Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

SIGNAL, IMAGE, ACOUSTIQUE ET OPTIMISATION

Présentée et soutenue par :
Chao LIN

le : lundi 2 juillet 2012

Titre :

Analyse des ondes P et T des signaux ECG à l'aide de méthodes Bayésiennes

---

Ecole doctorale :
Mathématiques Informatique Télécommunications (MITT)

Unité de recherche :
IRIT - Institut de Recherche en Informatique de Toulouse

Directeur(s) de Thèse :
Prof. Corinne MAILHES, INP Toulouse - ENSEEIHT
Prof. Jean-Yves TOURNERET, INP Toulouse - ENSEEIHT

Rapporteurs :

Prof. David BRIE, Université Henry Poincaré de Nancy
Prof. Lotfi SENHADJI, Université de Rennes 1

Membre(s) du jury :
Prof. Denis KOUAME, Université Paul Sabatier (Président)
Dr. Audrey GIREMUS, Université Bordeaux 1 (Examinateur)
Prof. Franz HLAWATSCH, Vienna University of Technology (Examinateur)
Prof. José Carlos BERMUDEZ, Federal University of Santa Catarina (Invité)
Mme. Marie RAZANI, St. Jude Medical (Invité)

Acknowledgement / Remerciements

Il y a cinq ans, je n'aurais jamais imaginé qu'un jour, je pourrais rédiger ces quelques lignes de gratitude qui viennent du fond de mon coeur en utilisant cette langue que je ne connaissais pas auparavant. La présente étude n'aurait pas été possible sans le bienveillant soutien de certaines personnes. Et je ne suis pas non plus capable de dire dans les mots qui conviennent, le rôle qu'elles ont pu jouer à mes côtés pour en arriver là.

Je tiens dans un premier temps à remercier mes directeurs de thèse, Corinne Mailhes, professeur à l'INP-ENSEEIHT et directrice opérationnelle de l'association TéSA, ainsi que Jean-Yves Tourneret, professeur à l'INP-ENSEEIHT et responsable de l'équipe Signal et Communication à l'Institut de Recherche en Informatique de Toulouse (IRIT, UMR 5505 du CNRS), pour m'avoir confié ce travail de recherche, ainsi que pour leur aide et leurs précieux conseils au cours de ces années. Au travers de nos discussions, ils m'ont apporté une compréhension plus approfondie des divers aspects du sujet. Je salue aussi la souplesse et l'ouverture d'esprit de mes directeurs de thèse qui ont su me laisser une large marge de liberté pour mener à bien ce travail de recherche.

Denis Kouamé m'a fait l'honneur de présider le jury, qu'il soit remercié pour son intérêt pour mon travail. J'adresse mes sincères remerciements à David Brie et Lotfi Senhadji qui ont accepté de rapporter cette thèse, ainsi qu'à Audrey Girémus et Franz Hlawatsch pour l'avoir examinée. Merci enfin à José Carlos Bermudez et Marie Razani pour avoir accepté de participer à ce jury.

Durant ces trois ans, j'ai eu la chance de bénéficier de l'aide de nombreux chercheurs français ou étrangers, soit lors de visites dans leurs laboratoires, soit lors de discussions plus informelles. Ces collaborations ont permis de concrétiser plusieurs parties de ces travaux de thèse. Parmi les contributeurs à ce travail de thèse, je remercie chaleureusement Georg Kail et Franz Hlawatsch, chercheurs à Vienna University of Technology, pour les échanges fructueux que nous avons entretenus. Je remercie également Audrey Girémus, chercheur à l'Université de Bordeaux, pour son accueil, sa disponibilité et son expertise. Je salue alors Monica Bugallo, chercheur à Stony Brook University à New York, avec qui j'ai eu le plaisir de travailler lors de sa visite à Toulouse. Je tiens à remercier Philippe Maury, cardiologue de l'Hôpital de Toulouse à Rangueil, pour son expertise médicale qui a beaucoup apporté à ce travail. J'aimerais aussi adresser mes remerciements à Marie Razani, directrice scientifique de St. Jude Medical, pour avoir financé la dernière partie de ce travail. J'en profite enfin pour adresser une pensée amicale à Omid Sayadi, chercheur à Harvard Medical School, pour son aide pendant le review de mon premier article journal.

Ensuite je tiens à remercier tous les personnels de l'association TéSA et ceux de l'équipe SC à l'IRIT, David Bonacci, Francis Castanié, Marie Chabert, Martial Coulon, Nicolas Dobigeon, Nathalie Thomas, $\cdots$, sans qui cette thèse ne serait pas ce qu'elle est. Merci également à Sylvie

Armengaud, Sylvie Eichen, Marie-Josée Estepa et Yamina Nedjari, pour leur professionnalisme sans faille.

Je passe ensuite une dédicace spéciale à tous les jeunes que j'ai eu le plaisir de côtoyer durant ces années : Abderrahim, Abdelaziz, Adnan, Anchalee, Bouchra, Cécile, Fabian, Florian, Nesrine, Nil, Raoul, Razvan, Tony, Victor, Yoann, $\cdots$. Les moments magnifiques que nous avons passés ensemble au bureau, dans le couloir, à la cantine, au foyer et sur le terrain de foot, je les garde précieusement au fond de mon coeur.

De plus, mes remerciements seraient incomplets, si je ne fais pas mention de mes parents, Xiansheng Lin et Ruizhi Zhou, qui sont à l'autre bout du continent l'Eurasie et qui m'ont soutenu de tout leur coeur en dépit de notre éloignement. Je leur dois aujourd'hui beaucoup.

Enfin, il y a une personne de très grande importance dans ma vie, qui m'accompagne à mes côtés et qui me soutient tout le temps par la force et l'énergie qui lui sont disponibles. Il s'agit de Xiaoting Li, ma merveilleuse, splendide et vertueuse épouse, dont je ne pourrai mesurer l'apport dans l'accomplissement de cette formation. Dans tout ce que j'ai pu souffrir et obtenir, elle a été l'autre moi. Je te remercie infiniment ma chérie.

Chao.

Abstract

This thesis studies Bayesian estimation/detection algorithms for P and T wave analysis in ECG signals. In this work, different statistical models and associated Bayesian methods are proposed to solve simultaneously the P and T wave delineation task (determination of the positions of the peaks and boundaries of the individual waves) and the waveform-estimation problem. These models take into account appropriate prior distributions for the unknown parameters (wave locations and amplitudes, and waveform coefficients). These prior distributions are combined with the likelihood of the observed data to provide the posterior distribution of the unknown parameters. Due to the complexity of the resulting posterior distributions, Markov chain Monte Carlo algorithms are proposed for (sample-based) detection/estimation. On the other hand, to take full advantage of the sequential nature of the ECG, a dynamic model is proposed under a similar Bayesian framework. Sequential Monte Carlo methods (SMC) are also considered for delineation and waveform estimation. In the last part of the thesis, two Bayesian models introduced in this thesis are adapted to address a specific clinical research problem referred to as T wave alternans (TWA) detection. One of the proposed approaches has served as an efficient analysis tool in the Endocardial T wave Alternans Study (ETWAS) project in collaboration with St. Jude Medical, Inc and Toulouse Rangueil Hospital. This project was devoted to prospectively assess the feasibility of TWA detection in repolarisation on EGM stored in ICD memories.

Résumé

Cette thèse a pour objet l'étude de méthodes Bayésiennes pour l'analyse des ondes P et T des signaux ECG. Différents modèles statistiques et des méthodes Bayésiennes associées sont proposés afin de réaliser la détection des ondes P et T et leur caractérisation (détermination du sommet et des limites des ondes ainsi que l'estimation des formes d'onde). Ces modèles prennent en compte des lois a priori pour les paramètres inconnus (les positions des ondes, les amplitudes et les coefficients de ces formes d'onde) associés aux signaux ECG. Ces lois a priori sont ensuite combinées avec la vraisemblance des données observées pour fournir les lois a posteriori des paramètres inconnus. En raison de la complexité des lois a posteriori obtenues, des méthodes de Monte Carlo par Chaînes de Markov sont proposées pour générer des échantillons distribués asymptotiquement suivant les lois d'intérêt. Ces échantillons sont ensuite utilisés pour approcher les estimateurs Bayésiens classiques (MAP ou MMSE). D'autre part, pour profiter de la nature séquentielle du signal ECG, un modèle dynamique est proposé. Une méthode d'inférence Bayésienne similaire à celle développée précédemment et des méthodes de Monte Carlo séquentielles (SMC) sont ensuite étudiées pour ce modèle dynamique. Dans la dernière partie de ce travail, deux modèles Bayésiens introduits dans cette thèse sont adaptés pour répondre à un sujet de recherche clinique spécifique appelé détection de l'alternance des ondes T. Une des approches proposées a servi comme outil d'analyse dans un projet en collaboration avec St. Jude Medical, Inc et l'hôpital de Rangueil à Toulouse, qui vise à évaluer prospectivement la faisabilité de la détection des alternances des ondes T dans les signaux intracardiaques.

List of Abbreviations and Symbols

APC	atrial premature contraction
A-V	atrio-ventricular
BG	Bernoulli-Gaussian
CADe	computer-aided detection
CADx	computer-aided diagnosis
cpb	cycle-per-beat
DWT	discrete wavelet transform
EA	ensemble averaging
ECG	electrocardiogram
EGM	intracardiac electrogram
EKF	extended Kalman filter
EM	electrode motion
EMG	electromyograph
ETWAS	Endocardial T Wave Alternans Study
HMM	hidden Markov model
ICD	implantable cardioverter defibrillator
iid	independent and identically distributed
IS	importance sampling
KF	Kalman filters
KS	Kolmogorov-Smirnov
LDA	linear discriminant analysis
LPD	low-pass differentiation
MA	muscular activity
MAP	maximum a posteriori
MCMC	Markov chain Monte Carlo
MITBIH	MIT-BIH Arrhythmia database
MMA	modified moving average
MMSE	minimum mean square error
MPSRF	multivariate potential scale reduction factor
MPF	marginalized particle filter
NMSE	normalized mean square error
OCT	optical coherence tomography
PCGS	partially collapsed Gibbs sampler
pdf	probability density function
PF	particle filter
PFA	probability of false alarm
PVC	premature ventricular contractions
QTDB	QT database

ROC	receiver operating characteristic
SA	sino atrial
SM	spectral method
SMC	sequential Monte Carlo
SNR	signal to noise ratio
SPF	standard particle filter
ST	statistical test
STFT	short time Fourier transform
TWA	T wave alternans
TWAR	T wave alternans ratio
VF	ventricular fibrillation
VPC	ventricular premature contraction
VT	ventricular tachycardia
WT	wavelet transform

Contents

List of Figures

List of Tables

List of Algorithms

General Introduction

The human heart, a magical muscular pump which is typically about the same size as a fist, circulates blood through a human's body all along their life. Despite the tremendous efforts that have been dedicated to the cardiology study, the heart has not yet revealed all its secrets. According to a fact sheet released by the World Health Organization in September 2011, cardiovascular diseases are still the number one cause of death globally [WHO11].

The electrocardiogram (ECG), a graphical representation of temporal differences in potential electrical forces that lead to the cardiac muscle contraction, has been recognized as the simplest non-invasive method of collecting electrical signal from the beating heart that provides extremely useful information to the doctors. It all started when Augustus D. Waller published the first human ECG recorded with a capillary electrometer in 1887. Subsequently, Willem Einthoven, who received the Nobel Prize in 1924 for his development of the first lead system for ECG recording, identified the five deflection points in the cardiac cycle by naming them P, Q, R, S and T which are still being used in the present standards (see Fig. 1.7). Since then a huge knowledge base has been generated covering clinical and engineering aspects of electrocardiography. ECG analysis has become a routine part of any complete medical evaluation.

In the last few decades electronic recorders have been developed for digital recording of the ECG signal and they are available in such a compact form that the user can wear them for ECG recording without much of obstruction in the routine activities. Since a huge volume of ECG data is generated by the wearable ECG recorders, automated methods are preferred for analysis of the ECG signal. On the other hand, increasing comfort usually results in signals with reduced quality. For instance, electrodes that are incorporated in garments generally provide signals with a lower signal-to-noise ratio (SNR) and more artifacts than contact electrodes directly glued to the body [GHM07]. Therefore, the development of efficient and robust signal processing methods for ECG computer-aided detection (CADe) and computer-aided diagnosis (CADx) is a subject of major importance.

Most of the clinically useful information in ECGs can be obtained from the intervals, amplitudes, and wave shapes (morphologies). The QRS complex is the most characteristic waveform of the ECG signal. Its high amplitude makes QRS detection easier than other waves. Thus, it is generally used as a reference within the cardiac cycle. Algorithms for P and T waves detection and delineation (determination of peaks and limits of the individual P and T waves) usually begin with QRS detection. Search windows are then defined before and after the QRS location to seek for the other waves. Finally, an appropriate strategy is used to enhance the distinctive features of each wave in order to find its peaks and limits. Because of the low slope and magnitude of P and T waves, as well as the presence of noise, interference, and baseline

fluctuation, P and T wave estimation and detection remain a difficult task. It is also worthwhile to note that there is not any universally acknowledged clear rule to locate the beginning and the end of wave components, which complicates the P and T wave delineation. Furthermore, in addition to the estimation of wave peaks and limits, an accurate waveform estimation is certainly relevant for some medical diagnoses (such as T-wave alternans (TWA) detection [MO04]) or pathology analysis (such as arrhythmia detection [EG96]).

