Fig. 3(a) shows two consecutive beats from the QTDB dataset sele0136. The corresponding sample-based estimates of the marginal posterior probabilities of having a T

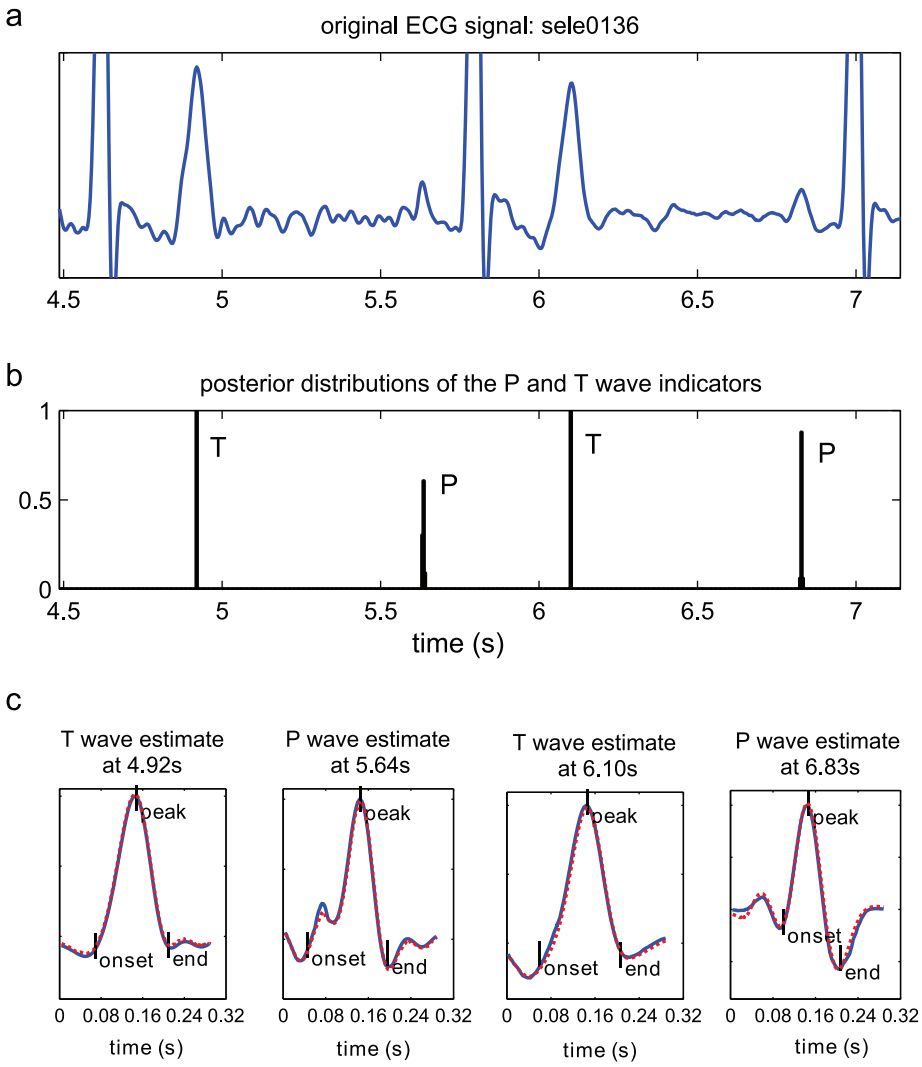

Fig. 3. Simulation results obtained with the proposed BGS method. (a) Two consecutive beats from QTDB dataset sele0136; (b) estimated marginal posteriors $P_S(b_k = 1)$; (c) estimated P and T waveforms (dotted red line) compared with the original P and T waveforms (blue) and delineation results. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