Bayesian interpretation of observations is a powerful method of relating measurements to their underlying sources. Bayesian models are well suited to the electrophysiological environment, allowing a direct and natural way to express what is known (and unknown) and to evaluate which one of many alternatives is most likely the source of the observations. For two centuries, Bayesian inference moved rapidly but unevenly from the domain of mathematical statistics into biomedical applications. In part that was because Bayesian interpretation often required too many steps to be feasible with hand calculation in real applications. As computer power became widespread in the last few years, use of Bayesian models is now growing rapidly in electrophysiology [BNP10].

The subject of this thesis is to study Bayesian estimation/detection algorithms suitable for P and T wave analysis in ECG signals. This problem is of great interest since it allows to express the relationships between unknown parameters established or limited by physiology or medicine in a probabilistic sense and "get it into the problem". Doing so is, in general, rewarded by improved performance in estimating the unknown parameters from the measurement data. In this work, different statistical models and associated Bayesian methods are proposed to solve simultaneously the P and T wave delineation task (determination of the positions of the peaks and boundaries of the individual waves) and the waveform-estimation problem. These models take into account appropriate prior distributions for the unknown parameters (wave locations and amplitudes, and waveform coefficients). These prior distributions are combined with the likelihood of the observed data to provide the posterior distribution of the unknown parameters. The posterior distribution depends on hyperparameters that can be fixed *a priori* or estimated from the observed data. The usual Bayesian estimators related to the posterior distribution are the minimum mean square error (MMSE) estimator and the maximum a posteriori (MAP) estimator [Van68]. Nevertheless, the posterior distributions of the parameters are usually too complex to obtain analytical expressions of the classic Bayesian estimator (MMSE or MAP). To alleviate numerical problems related to the posterior distributions associated to the P and T wave analysis problem, we first propose to resort to Markov chain Monte Carlo (MCMC) methods [RC04]. MCMC is a powerful sampling strategy, appropriate to solve complex Bayesian inference problems. It allows to generate samples asymptotically distributed according to the target distribution. The MMSE or MAP estimators of the unknown parameters are then computed using the generated samples. On the other hand, to take full advantage of the sequential nature of the ECG, sequential Monte Carlo (SMC) methods [DdFG01] are also investigated in this thesis under the same Bayesian inference. SMC allows one to update the posterior distribution as data become available, thus it allows to make the Bayesian models suitable for real-time ECG analysis. An additional benefit of sequential methods is their lower memory requirement compared to the window based methods since the data are dealt in a sequential manner.

The organization of this thesis is described in the following. Chapter 1 introduces the basis

of the cardiovascular system, electrocardiography and some physiology notions. A literature review of ECG processing methods is also presented. In Chapter 2, a multiple-beat processing window based Bayesian model for P and T wave delineation and waveform estimation is studied. Two Gibbs-type sampling methods (a partially collapsed Gibbs sampler and a block Gibbs sampler) are proposed to generate samples distributed according to the posterior of the proposed Bayesian model. The generated samples are used to estimate the unknown model parameters and hyperparameters. The wave detection and delineation criteria based on the posterior distributions are also presented. Simulation results performed on the standard annotated QT database (QTDB) [LMGM97] as well as a comparison with other state-of-art methods are also given. In the purpose of making the model more suitable for real-time applications, a beat-to-beat Bayesian model for P and T wave delineation is proposed in Chapter 3. Gibbs sampling methods and sequential Monte Carlo method (particle filters) are studied to estimate the unknown parameters of the beat-to-beat model. A comparison with the window based approach as well as other alternative methods on QTDB is reported. In Chapter 4, the window based Bayesian model and the Gibbs sampling methods are first adapted to deal with TWA detection in surface ECGs. Then, the beat-to-beat approach is applied on real intracardiac electrograms (EGMs) provided by St. Jude Medical, Inc to deal with endocardial TWA analysis.

The main contributions of this thesis are summarized below.

- **Chapter 1.** The physiological basis of the ECG is briefly introduced as four aspects: (1) physiology of the specific structures of the heart, (2) electrophysiology of the heart and the origin of the ECG, (3) ECG measurement and registration, (4) ECG interpretation in clinical context. Based on the initiative of ECG computer-aided detection and computer-aided diagnosis, a brief review of the ECG signal processing methods proposed in the literature is given, with the emphasis on P and T wave detection and delineation.

- **Chapter 2.** We introduce a new Bayesian model based on a *multiple-beat processing window* which simultaneously solves the P and T wave delineation and the waveform estimation problems. This model is based on a modified Bernoulli-Gaussian sequence with minimum distance constraint [KTHD12] for the wave locations and appropriate priors for the amplitudes, wave impulse responses and noise variance. A recently proposed partially collapsed Gibbs sampler which exploits this minimum distance constraint is adapted to the proposed model to estimate the unknown parameters [LMT10]. Then, a modified version of this Bayesian model is proposed to consider the baseline within each non-QRS component and to represent P and T waves by their respective dimensionality reducing expansion according to Hermite basis functions. The local dependency of the ECG signal is expressed by a block constraint. To alleviate numerical problems related to the modified Bayesian model, a block Gibbs sampler is studied [LKT+11, LTM+11]. The proposed PCGS and block Gibbs sampler overcome the slow convergence problem encountered with the classical Gibbs sampler. The resulting algorithms are validated using the entire annotated QT database. A comparison with other benchmark methods showed that the proposed Bayesian methods provide a reliable detection and an accurate delineation for a wide variety of wave morphologies. In addition, the proposed Bayesian methods can provide accurate waveform estimation and allow for the determination of confidence intervals which indicate reliability information about the estimates.

- **Chapter 3.** This chapter introduces a modified Bayesian model that enables P and T wave delineation and waveform estimation on a beat-to-beat basis. The *beat-to-beat* 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. Compared to the window based model which relies on a non-overlapped multiple-beat processing window to estimate the waveforms, the beat-to-beat model is more suitable to the pseudo-stationary nature of the ECG signal and to real time applications. The block Gibbs sampler is used to estimate the parameters of the resulting Bayesian model [LKT+12]. Then, in order to consider all the available former beats instead of using the last beat only, a SMC method is studied. Following the SMC analysis principle, the sequential nature of the ECG is exploited by using a dynamic model under a similar Bayesian framework. Particle filters (PFs) are then proposed to resolve the unknown parameters of the dynamic model [LBMT11, LGMT12].

- **Chapter 4.** The Bayesian model introduced in Chapter 2 is modified to account for a possible distinction between odd and even beats since a difference between successive T waves is a sign of a potential cardiac risk. The modified Bayesian model is then used to perform surface ECG T-wave alternans (TWA) analysis. The odd and even T-wave amplitudes generated by the block Gibbs sampler are used to build statistical tests for TWA detection [LMT11]. Furthermore, the beat-to-beat approach is adapted to deal with TWA analysis in an endocardial context that has special constraints compared to surface ECGs. The proposed strategies are validated on both synthetic data and and real EGMs provided by ST. JUDE MEDICAL, INC in collaboration with a cardiologist from the University Hospital Centre of Toulouse.

Personal Publication List

Journal publications

[LKT+12] C. Lin, G. Kail, J.-Y. Tourneret, C. Mailhes, and F. Hlawatsch, "Endocardial T wave alternans detection using a beat-to-beat Bayesian approach and a block Gibbs sampler," *IEEE Trans. Biomed. Eng.*, 2012, to be submitted.

[LMT10] C. Lin, C. Mailhes, and 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.*, vol. 57, no. 12, pp. 2840–2849, Dec. 2010.

Conference publications

[LBMT11] C. Lin, M. Bugallo, C. Mailhes, and J.-Y. Tourneret, "ECG denoising using a dynamical model and a marginalized particle filter," in *Proc. IEEE Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2011.

[LGMT12] C. Lin, A. Giremus, C. Mailhes, and J.-Y. Tourneret, "Beat-to-beat P and T wave delineation and waveform estimation in ECG signals using a marginalized particle filter," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Bucharest, Romania, Aug. 2012, to appear.

[LKT+11] C. Lin, G. Kail, J.-Y. Tourneret, C. Mailhes, and F. Hlawatsch, "P and T wave delineation and waveform estimation in ECG signals using a block Gibbs sampler," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 537–540.

[LMT11] C. Lin, C. Mailhes, and J.-Y. Tourneret, "T-wave alternans detection using a Bayesian approach and a Gibbs sampler," in *Proc. Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc. (EMBC)*, Boston, MA, Aug. 2011, pp. 5868–5871.

[LTM+11] C. Lin, J.-Y. Tourneret, C. Mailhes, G. Kail, and F. Hlawatsch, "Segmentation des signaux ECG et caractérisation des ondes P et T à l'aide d'un échantillonneur de Gibbs par bloc," in *Actes du XXIIIième Colloque GRETSI*, Bordeaux, France, Sep. 2011, in french.

Chapter 1

Introduction to Cardiac Electrophysiology

Contents

1.1 Introduction

Before attempting any ECG signal processing, it is important to first understand the physiological basis of the ECG, to review measurement conventions of the standard ECG, and to review how a clinician uses ECGs for patient care. Understanding the basis of a normal ECG requires appreciation of four aspects:

- physiology of the specific structures of the heart,
- electrophysiology of the heat and the origin of the ECG,
- ECG measurement and registration,
- ECG interpretation in clinical context.

Around these four topics, we will discuss in the first two sections how the wave of electrical current propagates through the heart muscle, the physiology of the specific structures of the heart through which the electrical wave travels, how that leads to a measurable signal on the

surface of the body, producing the normal ECG, and last how a cardiologist interprets the ECG. It should be noted that, through decades of investigation, much detail is available about the electrophysiologic activity of the heart and the introduction in this thesis is therefore only a highly abbreviated summary. Some of the material and figures in this chapter are taken from [Mar90, CAM06, CR90, CPD09], to which the reader is referred for a more detailed overview of this subject.

As mentioned in the general introduction, the detection/estimation and delineation of each wave component of the ECG signal is a crucial step for ECG interpretation. In the last section of this chapter, a brief review of ECG signal processing methods proposed in the literature is given, with the emphasis on wave detection and delineation.

1.2 Cardiology basis

The heart is a muscle that is rhythmically driven to contract and hence drive the circulation of blood throughout the body. Before every normal heartbeat, or *systole*, a wave of electrical current passes through the entire heart, which triggers myocardial contraction. The pattern of electrical propagation is not random, but spreads over the structure of the heart in a coordinated pattern which leads to an effective, coordinated systole. This results in a measurable change in potential difference on the body surface of the subject. The resultant amplified (and filtered) signal is known as an ECG. To understand the origin of the ECG, one should begin with the heart structure and its functioning.

1.2.1 Heart anatomy

The heart consists of four chambers (see Fig. 1.1: two atrial chambers (*atrium*) in the upper heart and two ventricular chambers (*ventricles*) in the lower part.). The atrium are the receiving chambers and the ventricles are the discharging chambers. The heart is mainly constituted by a specialized type of striated muscle, with properties that functions somewhat differently than skeletal muscle. Heart cells are connected by gap junctions that allow ions to flow from one cell to another, allowing for rapid spreading of depolarization. A collection of heart cells connected in this way constitutes a syncytium. The heart is composed of two syncytiums: an atrial syncytium and a ventricular syncytium.

The muscle, which is like a wall around the heart, is called the *myocardium*. The two ventricles are separated by the ventricular septum, which is thick muscular toward the bottom, and thick fibrous at the top. Under the myocardium is the *endocardium*, which lines the chambers of the heart, and which is mainly composed of epithelial cells. The *pericardium* is a double-walled sac that contains the heart and the roots of the great vessels. The outer wall of the pericardium is composed of dense connective tissue. The portion of the inner wall of the pericardium that is in contact with the heart (but not in contact with the great vessels) is the epicardium. The epicardium contains the blood vessels supplying blood to heart muscle (the coronary arteries) as well as nerve fibers.

The veins of the body terminate in two great vessels that empty into the right atrium: the

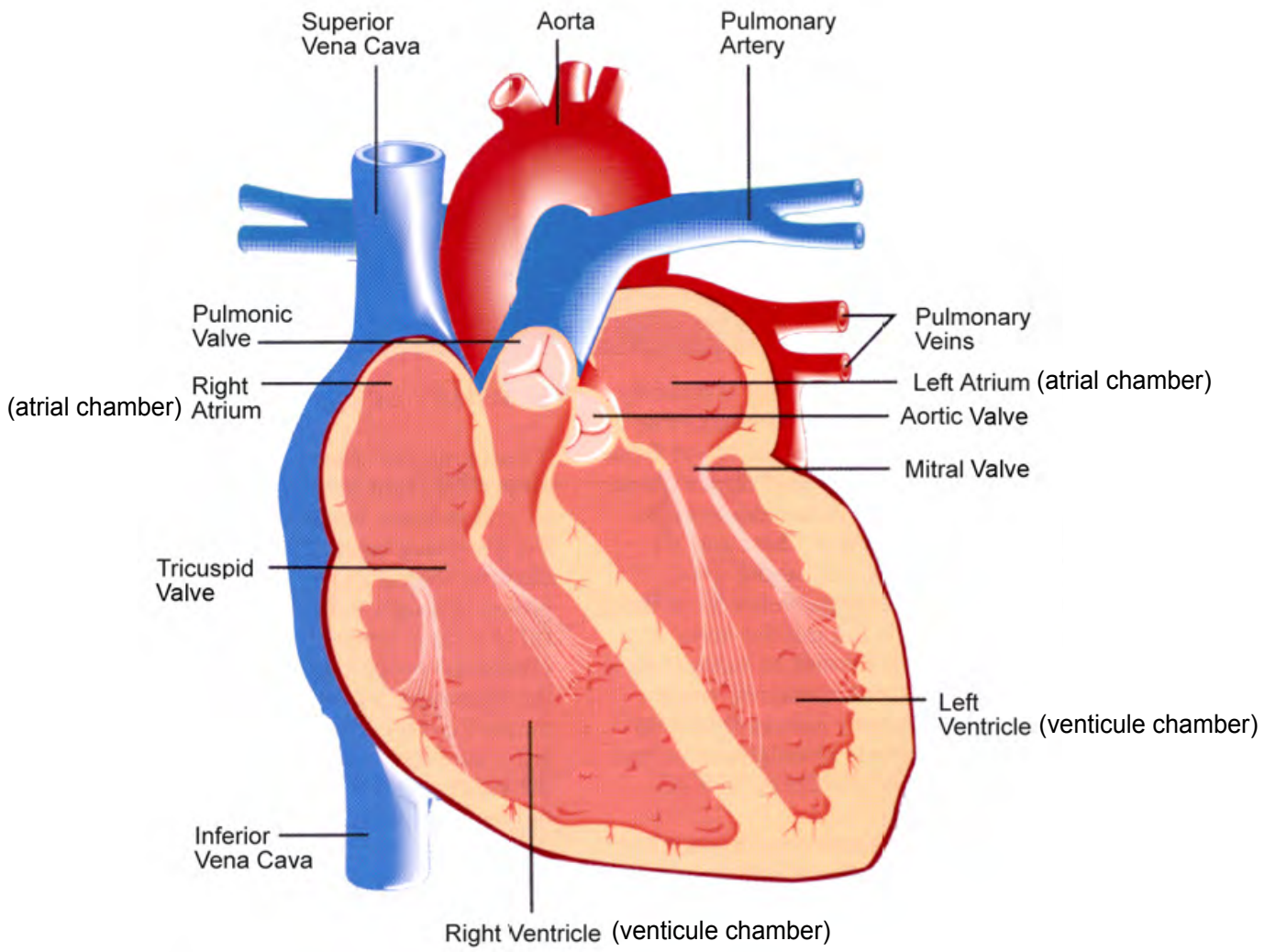


Figure 1.1: Structure diagram of the human heart. Image adapted from [CR90].