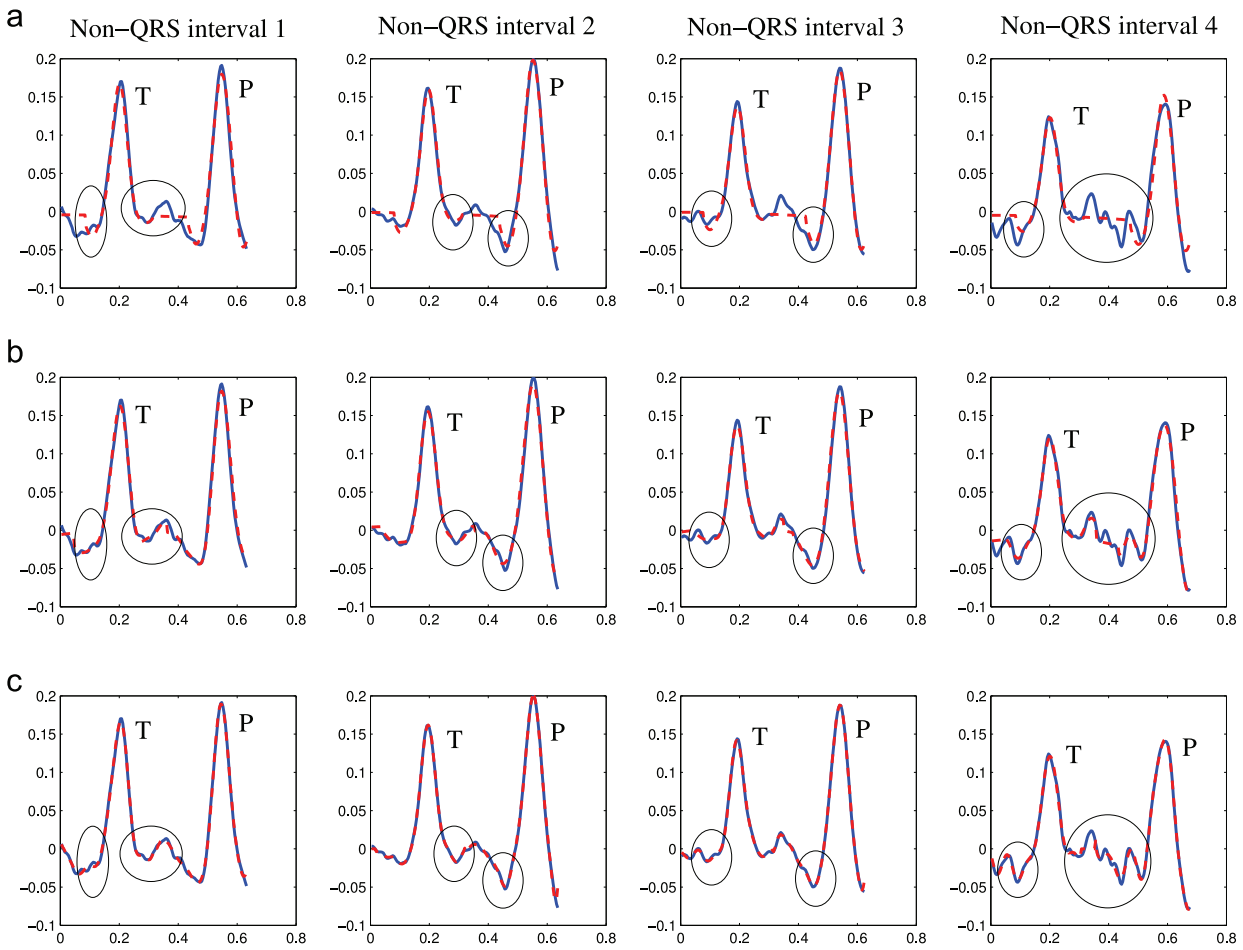

Fig. 4. Four consecutive segments from QTDB dataset sele0136 (blue) superimposed with their reconstructions using the estimated parameters (dotted red). (a) PCGS multi-beat method of [10]; (b) proposed beat-to-beat BGS method; (c) proposed beat-to-beat MPF method. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