superior vena cava (from the upper body) and the *inferior vena cava* (from the lower body). As can be seen in Fig. 1.1 and Fig. 1.2, blood exits the heart through the *pulmonary artery* (which takes unoxygenated blood to the lungs from the right ventricle) and through the *aorta* (which distributes oxygenated blood to the body from the left ventricle). Oxygenated blood from the lungs enters the left atrium from the *pulmonary vein*. The pulmonary artery is the only artery that carries unoxygenated blood and the pulmonary vein is the only vein that carries oxygenated blood.

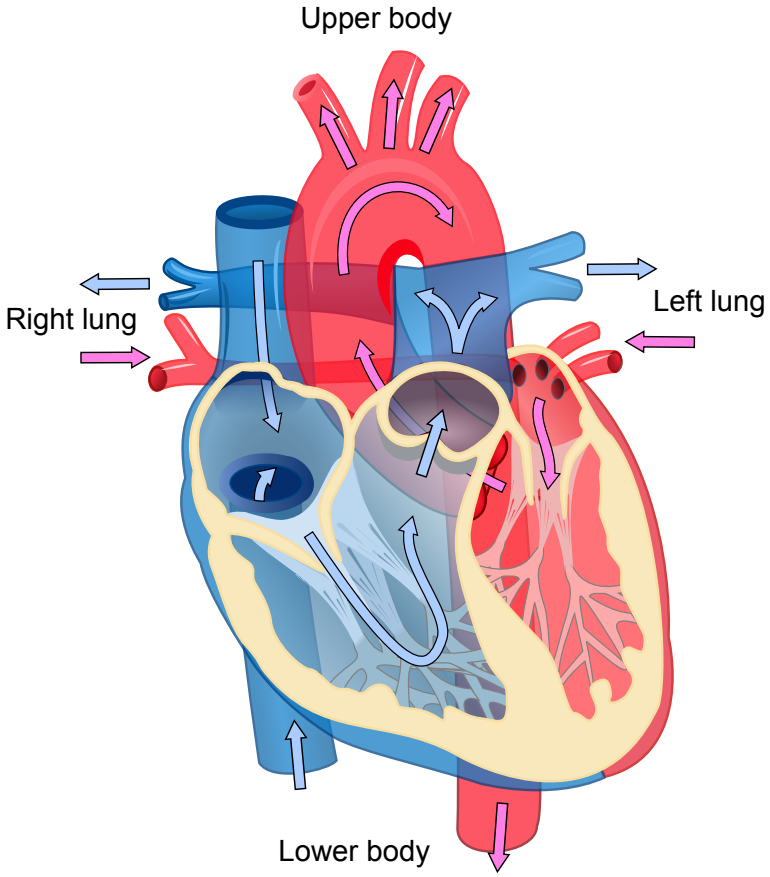


Figure 1.2: Blood flow diagram of the human heart. Blue components indicate de-oxygenated blood pathways and red components indicate oxygenated pathways. (Image adapted from [CR90])

There is a one-way flow of blood from the atria to the ventricles. Reverse flow is prevented by the atrio-ventricular (A-V) valves: the *tricuspid valve* (having 3 flaps or "cusps") on the right and the *mitral valve* on the left. The inner flaps of the A-V valves have connective tissue attachments ("heart strings") from muscular mounds on the ventricles that prevent the flaps from bulging too much into the atria during contraction of the ventricles. Flow of blood back to the ventricles following ejection of blood from the ventricles is prevented by the semilunar valves: the *pulmonary valve* on the right and the *aortic valve* on the left, each of which have three flaps shaped somewhat like half-moons.

The openings to the right and left coronary arteries (which supply blood to the heart) are

in the aorta just beyond the cusps of the aortic semilunar valve. The right coronary artery supplies the right atrium, right ventricle, and most of the posterior surface of the left ventricle with blood. The left coronary artery branches into the left anterior descending artery (which supplies the anterior surface of the left ventricle) and the circumflex artery (which supplies the left atrium and parts of the left ventricle)

1.2.2 Heart operation

The heart pumps blood in a two-step process. The atria pump blood into the ventricles and then the ventricles pump blood out of the heart. Chambers of the heart fill with blood during a relaxation phase (diastole) and eject blood during a contraction phase (systole). Atrial systole precedes ventricular systole. The systole and diastole of the left ventricle correspond to the *systolic* (high pressure) and *diastolic* (lower pressure) phases of blood pressure in the body. As shown in Fig. 1.2, the right atrium pumps blood from the body into the right ventricle. The right ventricle pumps blood into the lungs. The left atrium pumps blood from the lungs into the left ventricle. The left ventricle pumps blood to the body through the aorta. This cycling of blood by the heart is called the *cardiac cycle*.

The left ventricle has the most powerful musculature insofar as it must pump blood into the entire body. For the left ventricle the cardiac cycle can be described in four phases:

1. **Period of filling**. Volume increases from about 50 milliliters of blood to about 120 milliliters of blood, after which the mitral valve closes. 70% of ventricular filling occurs passively as blood flows into the atria and through the ventricles while the tricuspid valve is open. Atrial contraction pushes the final 30% of blood into the ventricle, after which the tricuspid valve closes.

2. **Period of isovolumetric contraction**. All valves are closed as pressure rises from 5 mmHg to 80 mmHg while blood volume remains unchanged.

3. **Period of ejection**. The aortic valve opens and blood is ejected into the body. 70% of blood emptying occurs in the first third of the ejection period as pressure rises to 120 mmHg. Pressure drops to 100 mmHg during the last two-thirds of the ejection period. (Stroke volume can double during exercise.)

4. **Period of isovolumetric relaxation**. Pressure drops to zero after the aortic valve has closed.

The heart electrical conduction system is presented in Fig. 1.3. Depolarization and contraction of cardiac muscle is preceded by depolarization of the cardiac nerve bundles. Depolarization of the heart is initiated at the *Sino Atrial node* (SA node) located on the right atrium. Depolarization is transmitted through the *Atrio-Ventricular node* (A-V node), the *His Bundle*, and then through the ventricles by the *Purkinje fibers*.

The cardiac cycle is initiated in the SA node in the atrium, called *atrial depolarization*, which causes contraction of the atrial syncytium. Unlike cardiac cells which have mechanical

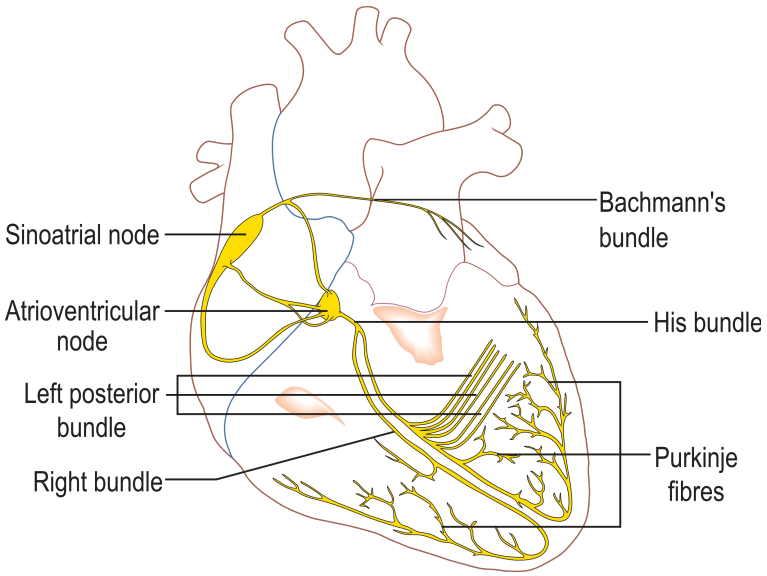


Figure 1.3: Heart electrical conduction system. Image adapted from [hea].

(contractile) function, the cardiac cells having an electrical conduction function not only conduct electrical impulses, but spontaneously generate impulses. In a normal heart, the cardiac cells of the sinoatrial node have the highest rate of spontaneous impulse generation, which makes the sinoatrial node the ultimate pacemaker of the electrical system of a healthy heart. Next, the depolarization is delayed at the A-V node before passing to the His bundle and splitting into the bundle branches. The A-V node delay, during which time there is no measurable electrical activity at the body surface, prevents the atria and ventricles from contracting at the same time. After activity emerges from the A-V node it sends depolarization to the muscle of the right ventricle and the left ventricle, respectively. Next, there is the *septal depolarization*. The septum is the wall between the ventricles, and a major bundle of conducting fibers runs along the left side of the septum. Then comes the *apical depolarization*, and the wave of depolarization moving left is balanced by the wave moving right. Note that in *left ventricular depolarization* and *late left ventricular depolarization*, there is also electrical activity in the right ventricle, but since the left ventricle is much more massive its activity dominates. After the ventricular depolarization, there is a plateau period during which the myocardium has depolarized (ventricles depolarized) where no action potential propagates. Finally, the individual cells begin to repolarize and another wave of charge passes through the heart, this time originating from the dipoles generated at the interface of depolarized and repolarizing tissue (i.e., *ventricular repolarization*). The heart then returns to its resting state (such that the ventricles are repolarized), awaiting another electrical stimulus that starts the cycle anew.

1.3 Electrocardiography

The heart electrical conduction can be projected onto different lines of well defined orientation on the skin surface. The lead in electrocardiography is a line that connects two observation

points of the electrical activity of the heart from which to measure a difference electric potential. Each lead reveals the magnitude of the electrical conduction in the direction of that lead at each instant of time. In this section, a description of how the electrical conduction of a cardiac cycle results in the pattern of a normal scalar ECG is described, and the clinical interpretation of the ECG signal is presented.

1.3.1 ECG measurement

ECG monitoring machines generally can register several potential differences at the same time according the location and number of electrodes located on the body. Additional electrodes provide more sensitive monitoring of depolarization. Each measurement of the potential corresponds to a lead of the ECG. Note that the term "lead" can be confusing insofar as the standard limb leads are bipolar (measure the potential difference between two electrodes) whereas the others are unipolar (single electrode). The 12-lead ECG, which contains six front and six precordial leads, has been standardized by an international convention. The different vantage points offered by the 12-lead ECG allow for a three-dimensional representation of the electrical activity of the heart. The standard 12-lead ECG system is composed as follows:

- 3 Einthoven bipolar limb leads: I, II and III,
- 3 unipolar augmented limb leads: aVR, aVL and aVF,
- 6 unipolar precordial chest leads: V1, V2, V3, V4, V5 and V6.

In the following paragraphs, we present the electrode positions corresponding to each type of lead.

Einthoven leads

Back to the beginning of the 20th century, Willem Einthoven discovered that depolarization of the heart can be monitored with three electrodes attached to the body: one on each arm and one on the left leg. Depolarization spreads from the right atrium (corresponding to the electrode on the right arm) to the left ventricle (corresponding to the electrode on the left leg). The electrode on the left arm is neutral. Based on this principle, he invented a three-lead system of the ECG recording–the standard limb leads–to which has been added three augmented limb leads and six chest leads in the modern 12-lead ECG. The standard limb leads are:

- Lead I: potential difference between the left arm (positive electrode) and the right arm (negative electrode),
- Lead II: potential difference between the left leg (positive electrode) and the right arm (negative electrode),
- Lead III: potential difference between the left leg (positive electrode) and the left arm (negative electrode).

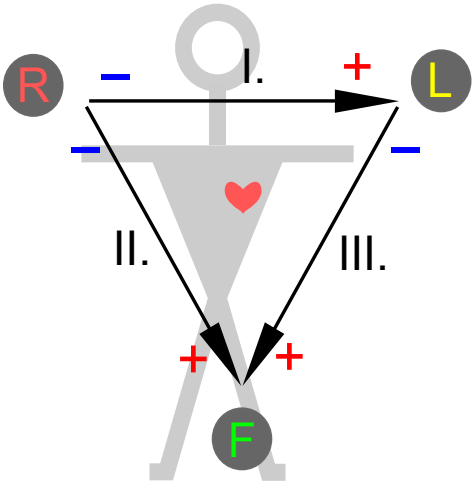


Figure 1.4: The Einthoven's triangle. Image adapted from [CAM06].

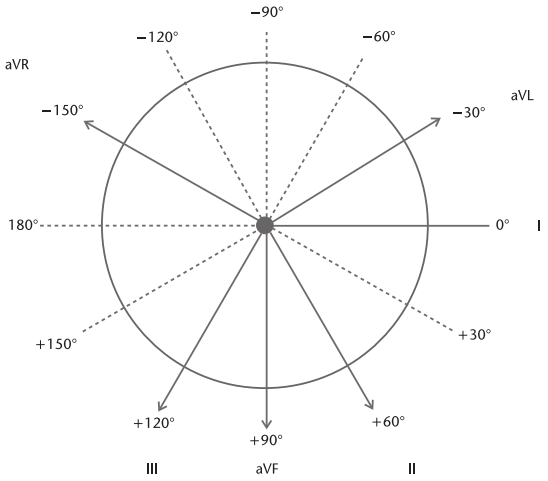


Figure 1.5: Frontal plane limb leads. Image adapted from [CAM06].

Note that the left arm is positive relative to the right arm and negative relative to the left leg. There is redundancy in using these three bipolar leads insofar as the voltage recorded at Lead II is the sum of the voltages of the other two leads (I + III = II). A positive waveform (positive voltage) is recorded when depolarization flows from a negative electrode to a positive electrode. For this reason, Lead II typically gives the canonical ECG patterns because the flow of depolarization from negative electrode (right arm) to positive electrode (left leg) matches the typical alignment of the heart in the chest: sloping downward toward the left side of the body.

Augmented leads

Leads I, II, and III give a view of the heart called *Einthoven's triangle* (see Fig. 1.4), that divides a 360° circle centered on the heart into three depolarization axes separated by 120°. Adding the three augmented Goldberger limb leads divides the heart into six depolarization axes separated by 60° (as illustrated in Fig. 1.5). The augmented limb leads are:

- aVR : potential difference between the right arm (positive electrode) and the center of the heart,
- aVL : potential difference between the left arm (positive electrode) and the center of the heart,
- aVF : potential difference between the left leg (positive electrode) and the center of the heart.

Here "a" denotes augmented, "V" denotes Voltage, and the final letter designates right or left arm, or the left leg. Including both positive and negative directions of depolarization of both

the standard and augmented limb leads gives a 360° circle divided into twelve 30° sections that is called the *hexaxial reference system*. Lead I (pointing toward the left arm) is taken as the reference for 0° lead orientation, with increasing positive values in a clockwise direction to 180° and increasing negative values in a counterclockwise direction such that −180° is the same orientation as +180°. Although Lead II, with an orientation of about 60° is the most typical orientation of an axis pointing from atria to ventricles, it is not abnormal for heart orientation to be in the range of −30° to +110° which can make Leads aVL, aVF, or III rather than Lead II the primary indicator of normal cardiac depolarization.

Precordial leads

The 360° hexaxial reference system defines what is called the frontal plane of ECG monitoring - a view of the heart from the front. To provide a view through the horizontal plane - viewing the heart as if the body was sliced horizontally through the middle of the chest - six additional "unipolar" precordial (chest) leads (designated V1 to V6) are placed on positions of the left rib cage forming a quarter-circle around the heart (see Fig. 1.6). The ECG represents the difference between each of these electrodes (V1-V6) and the central terminal.

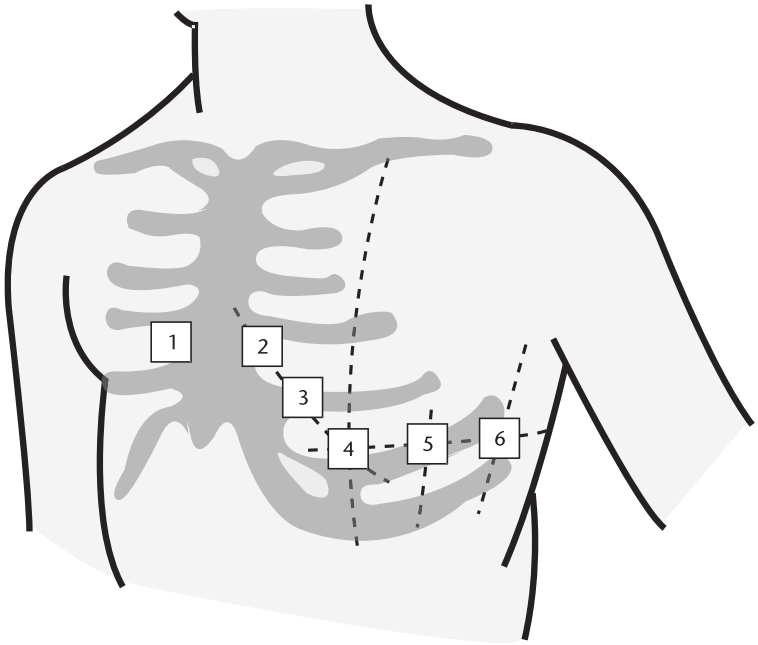


Figure 1.6: The six standard precordial leads. Image adapted from [CAM06].

1.3.2 ECG interpretation

Of all the 12 leads, the first six are derived from the same three measurement points. Therefore, any two of these six leads include exactly the same information as the other four. Over 90% of the heart's electric activity can be explained with a dipole source model. To evaluate this

dipole, it is sufficient to measure its three independent components. In principle, two of the limb leads (I, II, III) could reflect the frontal plane components, whereas one precordial lead could be chosen for the anterior-posterior component. The combination should be sufficient to describe completely the electric heart vector. (The lead V2 would be a very good precordial lead choice since it is directed closest to the x axis. It is roughly orthogonal to the frontal plane.) The main reason for recording all 12 leads is that it enhances pattern recognition. This combination of leads gives the clinician an opportunity to compare the projections of the resultant vectors in two orthogonal planes and at different angles. The useful information that can be extracted from the ECG signal as well as the clinical ECG interpretation procedure are presented in the following paragraphs.

P, QRS and T waves

As explained in 1.2.2, in healthy subjects, each heart cycle consists of the same depolarization / repolarization phase from the atria to the ventricles. The surface projection of the heart electrical action potential is thus a pseudo-periodic signal in the sense that the cardiac cycle repeats according to heart rate. In each cycle, the locations of different waves on the ECG are arbitrarily marked by the letters P, Q, R, S, T and U. (see Fig. 1.7) Note that the baseline voltage of the electrocardiogram is known as the isoelectric line (which is not represented in Fig. 1.7). Typically the isoelectric line is measured as the portion of the tracing following the T wave and preceding the next P wave. Here are the definitions of the ECG waves:

- **P wave** is related to the atrial depolarization. The main electrical vector is directed from the sinoatrial node toward the A-V node, and spreads from the right atrium to the left atrium. The P wave generally has small wave amplitude, with rounded, sometimes biphasic waveform.

- **QRS complex** reflects the rapid depolarization of the right and left ventricles. They have a large muscle mass compared to the atria and so the QRS complex usually has a much larger amplitude than the P and the T waves. By definition, the Q wave is the first negative wave, R wave is the first positive wave of the complex and the S wave is the first negative wave after the R wave.

- **T wave** represents the repolarization of the ventricles. This wave follows the QRS complex after returning to the isoelectric line. The interval from the beginning of the QRS complex to the apex of the T wave is referred to as the absolute refractory period.

- **U wave** is hypothesized to be caused by the repolarization of the interventricular septum. It normally has a low amplitude, and even more is often completely absent. It always follows the T wave and also follows the same direction in amplitude.

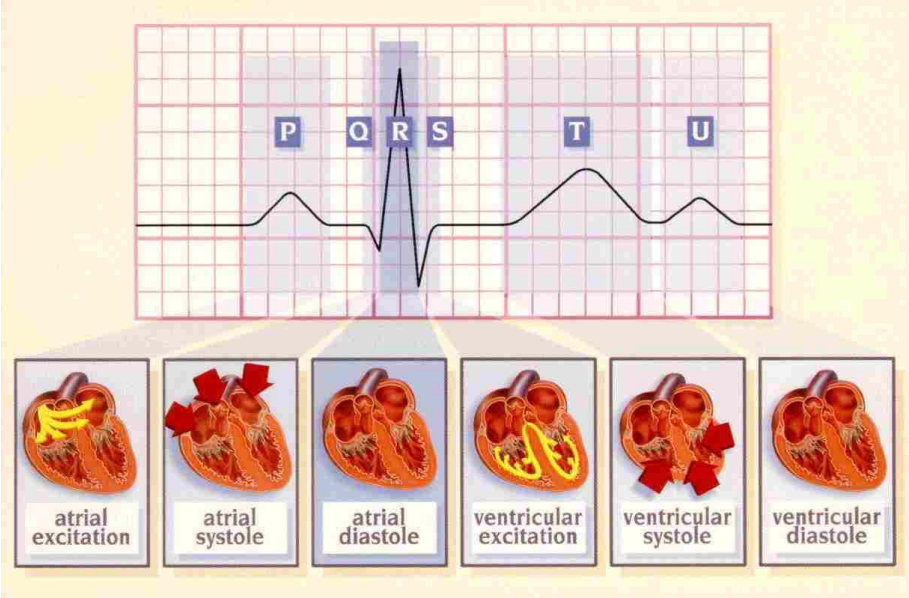


Figure 1.7: The P, Q, R, S, T, U waves resulting from one depolarization / repolarization cycle. Image adapted from [Ala].

ECG intervals

Intervals and segments of the ECG are important parameters for assessing the normality or abnormality of the space between two electrical events, thus they are generally clinically relevant. Fig. 1.8 illustrates the normal clinical features of the electrocardiogram, which include wave amplitudes and inter-wave timings (ECG intervals). Note that the inter-beat timing (RR interval) is not marked. The illustration uses the typical graph-paper presentation format, which stems from the early clinical years of electrocardiography, where analysis was done by hand measurements of hard copies. Each box is 1 mm^2 and the ECG paper is usually set to move at 25 mm/s. Therefore, each box represents 0.04 second in time. The amplitude scale is set to be 0.1 mV per square, although there is often a larger grid overlaid at every five squares (0.20 $second$/0.5 mV) The values for the clinical features indicated on the graph in Fig. 1.8 are typical, although they can vary based upon gender, age, activity, and health.

The clinical interpretation of the ECG segments and intervals can be summarized as following:

- **PR (or PQ) interval** is measured from the beginning of the P wave to the beginning of the QRS complex. The PR interval reflects the time the electrical impulse takes to travel from the sinoatrial node through the A-V node and entering the ventricles. The PR interval is therefore a good estimate of A-V node function.
- **PR segment** connects the P wave and the QRS complex. This coincides with the electrical conduction from the A-V node to the His bundle and the bundle branches and then to the Purkinje Fibers. This electrical activity does not produce a contraction directly and is merely traveling down towards the ventricles and this shows up flat on the

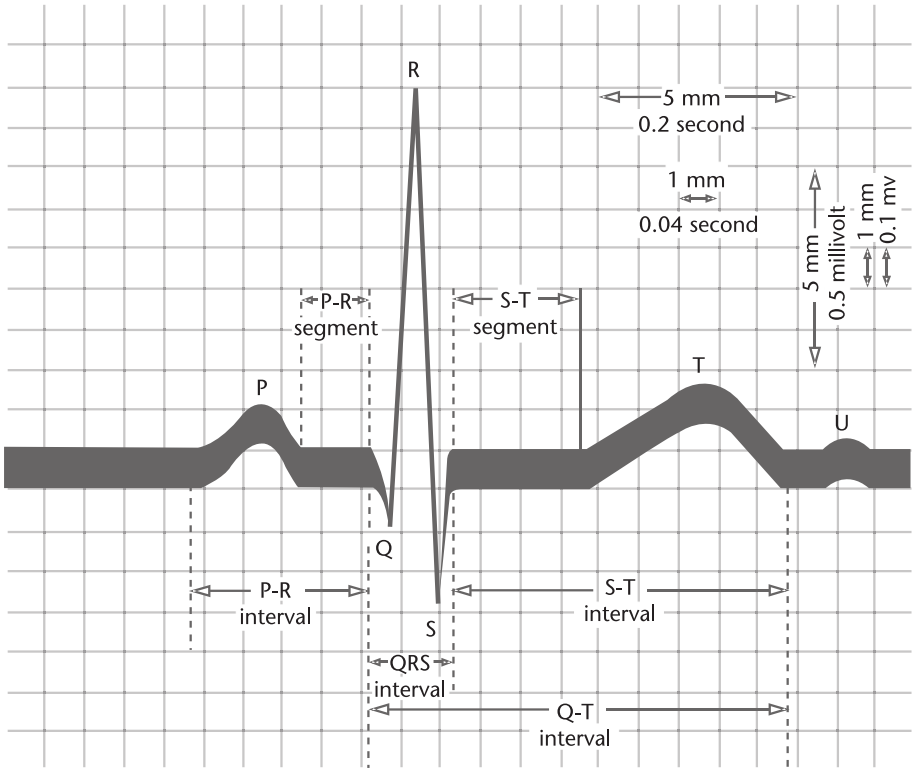


Figure 1.8: Normal features and the intervals of the ECG. Image adapted from [CAM06].

ECG.