or P wave at a given location k, $P_{\mathcal{S}}(b_{\mathrm{T},n,k}=1)$ and $P_{\mathcal{S}}(b_{\mathrm{P},n,k}=1)$, are depicted in Fig. 3(b). (For $k\in\mathcal{J}_{\mathrm{T}}$, $P_{\mathcal{S}}(b_k=1)$ equals the probability $P_{\mathcal{S}}(\boldsymbol{b}_{\mathrm{T}})$ of the specific hypothesis $\boldsymbol{b}_{\mathrm{T}}$ that contains a 1-entry at location k, and similarly for $k\in\mathcal{J}_{\mathrm{P}}$.) Fig. 3(c) shows the actual P and T waveforms and their estimates obtained by the proposed BGS method for each search interval, along with the corresponding delineation results (i.e., the estimated wave onsets, peaks, and ends, which were determined as described in Section 4). One can observe noticeable differences between the two consecutive T waveforms (at time instants 4.92 s and 6.10 s), as well as between the two consecutive P waveforms (at time instants 5.64 s and 6.83 s). This confirms the pseudo-cyclostationary nature of the ECG signal and justifies our introduction of a beat-to-beat processing scheme that allows for beat-to-beat variations of the P and T waveforms. The results displayed in Fig. 3(c) show that the BGS algorithm is able to estimate these P and T waveforms with good accuracy.

Next, we present a qualitative comparison of the proposed BGS and MPF methods with the multi-beat method of [10] (based on a partially collapsed Gibbs sampler (PCGS)) to highlight the benefits of beat-to-beat processing. To evaluate the methods under real physiological noise conditions, we added muscular activity noise from the MIT-BIH noise stress test database. The estimated non-QRS signal components obtained with the different methods are displayed in Figs. 4 and 5 for eight successive beats of a segment of QTDB dataset sele0136. The original ECG signal is also shown for comparison. It can be seen that the proposed beat-to-beat methods (BGS and MPF) provide closer agreements with the original ECG signal when compared to the multi-beat method, especially at the onsets and ends of the waves, which is a desirable property for wave delineation. These results show that, contrary to the multi-beat method of [10], the beat-to-beat BGS and MPF methods are able to capture the changes affecting the P and T waveforms. Additional results are available in a technical report [23]. In particular, both proposed methods are compared with the method of [7] (which is based on an extended Kalman filter) and are shown to be able to handle specific pathologies such as

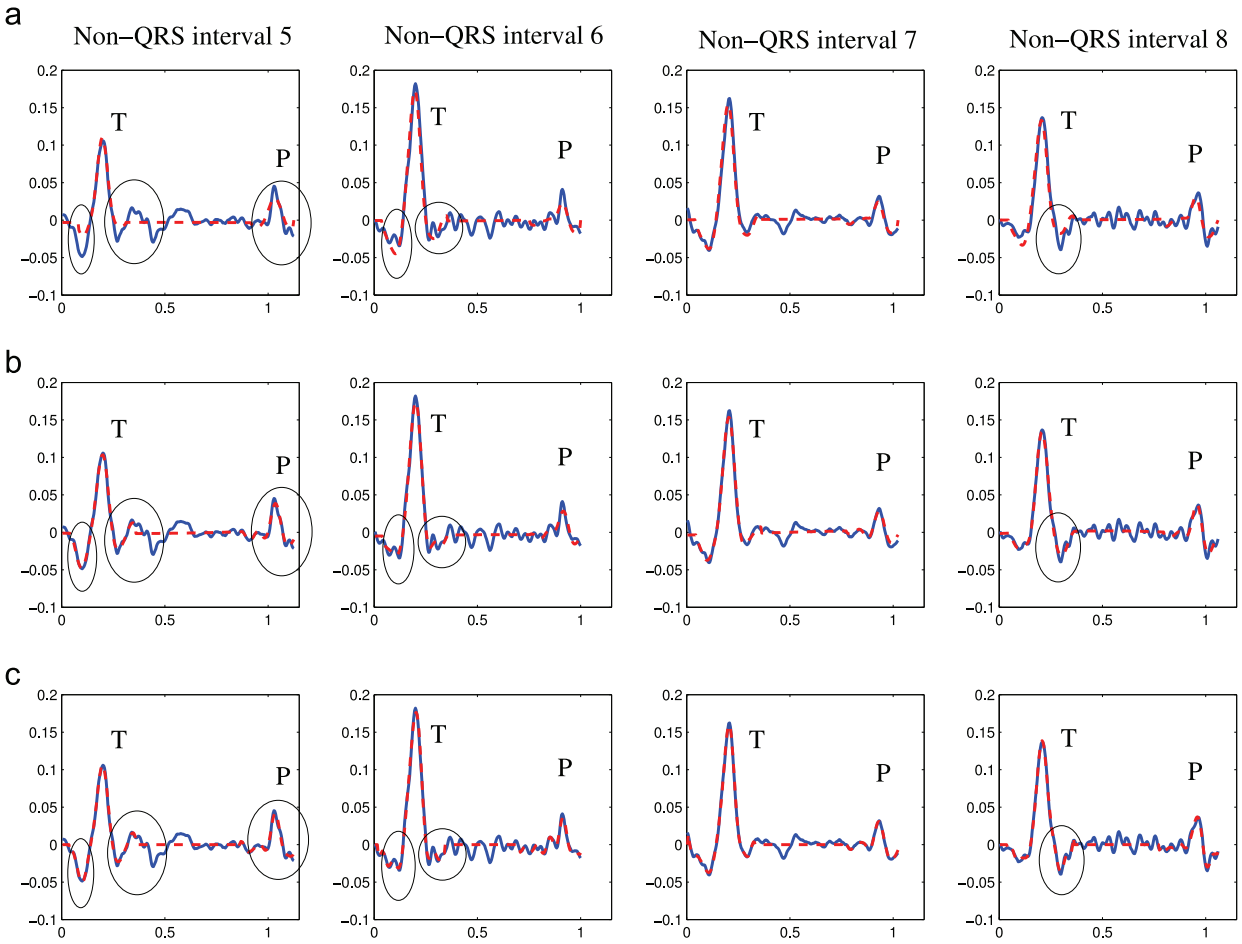

Fig. 5. Four other consecutive segments from QTDB dataset sele0136 (blue) superimposed with their reconstructions using the estimated parameters (dotted red). (a) PCGS multi-beat method of [10]; (b) proposed beat-to-beat BGS method; (c) proposed beat-to-beat MPF method. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

premature ventricular contractions, a pathology in which parts of the T waves are crossing the interval border and the P waves are missing.

5.3. Quantitative analysis

Next, we provide a quantitative performance comparison of the two proposed methods with the multi-beat method of [10] and three alternative methods [2,4,5], based on an exhaustive evaluation performed on the entire QTDB. For a quantitative analysis of the performance of P and T wave detection, as in [2,4,5,10,11], we computed the *sensitivity* (also referred to as *detection rate*) $\mathrm{Se} = \mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$ and the *positive predictivity* $P^+ = \mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$, where TP denotes the number of true positive detections (wave was present and was detected), FN stands for the number of false negative detections (wave was present but was missed), and FP for the number of false positive detections (wave was not present but was detected). The performance of wave delineation was measured by the average (denoted as m) and standard deviation (denoted as s) of the time differences between the results of the considered method and the corresponding cardiologist annotations. The indicated time values (in ms) are based on a sampling frequency of 250 Hz. The quantities m and s were computed separately for the wave onset times $t_{\mathrm{P,on}}$ and $t_{\mathrm{T,on}}$, the wave peak times $t_{\mathrm{P,peak}}$ and $t_{\mathrm{T,peak}}$, and the wave end times $t_{\mathrm{P,end}}$ and $t_{\mathrm{T,end}}$. We note that while the QTDB includes annotations made by two cardiologists, we considered only those of the first cardiologist, who provided annotations for at least 30 beats per dataset.

Table 2 shows the results for Se, P^+, and $m \pm s$ obtained for the entire QTDB. It can be seen that the two proposed methods detect the P and T waves annotated by the cardiologist with high sensitivity: the sensitivity Se is 100% for the T waves and 99.93% or 99.95% for the P waves. Similarly good results were obtained for the positive predictivity P^+, which is between 98.01% and 99.30% for the T waves and 99.10% or 99.23% for the P waves. Both the Se values and the P^+ values are typically better than those obtained with the other methods, including the recently proposed multi-beat method of [10]. Regarding the delineation performance, it is seen from Table 2 that

Table 2
Comparison of the detection and delineation performance of the proposed beat-to-beat BGS and MPF methods with that of the PCGS multi-beat method of [10], the wavelet transform based method of [4] (WT), the low-pass differentiation based method of [2] (LPD), and the action potential based method of [5]. The variances of these methods are compared with the delineation error tolerance of [32], which is provided in the last row. (N/A: not available).