- **QT interval** is measured from the beginning of the QRS complex to the end of the T wave. It corresponds to the time of ventricular systole from the beginning of the excitation of the ventricles to the end of their relaxation. A prolonged QT interval is a risk factor for ventricular tachyarrhythmias and sudden death.
- **ST segment** connects the QRS complex and the T wave. The ST segment represents the period when the ventricles are depolarized. Since there is *a priori* no electrical propagation, the segment is isoelectric.
- **ST interval** is measured from the end of the QRS complex (J point) to the end of the T wave.
- **RR interval** is measured between two peaks of successive R waves and represents the cycle of ventricular repolarization. It is associated with the cardiac period.

ECG abnormal patterns

The clinician who uses the electrocardiogram as a diagnostic test wishes to determine cardiac abnormalities from the body surface potentials. As a rough framework, it is worth thinking of the heart as three separate systems: a functional electrical system, a functional system of cardiac arteries to channel nourishing blood to every cell of the myocardium, and a culmination in an effective mechanical pump. To a first approximation, electrical problems come in two forms: those which make the heart pump too slowly or infrequently (*bradycardias*), and those with make the heart pump too quickly (*tachycardias*). If the pumping is too slow, the cardiac output of life-sustaining blood can be dangerously low. If too quick, the cardiac output can also be too low since the heart does not have time to fill, and also because the heart can suffer damage (e.g., *demand ischemia*) when it tries to pump too rapidly. In this paragraph, several ECG abnormal patterns are showed.

Sinus rhythm One class of heart rate abnormalities arises from abnormal function of the control system for heart rate. As discussed in Section 1.2.2, there are specialized cells in the SA node whose function is to act as the heart's pacemaker, rhythmically generating action potentials and triggering depolarization for the rest of the heart (recall that once any portion of the heart depolarizes, the wavefront tends to propagate throughout the entire myocardium). The SA node has an intrinsic rate of firing, but ordinarily this is modified by the *central nervous system*, specifically, the autonomic nervous system.

When the heart rate is controlled by the SA node's rate of firing, the sequence of beats is known as a sinus rhythm (see Fig. 1.9). When the SA node fires more quickly than usual (for instance, as a normal physiologic response to fear), the rhythm is termed sinus tachycardia (see Fig. 1.10). When the SA node fires more slowly than usual (for instance, either as a normal physiologic response in a very well-conditioned athlete, or an abnormal response in an older patient taking too much heart-slowing medication), the rhythm is known as sinus bradycardia (see Fig. 1.11). An *arrhythmia* is any abnormal cardiac rhythm. There may be

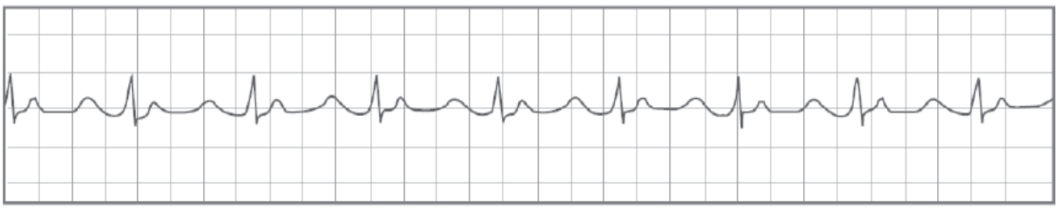


Figure 1.9: Normal sinus rhythm. Image adapted from [CAM06].

cyclic variations in heart rate due to breathing, known as sinus arrhythmia (see Fig. 1.12). This nonpathologic pattern is caused by activity of the parasympathetic system (the sympathetic system responds too slowly to alter heart rate on this time scale), which is responding to subtle changes in arterial blood pressure, cardiac filling pressure, and the lungs themselves, during the respiratory cycle.

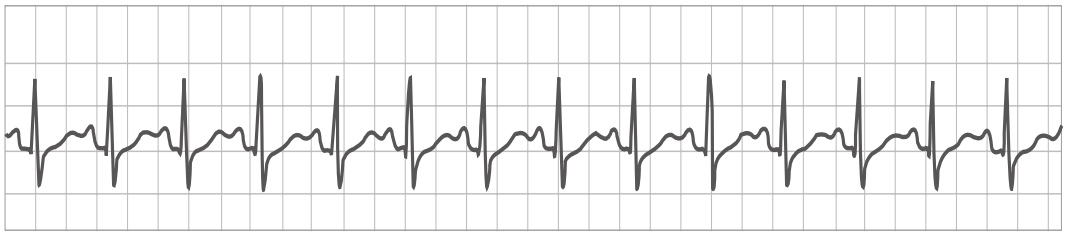


Figure 1.10: Sinus tachycardia. Image adapted from [CAM06].

Figure 1.11: Sinus bradycardia. Image adapted from [CAM06].

Ectopy and Fibrillation One category of arrhythmias occurs when the trigger to depolarize originates outside of the SA node, in another part of the myocardium (known as ectopic depolarization). Common causes of ectopy include a drug effect (e.g., caffeine) or a viral infection of the myocardium, or other inflammation or damage of part of the heart. When the ectopic beat originates in the atrium, it leads to an atrial premature contraction (APC) (see Fig. 1.13). When it originates in the ventricles, it leads to a premature ventricular beat or ventricular premature contraction (VPC) (see Fig. 1.14).

In some cases, the unified wavefront of depolarization can break down into countless smaller wavefronts which circulate quasi-randomly over the myocardium. This leads to a total breakdown of coordinated contraction, and the myocardium will appear to quiver. This is termed *fibrillation*. In atrial fibrillation (see Fig. 1.15), the A-V node will still act as a gatekeeper for

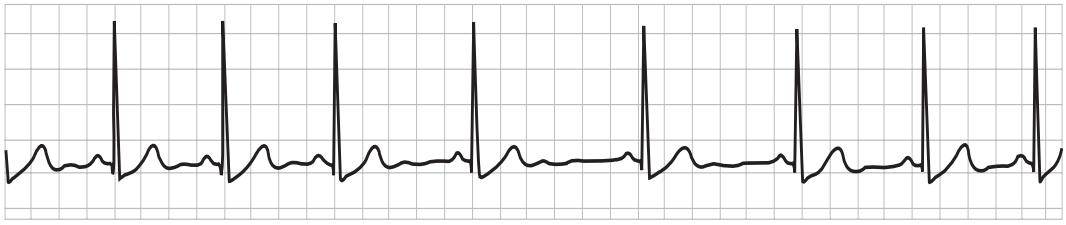


Figure 1.12: Sinus arrhythmia. Image adapted from [CAM06].

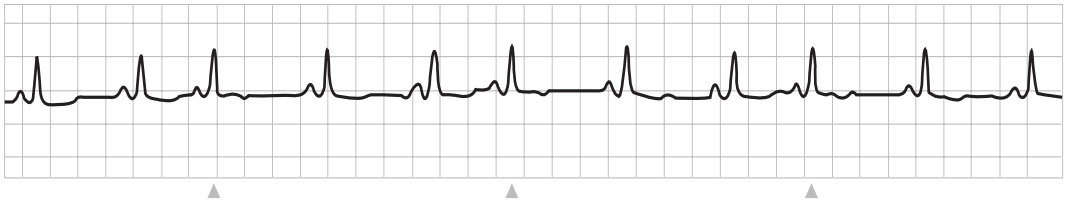


Figure 1.13: Atrial premature contractions (indicated by arrowheads). Image adapted from [CAM06].

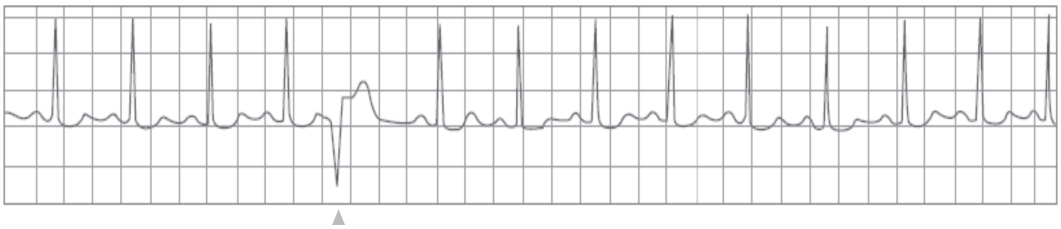


Figure 1.14: Ventricular premature contractions (indicated by arrowheads). Image adapted from [CAM06].

these disorganized atrial wavefronts, maintaining organized ventricular depolarization distal to the A-V node with normal QRS complexes. In contrast to atrial fibrillation, the ventricular fibrillation (see Fig. 1.16) is fatal in seconds to minutes: the appearance of fibrillating ventricles has been likened to a "bag of worms" and this causes circulatory arrest, the termination of blood flow through the cardiovascular circuit. Note that as pointed out by P. Laguna *et al.* in [LJC94], during fibrillation, it is meaningless to measure P, QRS and T wave boundaries and to analyze the waveforms. This kind of ECG pattern is out of the scope of this thesis work.

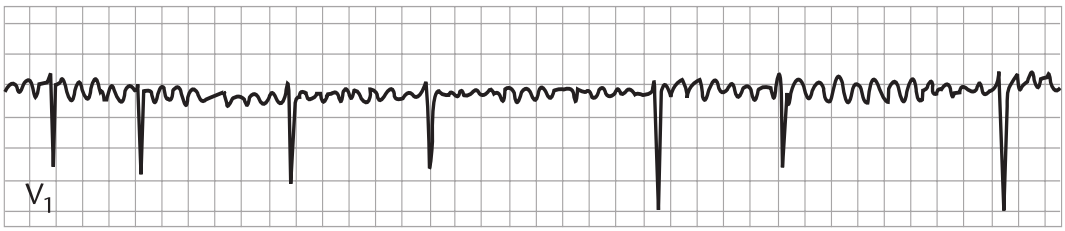


Figure 1.15: Atrial fibrillation. Image adapted from [CAM06].

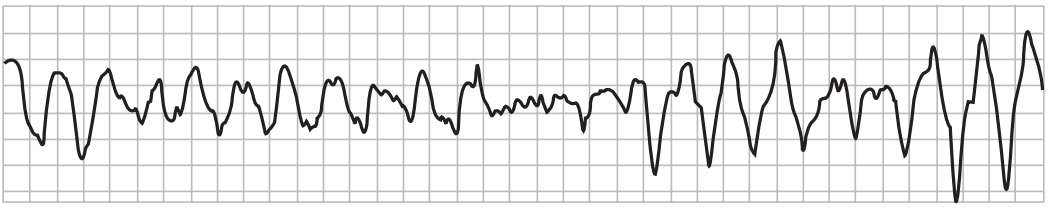


Figure 1.16: Ventricular fibrillation. Image adapted from [CAM06].

Escape Rhythms The other category of arrhythmias is related to excessively slow rhythms, and abnormal blockages of wavefront propagation. As an extreme case of the heart blockages, an escape beat (see Fig. 1.17) is similar to an ectopic beat, in that it is the initiation of a depolarization wavefront outside of the SA node. However, the difference is that the escape beat is a normal, compensatory response, a normal fail-safe functionality of the heart: there is a network of cardiac cells able to initiate heart beats, so that a key life-sustaining function (e.g., pacing the heart) is not exclusively relegated to a microscopic collection cells in the SA node.

Clinical ECG interpretation procedure

In analyzing the clinical electrocardiogram, it is important to use a systematic approach. Meanwhile, it may also provide a relevant framework in which to apply numerical analysis techniques to the ECG. The following clinical approach, which is extracted from [CAM06], should not be considered completely thorough, but simply as a guide to understanding how clinicians identify abnormalities in the ECG.

1. **Identify the QRS complexes.** The following observations should be made:

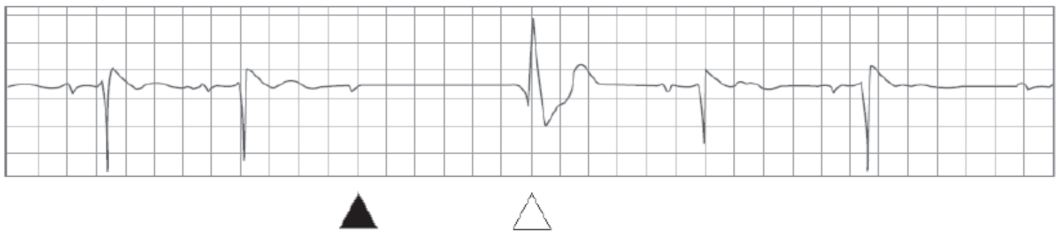


Figure 1.17: Ventricular escape beat. Note the atrial P wave (black arrowhead) followed by an evident pause, indicating a failure to conduct through the A-V node. The ventricular escape beat (white arrowhead) is a fail-safe mechanism so that conduction blocks do not cause ventricular cardiac arrest. Image adapted from [CAM06].

- What is the ventricular rate?
- Are the QRS complexes spaced at regular intervals?
- Are the QRS complexes identical in shape in a given lead? Are they of normal size and morphology?

2. **Identify the P waves.** In some cases this will require careful observation, and the following questions should be explored:

   - Is there a one-to-one relationship between P-waves and QRS complexes?
   - Is the PR interval of normal duration?
   - Are the P waves identical in shape in a given lead? Are they of normal size and shape?

3. **Examine the QRS complex.** Is the QRS axis normal? Overall, are the QRS widths and amplitudes normal?

4. **Examine the ST-T segments.** Are there abnormalities (such as elevation or depression)? Is the abnormality suggestive of ischemia, infarction, or hypothermia?

5. **Examine the T waves.** Are their shapes normal? In each lead, are they oriented in the same direction as the QRS complex? If not, is it suggestive of ischemia or ventricular conduction abnormalities, or a potassium abnormality?

6. **Examine the QT interval.** Is it over half the RR interval?