Method	Parameters	$t_{P,on}$	$t_{P,peak}$	$t_{P,end}$	$t_{T,on}$	$t_{T,peak}$	$t_{T,end}$
Beat-to-beat BGS (proposed)	Annotations	3176	3176	3176	1345	3403	3403
	Se (%)	99.93	99.93	99.93	100	100	100
	P^+ (%)	99.10	99.10	99.10	98.01	99.30	99.30
	$m \pm s$ (ms)	3.4 ± 14.2	1.1 ± 5.3	-2.1 ± 9.8	6.8 ± 16.3	-0.8 ± 4.1	-3.1 ± 14.0
Beat-to-beat MPF (proposed)	Annotations	3176	3176	3176	1345	3403	3403
	Se (%)	99.95	99.95	99.95	100	100	100
	P^+ (%)	99.23	99.23	99.23	98.67	99.20	99.20
	$m \pm s$ (ms)	3.1 ± 8.3	1.2 ± 5.3	2.7 ± 9.8	6.5 ± 16.3	-0.4 ± 4.8	-3.8 ± 14.2
Multi-beat partially collapsed Gibbs sampler [10]	Annotations	3176	3176	3176	1345	3403	3403
	Se (%)	99.60	99.60	99.60	100	100	100
	P^+ (%)	98.04	98.04	98.04	97.23	99.15	99.15
	$m \pm s$ (ms)	1.7 ± 10.8	2.7 ± 8.1	2.5 ± 11.2	5.7 ± 16.5	0.7 ± 9.6	2.7 ± 13.5
WT [4]	Annotations	3194	3194	3194	N/A	3542	3542
	Se (%)	98.87	98.87	98.75	N/A	99.77	99.77
	P^+ (%)	91.03	91.03	91.03	N/A	97.79	97.79
	$m \pm s$ (ms)	2.0 ± 14.8	3.6 ± 13.2	1.9 ± 12.8	N/A	0.2 ± 13.9	-1.6 ± 18.1
LPD [2]	Annotations	N/A	N/A	N/A	N/A	N/A	N/A
	Se (%)	97.70	97.70	97.70	N/A	99.00	99.00
	P^+ (%)	91.17	91.17	91.17	N/A	97.74	97.74
	$m \pm s$ (ms)	14.0 ± 13.3	4.8 ± 10.6	-0.1 ± 12.3	N/A	-7.2 ± 14.3	13.5 ± 27.0
Action potential based method [5]	Annotations	N/A	N/A	N/A	N/A	N/A	N/A
	Se (%)	N/A	N/A	N/A	N/A	92.60	92.60
	P^+ (%)	N/A	N/A	N/A	N/A	N/A	N/A
	$m \pm s$ (ms)	N/A	N/A	N/A	20.9 ± 29.6	-12.0 ± 23.4	0.8 ± 30.3
Delineation error tolerance	$2s_{CSE}$ (ms)	10.2	N/A	12.7	N/A	N/A	30.6

the two proposed methods delineate the annotated P and T waves with mean errors $|m|$ not exceeding 4 ms (except for $t_{T,on}$) and with smaller standard deviations s than the other methods (with two exceptions). We note that delineation error tolerances have been recommended by the CSE Working Party [32]. In particular, the standard deviation s for $t_{P,on}$, $t_{P,end}$, and $t_{T,end}$ should be at most $2s_{CSE}$, which is listed in the last row of Table 2. However, a stricter recommendation proposed in [4] is $s \leq s_{CSE}$. According to Table 2, the standard deviations for $t_{P,end}$ achieved by both proposed methods and the standard deviation for $t_{P,on}$ achieved by the proposed MPF method comply with the loose recommendation. For the $t_{T,end}$ results, both proposed methods comply with the strict recommendation. In Table 2, the advantage of the proposed beat-to-beat methods over the multi-beat method of [10] is not as clear as in Figs. 4 and 5. This is because only a small part of the signals evaluated in Table 2 exhibit obvious inter-beat waveform variations. From Table 2, it is furthermore seen that the detection and delineation results obtained with the two proposed methods are quite similar.