As can be seen from this procedure, the detection and delineation of the P and T waves as well as the estimation of their waveforms are very important steps to the clinical interpretation of the ECG. Most of the phases in the procedure require a good estimation of the wave fiducial points (onset, peak and end) and that of the waveforms (wave shapes). This justifies the vital role of automated P and T wave analysis via signal processing techniques in ECG computer-aided diagnosis.

1.4 Literature review of ECG signal processing methods

As presented in the previous section, the ECG data may be composed of single-lead or multiple-lead ECG signals depending upon the type and configuration of the ECG recorder. Accordingly, the method of analysis is also different. Single-lead ECG waveform analysis includes wave shapes (morphologies), spectra and repeatability of the cardiac cycle. On the other hand, multi-lead ECG processing algorithms can utilize additional information like simultaneous features from other leads. This may lead to a greater immunity against interference signals [LJC94]. The disadvantage of multiple-lead ECG monitoring lies in increased patient discomfort and stress, especially for ambulatory testing and wearable ECG recording. For the purpose of basic cardiac monitoring during ambulatory testing, it is desirable to have fewer leads and hence single-lead is more suitable for wearable ECG applications [CPD09]. This work focuses on the single-lead based P and T wave analysis in ECG signals. The following sections review some of the existing techniques developed for analysis of single-lead ECG signals, with the emphasis on the P and T wave delineation and waveform estimation.

1.4.1 Preprocessing

It is well known that noise from various sources like muscular activities, 50/60Hz powerline, skin stretching and electrode motion, movement of heart due to respiration, etc. can contaminate the ECG signal and hence affect the interpretation of ECG signal. In particular, an automated analysis requires noise free ECG signal for correct interpretation. However, it is difficult to control the environment and prevent the interference due to some physiological events like breathing. Consequently, a pre-processing stage which allows the reduction of noise is generally believed to be essential in ECG signals.

Powerline noise cancellation

Since spectral bandwidth of ECG can be considered to be between 0 to 250Hz, powerline is the most usual source of interference in the ECG recording. This kind of interference is caused due to powerline cords nearby, and its effect can be minimized by moving aways from such sources of this noise. Usually, the power-lines have a specific frequency of either 50 or 60Hz. Therefore, the interference is usually removed by using a narrow stop-band filter centered at the powerline frequency in the frequency response of the ECG equipment, which is usually from $0.05-100$Hz. In [AS85], authors proposed a technique for removing the powerline interference using a non-recursive finite impulse response. The notch filter for powerline is acceptable by the guidelines provided in [Gib02] for exercise monitoring ECG equipment. In many places, however, the powerline noise is not a pure 60Hz (or 50Hz) sinusoid, but is distorted. In this case, the adaptive filtering technique can be more effective. Furno and Tompkins [FT83] and Sahakian and Furno [SF83] described filter designs that subtract a 60 Hz sinusoid from ECG. In [TZ91], adaptive filtering techniques are applied for cancellation of powerline and the electromyograph (EMG) interference. The powerline interference appears as the common mode signal to the ECG amplifier and available from the right leg electrode. Hence the signal from the right leg electrode is used as the reference input signal to the adaptive filter for cancellation of powerline

noise. Moreover, it has been shown in [Ham96] that the adaptive implementation introduces less noise in measurement of the ST segment in comparison to that by a non-adaptive notch filter.

Baseline wandering reduction

Low frequency artifacts and baseline wander may be caused in the chest lead ECG signals by coughing or breathing, with large movements of the chest, or when an arm or leg is moved during the ECG data acquisition. Poor contact of the electrodes and perspiration of the patient under the electrodes may affect the electrode impedance which causes low frequency artifacts. Baseline drift may sometimes be caused by variations in temperature and bias in the instrumentation and amplifiers as well. The simplest approach for removal of the baseline is to filter the ECG signal using high-pass digital filters with a cutoff of ~ 0.8 Hz [AS85]. However, as with the powerline cancellation, such a filtering operation introduces distortions in the ST segment of the ECG which plays a vital role in the diagnosis of different life threatening cardiac disorders such as coronary artery heart disease. The selection of an optimal cutoff frequency such that the filter introduces minimum distortion in the ECG is an issue. Some sophisticated filtering based approaches use adaptive filtering [LJM+92], time varying digital filters using Short Time Fourier Transform (STFT) for estimation of cutoff filters [Pan96], cubic spline curve fitting [MK77], linear interpolation between isoelectric levels [CM07] or through wavelet transform [Zha05, SS07]. A comprehensive comparison of different ECG baseline removal techniques can be found in [ARA09]. Several widely used methods are introduced briefly as follows.

Adaptive filtering. Adaptive filtering has been used for baseline removal from the ECG in [LJM+92] using the architecture shown in Fig. 1.18. In the adaptive filtering based approach,

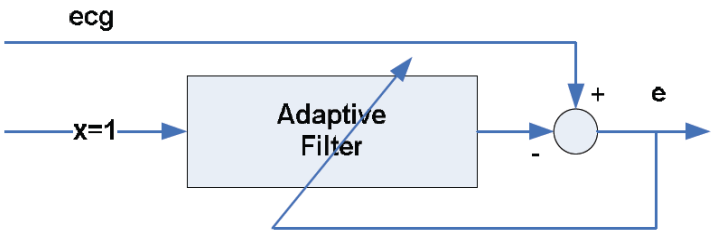


Figure 1.18: Adaptive Filtering for ECG Baseline Removal. Image adapted from [LJM+92].

only one weight is needed and the reference input is a constant with a value of one. The optimal weight w is determined using the Least Mean Squares (LMS) algorithm as follows

$$w(k+1) = w(k) + 2\mu e(k)x(k) \tag{1.1}$$

where $x(k)$ is the recorded ECG, $e(k)$ is the difference between the ECG and the output of the adaptive filter and μ is the step size. This filter has a zero at 0Hz and consequently it creates a notch with a bandwidth of $\left(\frac{\mu}{\pi}\right) F_s$ where F_s is the sampling frequency. However, as pointed out in [PLY98], this approach produces severe distortion in the ECG signal, especially in the

ST segment area.

Cubic spline curve fitting. This method is among the most commonly used approaches for removal of ECG baseline variation. In this approach [MK77], isoelectric fiducial points (generally the Q, R, and S peaks) are found in the ECG signal with baseline variation for each beat using an appropriate approach and a third order cubic spline is fitted on these points to obtain an estimate of the baseline which is then subtracted from the original ECG signal. Cubic spline interpolation based baseline removal and other interpolation based techniques adapt themselves automatically to the heart rate as more reference points become available with increase in heart rate. However, in the absence of any baseline variation in an ECG segment, an error in the calculation of the isoelectric reference point or the corresponding level causes undesired distortion in the ECG. Therefore accurate definition of the isoelectric reference point is mandatory for proper functioning which can become difficult in the presence of noise in the ECG signal. Furthermore, as pointed out in [BZB06], the cubic spline interpolation method and other interpolation based techniques only remove baseline fluctuations at frequencies lower than heart rate. This makes the residual local baseline in the non-QRS components problematic for many ECG analysis.

Median filtering. In [CM07], Chouhan *et al.* give a technique for baseline removal using median filtering. In this approach, the median of the ECG signal is firstly subtracted from the ECG signal. Then a fifth order polynomial is fitted to this shifted waveform to obtain a baseline estimate which is then subtracted from the ECG signal. The baseline drift is further removed by applying a content median correction, one by one, in each RR interval. This approach also offers the advantage that the signal is not distorted in the absence of baseline variation and is computationally efficient.

In-band noise cancellation

The numerous non-cardiac ECG contaminants, such as the EMG noise and the motion artifact overlap with the cardiac components in the frequency domain, particularly in the 0.01Hz to 100Hz range. It is still a challenging problem to remove this kind of in-band noise reliably without affecting the cardiac components of the ECG signal. Band-pass filtering is therefore inadequate to suppress such contaminants. Ensemble Averaging (EA) is a 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 [LB97]. As an improvement over EA, classical adaptive filter architectures have been used for the noise cancellation of ECGs containing EMG noise and motion artifacts [LJM+92, TZ91].

A nonlinear Bayesian filtering framework has been recently proposed for the filtering of single channel noisy ECG recordings [SSJC07]. This effective framework for the model-based filtering of noisy ECG recordings is based on a modified nonlinear dynamic model, previously suggested for the generation of a highly realistic synthetic ECG [MCTS03]. 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. Superior results compared with conventional ECG denoising approaches such as band-pass filtering, adaptive filtering, and wavelet denoising are reported in [SSJC07] on real non-stationary muscle artifacts.

1.4.2 QRS detection

As we have seen in 1.3.2, ECG is a pseudo-periodic signal in the sense that the cardiac cycle repeats according to heart rate. However, the heart rate may not remain constant. The variations in the heart rate may affect the durations of PQ and ST segments while the durations of P wave, QRS complex and T wave may still remain the same for a normal heart. The R peak in the QRS complex is the dominant feature of the cardiac cycle, which can be distinctly recognized from the sharp edges and a high amplitude as we have seen in Fig. 1.7. Therefore, in most circumstances it is relatively easy to locate the QRS complex in the ECG even in the presence of low frequency noise (like baseline wandering due to respiration) and hence this is used for determining the current heart beat. The QRS detection forms the basis of most ECG analysis algorithms, particularly those used for arrhythmia monitoring, wave delineation, interval measurement, etc. This explains the importance of QRS detection in cardiac monitoring using ECG.

Most extensively used real-time QRS detection algorithms, in general, use the relatively high energy contents of the QRS complex that lie in 5-25Hz band [PT85, KHO07]. The more complex QRS detection algorithms involve application of neural network, hidden Markov model (HMM), syntactic methods, etc. [CSCB90, HTUA93, Suz95, UM90], but they are rarely used in low computational cost ECG applications. Further details of the QRS detection methods and the comparisons of their performances in presence of noise and their computational complexities can be found in [FJJ+90, KHO07]. Most of the simple QRS detection algorithms are based on one of the following methods: derivatives, wavelets, filter-banks, mathematical morphology and correlation [FJJ+90, KHO07]. In the following paragraphs a few of the approaches in literature for QRS complex detection are discussed in brief.

Temporal derivative. The characteristic of higher slopes of the QRS complex inspires one to use temporal derivatives for its detection. In the derivative based methods such as the popular Pan and Tompkins method [PT85], the ECG signal is first smoothed with an appropriate moving average filter for suppressing any high frequency noise outside the 5-25Hz band. The smoothed signal is differentiated to emphasize the high slopes and to suppress smooth ECG waves and baseline wanders. The overall response of these two simple arithmetic operations results in a band-pass filter to match the spectral band of the QRS while suppressing the relatively low frequencies in P and T waves. The squared magnitude of the derivative signal is used to enhance further the high derivatives of the QRS complex. A moving average integration filter with the window length matching the duration of QRS complex is applied after the squaring operation. The integrated signal is then searched for the local maxima exceeding an appropriate threshold value. The search is further refined by eliminating the points which occur within the refractory period of a ventricular activity. A block diagram which describes different processing steps of the algorithm is shown in Fig. 1.19, and Fig. 1.20 shows signal examples at various steps.

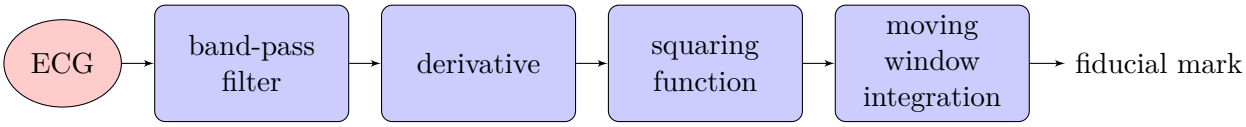


Figure 1.19: QRS detection processing steps of Pan and Tompkins algorithm in [PT85].

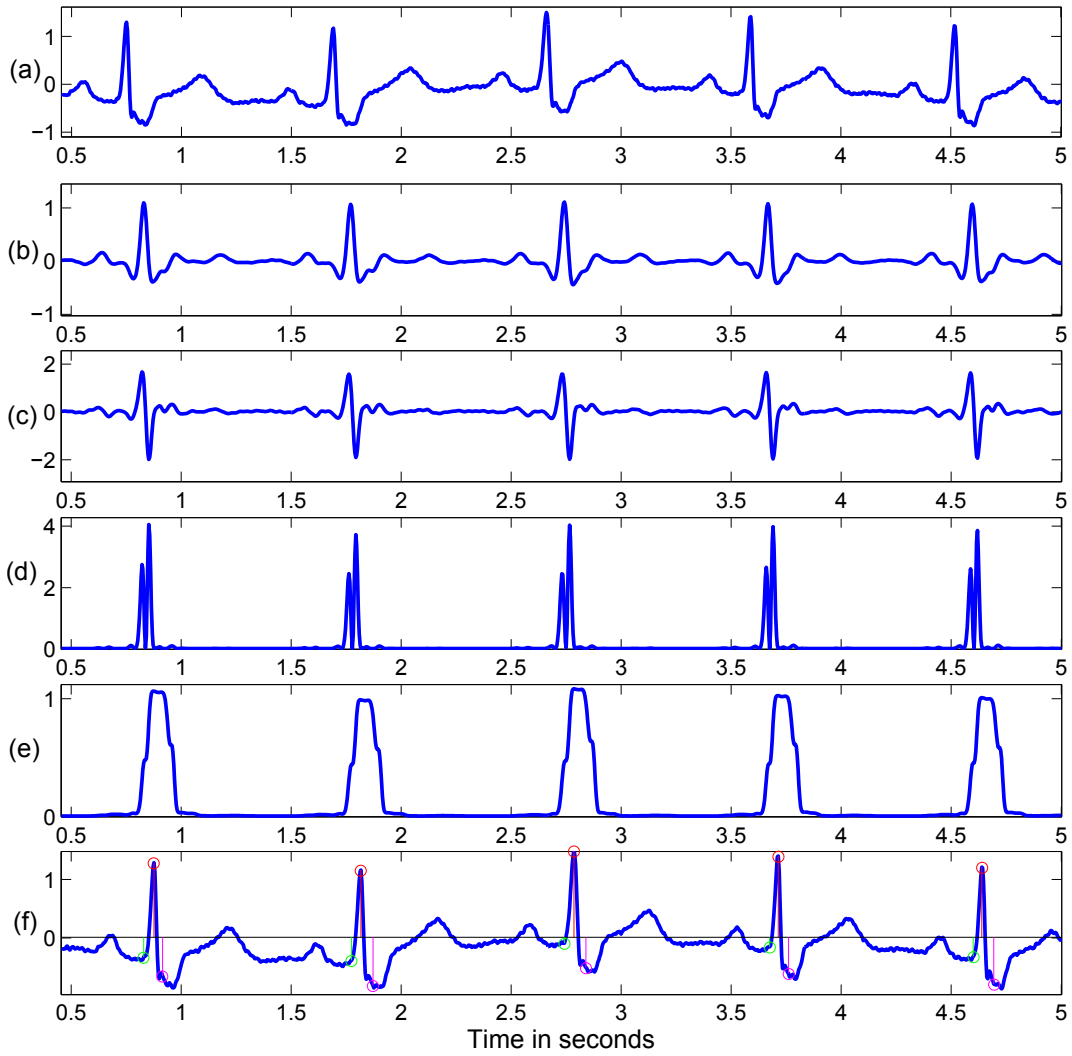


Figure 1.20: Results of the Pan and Tompkins algorithm [PT85]: (a) Original signal. (b) Output of band-pass filter. (c) Output of differentiator. (d) Output of squaring process. (e) Results of moving window integration. (f) Original signal delayed by the total processing time with the estimated QRS marks.