To further appreciate the differences between the two proposed methods, we conducted a quantitative comparison of their waveform estimation performance. First, in order to constitute our ground truth, we visually selected from the MIT-BIH Normal Sinus Rhythm Database 20 segments of ECG signals, each of duration 10 s, with a high SNR and no significant arrhythmia. Then, to generate realistic ECG signals, the ground truth was corrupted by adding muscular activity noise from the MIT-BIH noise stress test database with an SNR ranging from 20 to

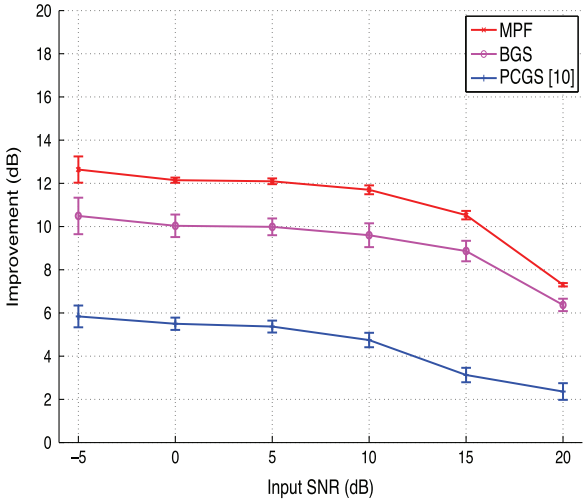

Fig. 6. Waveform estimation SNR improvement measure SNR_{imp} obtained with the proposed MPF and BGS methods and the multi-beat PCGS method of [10] versus the input SNR for 20 signal segments selected from the MIT-BIH Normal Sinus Rhythm database and corrupted by muscular activity noise.

−5 dB. For a quantitative evaluation, we considered the SNR improvement measure defined as

$$\mathrm{SNR}_{\mathrm{imp}} = 10 \log\left(\frac{\|\boldsymbol{x}-\boldsymbol{c}\|^2}{\|\boldsymbol{z}-\boldsymbol{c}\|^2}\right)$$

where $\boldsymbol{x}$ is the noisy signal, $\boldsymbol{c}$ is the clean signal, and $\boldsymbol{z}$ is the estimated signal. This evaluation was carried out only on the non-QRS intervals (P and T waveforms) of each signal. In order to obtain a fair performance comparison between the different algorithms, 20 Monte Carlo runs of each method were considered for each ECG signal segment. The output SNR was averaged over the 400 results for each input SNR (20 runs for each of the 20 signal segments). In Fig. 6, the means and standard deviations of the SNR improvement $\mathrm{SNR}_{\mathrm{imp}}$ are plotted versus the input SNR. It can be observed that the proposed BGS and MPF algorithms clearly outperform the PCGS method [10] in terms of SNR improvement. Furthermore, the MPF algorithm outperforms the BGS algorithm.

6. Conclusion

This paper presented and studied two Bayesian methods for beat-to-beat P and T wave delineation and waveform estimation. Instead of using a processing window that contains several successive beats involving the same P and T waveforms by assumption, the proposed methods account for beat-to-beat variations of the P and T waveforms by processing individual beats *sequentially* (i.e., with memory). First, a block Gibbs sampler (BGS) method was proposed to estimate the unknown parameters of the beat-to-beat Bayesian model. Alternatively, in order to take advantage of all the available information contained in the past of the beat to be processed, a dynamic model was proposed. This model exploits the sequential nature of the ECG signal by using a random walk model for the waveform coefficients. A marginalized particle filter (MPF) was then proposed to estimate the unknown parameters of the dynamic model.

The main features and contributions of this work can be summarized as follows:

1. Beat-to-beat BGS method
   - The proposed Bayesian model uses the P and T waveform estimates of the previous beat as prior information for detecting/estimating the current P and T waves.
   - By accounting for the local dependencies in, and the sequential nature of, ECG signals, the proposed BGS exhibits a faster convergence than the samplers used in [10,11].
   - The high accuracy of the proposed technique for P and T waveform estimation allows a threshold-free delineation technique to be used.
   - The beat-to-beat processing mode leads to smaller memory requirements and a lower computational complexity compared to the multi-beat Bayesian methods in [10,11].
2. Beat-to-beat MPF method
   - The sequential nature of the ECG signal is exploited by using a dynamic model within the Bayesian framework.
   - The proposed MPF method efficiently estimates the unknown parameters of the dynamic model. Thanks to the marginalization, a smaller number of particles is needed for good estimation performance, compared to the classical particle filter.
   - Compared to the BGS method, the MPF method is potentially advantageous in that it considers all the available beats in the waveform estimation.