Wavelet transform. Wavelet based methods for QRS detection use the principle of singularity detection in the wavelet coefficients. Wavelet coefficients of the ECG signal at several scales are analyzed [LZT95, STB97, KMBB99] to find the local maxima and positions of matching in two consecutive scales to locate QRS positions. This is based on the assumption that the energy of the QRS complex is continuously spread over the spectral bands as well as the temporal scales. The noise in the signal may not have this kind of property and hence false alarms due to such noise can be reduced using this multiscale approach.

Filter bank. In a filter bank approach of QRS detection, sub-bands at different scales are combined to confirm the positions of the local maxima [ATNL99]. The filter bank approach is based on the fact that the QRS complex has simultaneous presence in the sub-bands, whereas other ECG waves and noise may not exhibit this characteristic behavior. This is similar to the wavelet based approach of QRS detection. In [BPSN82], a generalized class of filter with a transform having two factors, $\left(1 - z^{-K}\right)\left(1 + z^{-1}\right)^{L}$ is given: the first implements a difference with integer delay K and the other is for a low-pass filter with band width controlled by an integer parameter L. The integer parameters (K, L) are determined depending upon the sampling rate. In [EZ79], $(K, L) = (5, 4)$ is used at the sampling rate of 250Hz.

Mathematical morphology. In [ZMJ03], a curve length transform is used for the detection of QRS complex. The curve length is defined as the sum of Euclidean distances between pairs of consecutive sample points in the ECG signal. The curve length of the ECG signal depends on the increments in the sample values and the sampling time of the ECG. For uniform sampling in time the curve length is a measure of the increments in sample values. For QRS detection the curve length of the ECG signal is evaluated in a window with the length matching with the widest possible QRS complex. When the window is in perfect alignment with the onset of QRS complex it produces the local maximum in the curve length feature and is utilized for locating the onsets of the QRS complex. A low-pass filter with 3dB cut-off at 16Hz is used as a preprocessing step to suppress noise.

It should be noted that it is hard to design a single comprehensive preprocessing technique for achieving simultaneous QRS enhancement and noise reduction effectively in practice. In [PT85], the trade off between misses and false detections relies on the selection of bandwidth of the filter and size of the moving-window integrator. The WT-based QRS detector [LZT95] has the choice problem of mother wavelet and scales to obtain QRS events. Therefore, most of the works paid attention on constructing suitable decision rules based on the preprocessing results.

In this thesis work, the Pan and Tompkins algorithm has been chosen to accomplish the QRS detection because it is extensively implemented in many real-time ECG monitoring devices and offers well-recognized performances.

1.4.3 P and T wave delineation

In the previous section we have reviewed some of the QRS detection techniques in the literature. For automated analysis of the ECG, detection of P and T waves is also important as the P wave represents atrial activity and the T wave is related to repolarization of ventricles. As

mentioned before, P and T wave analysis is more complicated because of the low slope and magnitude of P and T waves. In a cardiac cycle the sequence of occurrence of these waves is P-QRS-T. Therefore, one can search for P and T waves within appropriate time windows after the QRS complex is located. Meanwhile, it is recommended in [LJC94] that any fibrillation condition should be detected using procedures proposed in [TZP90] before proceeding with P and T wave analysis because it is meaningless to detect P and T waves in such circumstance.

The T wave is the wave with the next highest level of energy in cardiac cycle. The location of T wave search window from the R peak depends on the current beat period, measured as the time interval between two consecutive R peaks which is simply called RR interval. In [LTC+90], the T wave search window is defined from the R peak position in the interval from 140 to 500ms if the RR interval mean is bigger than 700ms and for smaller RR intervals the search window is defined in the range 100ms to 0.7 times RR interval. In [FS91], the P wave search window is assumed to occur in a specified time window of 240 to 400ms preceding the R wave of the QRS complex in each cardiac cycle. Once the search windows are defined before and after the QRS location, an appropriate strategy is used to enhance the distinctive features of each wave in order to find its peaks and limits. In the last two decades, a variety of approaches have been proposed to automatically detect and delineate P and T waves in the ECG [TZ91, LJC94, LTC+90, Str02, MN92, MP92, LZT95, MAO+04, MSV95, BNY01, TS90, TRC06, CV06, SS09, GG89, CM08, YLL98, VGP+00, DHC10]. Some of the classical approaches in the literature are reviewed in the following.

Low-pass differentiation (LPD) based methods

One of the classical methods to detect the ECG wave boundaries is based on low-pass differentiation [LTC+90]. The P and T wave fiducial points (onset, peak and end) are located using a low-pass filter from the differentiated ECG.

As show in Fig. 1.21, the ECG signal is processed by a derivative filter $G_1(z)$ and a smoothing low-pass filter $G_2(z)$ such that:

$$G_1(z) = \frac{1 - z^{-8}}{1 - z^{-1}}, \quad G_2(z) = 1 - z^{-6} \tag{1.2}$$

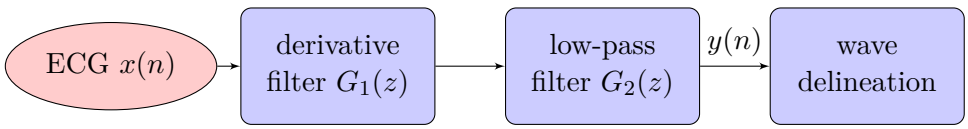


Figure 1.21: Block diagram for the P and T wave delineation based on LPD.

The output of the LPD $y(n)$ is then subjected to a thresholding method which adapts to different wave morphologies to determine the wave boundaries and wave peaks. Take T wave as an example, Laguna *et al.* [LTC+90] considered four types of T wave morphologies: normal T wave (monophasic and positive), inverted T wave (monophasic and negative), ascending biphasic T wave and descending biphasic T wave. The algorithm in [LTC+90] seeks in each

T wave search interval, maximum ($T_{\max}$) and minimum ($T_{\min}$) of signal $y(n)$. Three cases are then:

- The algorithm finds a T wave maximum before a T wave minimum. In this case, the method considers the T wave as an ascending biphasic T wave if $|T_{\max}| > 4\,|T_{\min}|$, otherwise the algorithm considers it as a normal T wave.
- The algorithm finds a minimum before a maximum. In this case, if $|T_{\max}| > 4\,|T_{\min}|$ the method considers the T wave as a descending biphasic T wave, otherwise it is considered to be inverted T wave.
- The algorithm finds a minimum before a maximum and another minimum after the maximum. In this case, if $|T_{\max}| < 4\,|T_{\min}|$ then the algorithm considers the T wave as normal, otherwise it is considered as inverted.

Once the T wave morphology type is determined, the algorithm locates the onset, peak and end points by using experimental thresholding based criteria. Fig. 1.22 illustrates one example on a normal T wave. The zero-crossing of $y(n)$ is considered as the wave peak (see T_{pic} in Fig. 1.22), and the point T_2 where $y(n)$ exceeds half the amplitude of the last summit of $y(n)$ denoted as $y(T_{\text{i}})$ is considered as the wave end.

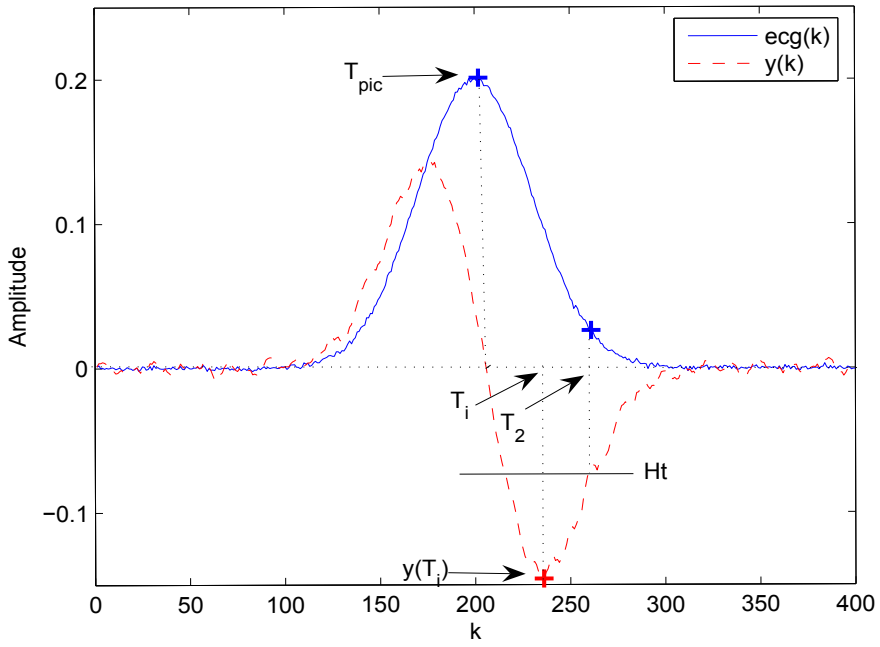


Figure 1.22: Detection of the peak and the end of a normal T wave by using the method [LTC+90]. Image adapted from [Cab08].

The advantage of the LPD based methods is that they are generally simple to implement and robust. The major drawback is related to the differentiation which is known to be sensitive

to noise. Another disadvantage is that the delineation criteria require *a priori* information on the waveform (the class of morphology) and the delineation performance relies on thresholds which are chosen arbitrarily according to experimental values.

Wavelet Transform (WT) based methods

Another class of P and T wave delineation methods are based on the wavelet transform (WT). This transform provides a description of the signal in time-scale domain, allowing the representation of the temporal features of a signal at different resolutions. Similar to the QRS detection, this technique has also shown interesting properties to P and T wave detection and delineation problems. As a pioneer work, Li *et al.* proposed a method for detecting the QRS complex by using WT [LZT95]. In their work, Li and coworkers also applied the WT to monophasic P and T waves delineation problem, although only the QRS detector was validated. In [MAO+04], a generalization of that method was presented, including the determination of the individual QRS waves, and a robust delineation of QRS, P and T waves for a wide range of morphologies. The performance of the WT based strategy was assessed using standard manually annotated ECG databases, which makes it easier to compare with other methods.

The wavelet transform is a decomposition of the signal as a combination of a set of basis functions, obtained by means of dilation a and translation b of a single prototype wavelet $\psi(t)$. Thus, the WT of a signal $x(t)$ is defined as

$$W_{a,b}[x(t)] = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\, \psi\left(\frac{t-b}{a}\right) \mathrm{d}t, \;\; a > 0 \tag{1.3}$$

The greater the scale factor a is, the wider is the basis function and consequently, the corresponding coefficient gives information about lower frequency components of the signal, and vice versa. In this way, the temporal resolution is higher at high frequencies than at low frequencies, achieving the property that the analysis window comprises the same number of periods for any central frequency. In practice, the scale factor a and the translation parameter b can be discretized. The usual choice is to follow a *dyadic grid* on the time-scale plane: $a = 2^k$ and $b = 2^k l$. The discrete wavelet transform (DWT) can be formed with basis functions

$$\psi_{k,l}(t) = 2^{-k/2} \psi(2^{-k}t - l), \;\; k, l \in \mathbf{Z}^+ \tag{1.4}$$

Fig. 1.23 shows several simulated waves similar to those in the ECG, together with the first five scales of their DWT ($k \in \{1, \ldots, 5\}$). The local maxima and minima of the DWT indicate the local singular points of the considered signal. Note that the same sampling rate is applied in all scales to keep the time-invariance and the temporal resolution. This is achieved by removing the decimation stages and interpolating the filter impulse responses of the previous scale [MAO+04].