The statistical models used for these two methods are similar, except for two minor differences: (1) the BGS processes P and T waves simultaneously, in each non-QRS interval, whereas the MPF processes them separately; (2) the MPF estimates the noise variance during a preprocessing step in order to obtain a reasonable computational complexity.

The proposed beat-to-beat Bayesian methods were validated using the QT database. A comparison with the method of [10] and with other benchmark methods demonstrated that both proposed methods can provide significant improvements regarding P and T wave detection rate, positive predictivity, and delineation accuracy. Moreover, whereas the delineation results obtained with the two proposed methods are quite similar, the MPF outperforms the BGS from a waveform estimation point of view, at the price of a higher computational cost. We note that the proposed methods are single-lead based ECG processing methods. They can be extended to multi-lead ECG signals by including post-processing decision rules to determine global marks from the single-lead delineation results [33].

Besides its suitability for real-time ECG monitoring, another advantage of the proposed beat-to-beat processing mode is the possibility of analyzing the beat-to-beat variation and evolution of the P and T waveforms. Potential clinic applications include T wave alternans (TWA) detection in intra-cardiac electrograms. This application is currently under investigation.

References

[1] N.V. Thakor, Y.S. Zhu, Application of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection, IEEE Trans. Biomed. Eng. 38 (1991) 785–793.
[2] P. Laguna, R. Jané, P. Caminal, Automatic detection of wave boundaries in multilead ECG signals: validation with the CSE database, Comput. Biomed. Res. 27 (1994) 45–60.
[3] C. Li, C. Zheng, C. Tai, Detection of ECG characteristic points using wavelet transforms, IEEE Trans. Biomed. Eng. 42 (1995) 21–28.
[4] J.P. Martínez, R. Almeida, S. Olmos, A.P. Rocha, P. Laguna, A wavelet-based ECG delineator: evaluation on standard databases, IEEE Trans. Biomed. Eng. 51 (2004) 570–581.
[5] J.A. Vila, Y. Gang, J. Presedo, M. Delgado, S. Barro, M. Malik, A new approach for TU complex characterization, IEEE Trans. Biomed. Eng. 47 (2000) 746–772.
[6] P. Trahanias, E. Skordalakis, Syntactic pattern recognition of the ECG, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 648–657.
[7] O. Sayadi, M.B. Shamsollahi, A model-based Bayesian framework for ECG beat segmentation, J. Physiol. Meas. 30 (2009) 335–352.
[8] J. Dumont, A.I. Hernandez, G. Carrault, Improving ECG beats delineation with an evolutionary optimization process, IEEE Trans. Biomed. Eng. 57 (2010) 607–615.
[9] M.B. Conover, Understanding Electrocardiography, Mosby, St. Louis, 2003.
[10] C. Lin, C. Mailhes, J.-Y. Tourneret, P- and T-wave delineation in ECG signals using a Bayesian approach and a partially collapsed Gibbs sampler, IEEE Trans. Biomed. Eng. 57 (2010) 2840–2849.

[11] C. Lin, G. Kail, J.-Y. Tourneret, C. Mailhes, F. Hlawatsch, P and T wave delineation and waveform estimation in ECG signals using a block Gibbs sampler, in: Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Process. (ICASSP), Prague, Czech Republic, 2011, pp. 537–540.
[12] A. Doucet, N. Gordon, V. Krishnamurthy, Particle filters for state estimation of jump Markov linear systems, IEEE Trans. Signal Process. 49 (2001) 613–624.
[13] T. Schön, F. Gustafsson, P.-J. Nordlund, Marginalized particle filters for mixed linear/nonlinear state-space models, IEEE Trans. Signal Process. 53 (2005) 2279–2289.
[14] P. Laguna, R. Mark, A. Goldberger, G. Moody, A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG, in: Proceedings of Computing in Cardiology, Lund, Sweden, 1997, pp. 673–676.
[15] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Trans. Biomed. Eng. 32 (1985) 230–236.
[16] V.S. Chouhan, S.S. Mehta, Total removal of baseline drift from ECG signal, in: Proceedings of the International Conference in Computing: Theory and Applications, Kolkata, India, 2007, pp. 512–515.
[17] Q. Cheng, R. Chen, T.-H. Li, Simultaneous wavelet estimation and deconvolution of reflection seismic signals, IEEE Trans. Geosci. Remote Sensing 34 (1996) 377–384.
[18] G. Kail, J.-Y. Tourneret, F. Hlawatsch, N. Dobigeon, Blind deconvolution of sparse pulse sequences under a minimum distance constraint: a partially collapsed Gibbs sampler method, IEEE Trans. Signal Process. 60 (2012) 2727–2743.
[19] L. Sörnmo, P. Börjesson, M. Nygårds, O. Pahlm, A method for evaluation of QRS shape features using a mathematical model for the ECG, IEEE Trans. Biomed. Eng. 28 (1981) 713–717.
[20] R. Jané, S. Olmos, P. Laguna, P. Caminal, Adaptive Hermite models for ECG data compression: performance and evaluation with automatic wave detection, in: Proceedings of Computing in Cardiology, London, UK, 1993, pp. 389–392.
[21] N. Dobigeon, J.-Y. Tourneret, M. Davy, Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a Bayesian sampling approach, IEEE Trans. Signal Process. 55 (2007) 1251–1263.
[22] C.P. Robert, The Bayesian Choice, Springer, New York, NY, 1996.
[23] C. Lin, G. Kail, A. Giremus, C. Mailhes, J.-Y. Tourneret, F. Hlawatsch, Sequential Beat-to-Beat P and T Wave Delineation and Waveform Estimation in ECG Signals: Block Gibbs Sampler and Marginalized Particle Filter, Technical Report, IRIT/ENSEEIHT/TéSA, http://mailhes.perso.enseeiht.fr/documents/TechReport2013.pdf, 2013.
[24] C. Lin, A. Giremus, C. Mailhes, J.-Y. Tourneret, Beat-to-beat P and T wave delineation in ECG signals using a marginalized particle filter, in: Proceedings of the EUSIPCO-12, Bucharest, Romania, 2012, pp. 479–483.
[25] A. Doucet, N. de Freitas, N. Gordon, Sequential Monte Carlo Methods in Practice, Springer, New York, NY, 2001.
[26] R. Sameni, M.B. Shamsollahi, C. Jutten, G.D. Clifford, A nonlinear Bayesian filtering framework for ECG denoising, IEEE Trans. Biomed. Eng. 54 (2007) 2172–2185.
[27] A. Doucet, S. Godsill, C. Andrieu, On sequential Monte Carlo sampling methods for Bayesian filtering, Stat. Comput. 10 (2000) 197–208.
[28] W.R. Gilks, S. Richardson, D.J. Spiegelhalter, Introducing Markov chain Monte Carlo, in: W.R. Gilks, S. Richardson, D.J. Spiegelhalter (Eds.), Markov Chain Monte Carlo in Practice, Chapman & Hall, London, UK, 1996, pp. 1–19.
[29] C. Labat, J. Idier, Sparse blind deconvolution accounting for time-shift ambiguity, in: Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Process. (ICASSP), Toulouse, France, 2006, pp. 616–620.
[30] P.-C. Chen, S. Lee, C.-D. Kuo, Delineation of T-wave in ECG by wavelet transform using multiscale differential operator, IEEE Trans. Biomed. Eng. 53 (2006) 1429–1433.
[31] F.J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, Proc. IEEE 66 (1978) 51–83.
[32] The CSE Working Party, Recommendations for measurement standards in quantitative electrocardiography, Eur. J. Heart 6 (1985) 815–825.
[33] R. Almeida, J.P. Martinez, A.P. Rocha, P. Laguna, Multilead ECG delineation using spatially projected leads from wavelet transform loops, IEEE Trans. Biomed. Eng. 56 (2009) 1996–2005.