Fig. 1.24 summarizes the block diagram of a WT based P and T wave delineator. First, the DWT of the ECG signal is computed on five scales, $w_{x,k}(n)$ denotes the DWT of $x(n)$ at scale 2^k. The detection of the fiducial points is then carried out across the adequate WT scales, attending to the dominant frequency components of each ECG wave: QRS waves correspond

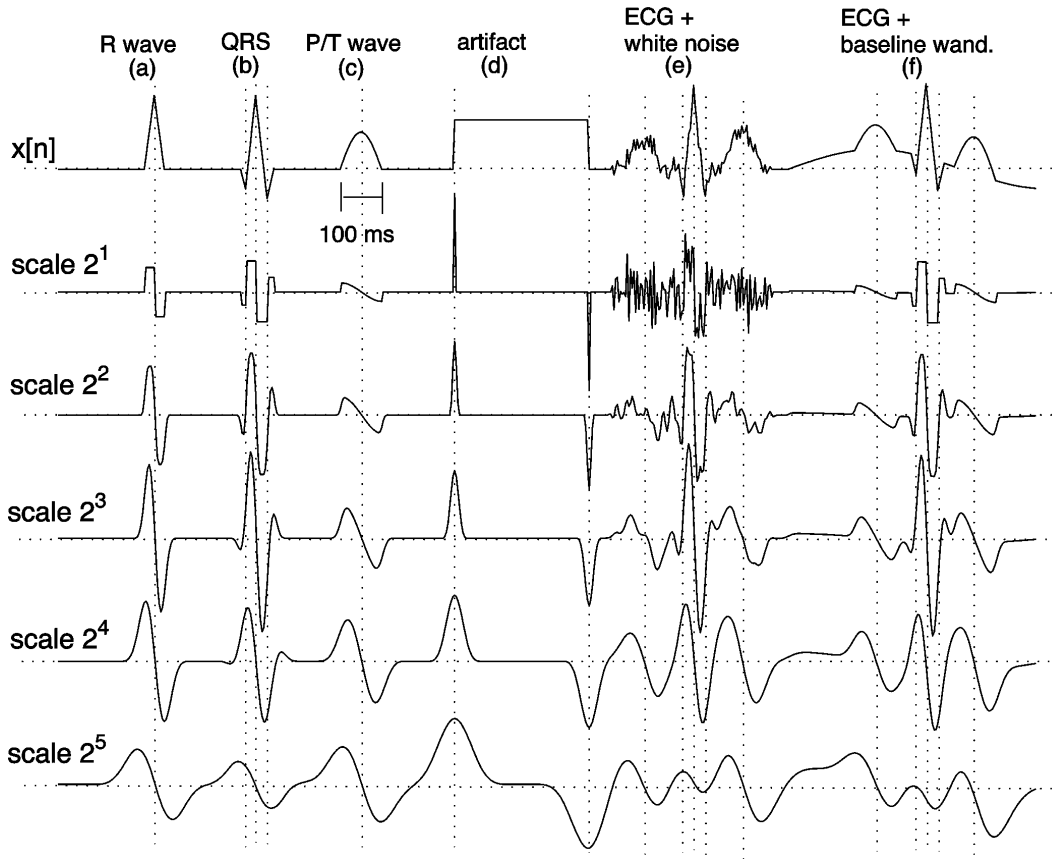


Figure 1.23: WT at the first five scales of ECG-like simulated waves. Image adapted from [MAO+04].

to a simultaneous effect in scales $2^1 - 2^4$, while the T and P waves affect mainly scales 2^4 or 2^5. Using the information of local maxima, minima and zero crossings at different scales, the algorithm identifies the significant points of the QRS complex first, and then those of the P and T waves.

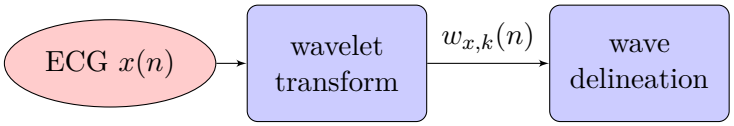


Figure 1.24: Block diagram for the P and T wave delineation based on WT.

Depending on the number and polarity of the slopes found, a wave morphology is assigned and boundaries are located using threshold-based criteria [MAO+04]. Let us take a T wave as an example (Fig. 1.25), the onset of a wave, denoted as n_o, occurs before the first significant slope associated with the wave (the first maximum of $|w_{x,4}(n)|$), at sample n_f. In the same manner, the end of a wave, denoted as n_e, occurs after the last significant slope associated with the wave (the last maximum of $|w_{x,4}(n)|$), at sample n_l. Each boundary is located by selecting the sample nearest to n_f (n_l) where one of the following two criteria is satisfied: 1) $|w_{x,4}(n)|$ is below a threshold $\xi_o(\xi_e)$ relative to $|w_{x,4}(n_f)|$ ($|w_{x,4}(n_l)|$) and 2) a local minimum of $|w_{x,4}(n)|$ exists before n_f (n_l).

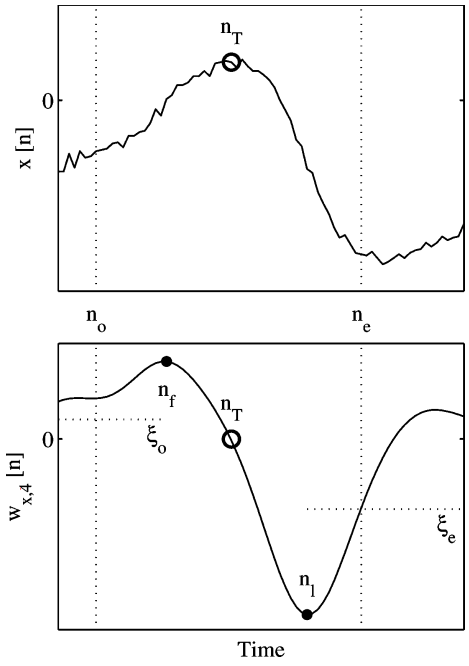


Figure 1.25: Detection of a biphasic T wave boundaries by using its DWT. Image adapted from [AMRL09].

Regarding the purpose of locating different waves with typical frequency characteristics, the WT is a suitable tool for ECG automatic delineation. The major drawbacks of the WT

based approach are that a priori information is required on the waveform (delineation criteria depend on the class of morphology) and width, and rigid arbitrary thresholds are applied to determine the significance of the wave components to detect the wave and their boundaries. The performance of the algorithm can be affected by the choice of the threshold values which may differ among specified ECG datasets. This issue has been recently addressed by Dumont *et al.* in [DHC10]. An evolutionary optimization process based on a learning period was proposed to determine the proper threshold values for WT based methods.

Model-based approaches

The wave delineation can also be based on the concept of fitting a realistic model to the ECG signal and extracting parameters from the model to determine waveform fiducial points. Particular attention has been devoted to Gaussian mixture state-space models whose parameters can be estimated with nonlinear gradient descent [CV06], Kalman filters [SS09] or particle filters [LBMT11]. This idea originates from the work by McSharry *et al.* [MCTS03], which suggested the use of Gaussian mixture models to generate realistic synthetic ECGs. The model generates a three-dimensional trajectory which consists of a circular limit cycle in the polar plane that is pushed up and down as it approaches each of the wave centers, as shown in Fig. 1.26. The simplified discrete polar form [SSJC07] of the model is shown as

$$\begin{aligned} \theta_{k+1} &= \theta_k + \omega T s \\ x_{k+1} &= -\sum_{j=1}^{J} T s \frac{\alpha_j \omega}{b_j^2} \Delta\theta_{j,k} \exp\left(-\frac{\Delta\theta_{j,k}^2}{2b_j^2}\right) + x_k + \eta_k \end{aligned} \tag{1.5}$$

where $x_1, ..., x_K$ denote the ECG samples and $\Delta\theta_{j,k} = (\theta_k - \nu_j) \mod (2\pi)$, ν_j denotes some specific time instants and α_j and b_j are model parameters which are characterized in the following. Here, $T_s = 1/F_s$ is the sampling time, ω is the angular velocity and η_k is a random time variant noise which has been placed to represent the baseline wander. The summation over j is taken over the number of Gaussian functions (or turning points) J used for modeling each of the ECG component. In [SSJC07], the authors proposed to use five Gaussian functions to model the ECG channels containing the P, Q, R, S and T waves such that $J = 5$. Fig. 1.27 shows the five Gaussian functions with arrows indicating the kernels' effect intervals. Moreover, the corresponding parameters of the Gaussian functions, i.e., the amplitude $\boldsymbol{\alpha} = [\alpha_1 \ \dots \ \alpha_5]^T$, the angular spread $\boldsymbol{b} = [b_1 \ \dots \ b_5]^T$ and the location $\boldsymbol{\nu} = [\nu_1 \ \dots \ \nu_5]^T$ were fixed as presented in Table 1.1.

Table 1.1: Fixed ECG feature parameters proposed in [SSJC07]

Index (j)	P	Q	R	S	T
α_j	1.2	-5.0	30	-7.5	0.75
ν_j	$-\pi/3$	$-\pi/12$	0	$\pi/12$	$\pi/2$
b_j	0.25	0.1	0.1	0.1	0.4

The space-state model presented in (1.5) can be modified and extended to include the evolution of the Gaussian function parameters. In the modified dynamical model, the Gaussian

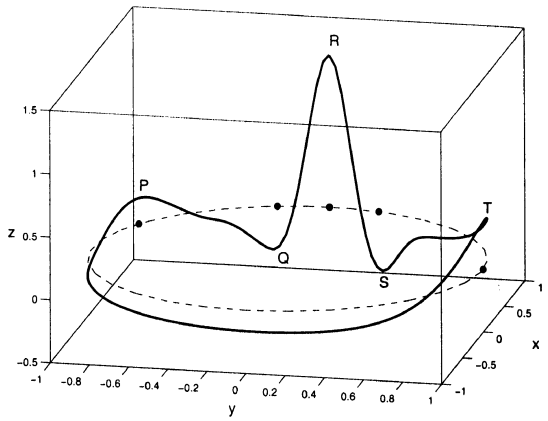


Figure 1.26: Typical trajectory generated by the ECG dynamical model [MCTS03] in the 3D space given by (x, y, z). Image adapted from [MCTS03].

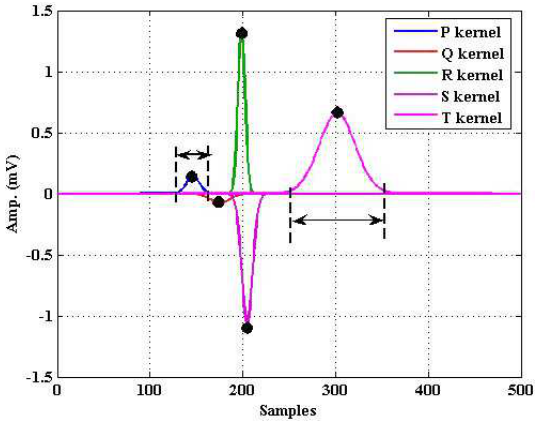


Figure 1.27: Five Gaussian functions with arrows indicating the kernels' effect intervals. Image adapted from [SS09].

function parameters are considered as hidden-state variables with first order autoregressive dynamics and no corresponding observations. This modified model is summarized as follows

$$\begin{aligned}
\theta_{k+1} &= \theta_k + \omega T s \\
x_{k+1} &= -\sum_{j=1}^{5} T s \frac{\alpha_{j,k}\omega}{b_{j,k}^2} \Delta\theta_{j,k} \exp\left(-\frac{\Delta\theta_{j,k}^2}{2b_{j,k}^2}\right) + x_k + e_{z,k} \\
\boldsymbol{\alpha}_{k+1} &= \boldsymbol{\alpha}_k + \boldsymbol{e}_{\alpha,k} \\
\boldsymbol{\nu}_{k+1} &= \boldsymbol{\nu}_k + \boldsymbol{e}_{\nu,k} \\
\boldsymbol{b}_{k+1} &= \boldsymbol{b}_k + \boldsymbol{e}_{b,k}
\end{aligned} \tag{1.6}$$

where $\boldsymbol{e}_{\alpha,k}, \boldsymbol{e}_{\nu,k}$ and $\boldsymbol{e}_{b,k}$ are additive mutually independent white noise vectors whose variances determine how fast the parameters are expected to change with time. Concerning the observation equation, besides the ECG observations x, Sameni *et al.* proposed in [SSJ05] to add the phase ϕ as a second observation, which can be simply obtained by detecting the R peaks. Hence, we have two noisy observation vectors $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_K)$ and $\boldsymbol{s} = (s_1, \cdots, s_K)$ (where K is the number of observed samples), corresponding to the state variables θ and x. Other state variables are considered as hidden states. The resulting observation equation of the dynamical model can be defined as

$$\begin{aligned}
\phi_k &= \theta_k + u_{\phi,k} \\
s_k &= x_k + u_{s,k}
\end{aligned} \tag{1.7}$$

where $u_{\phi,k}$ and $u_{s,k}$ are observation noises which are assumed to be Gaussian. An EKF-based Bayesian filtering method was proposed to estimate the unknown parameters of the model (1.6) in [SS09]. Note that only Gaussian noise is considered in the work of [SS09]. However, in the presence of motion artifacts, environmental noises and bioelectrical artifacts, non-linearities or/and non-Gaussian noises can be introduced in the ECG dynamical model. This problem

was addressed in [LBMT11], where a particle filtering based method is proposed to resolve a similar Bayesian framework.

Based on the estimates of the model parameters, the locations of characteristic waveforms and their corresponding fiducial points can be determined by using certain criteria. It is assumed in [SS09] that each ECG characteristic waveform is symmetric and thus the ECG waveform peaks correspond to the center parameters of Gaussian functions, where the trajectory reaches to its local maxima/minima. This way, the ν_j parameter, which stands for the mean of the j-th Gaussian kernel, determines the wave peak locations. The center positions for a typical synthetic trajectory are shown in Fig. 1.27 with small filled circles. Concerning the onset and offset of P and T waves, the spread parameter b_j is used. The approximately 99% confidence bound towards the spread parameters is considered as the termination of the two Gaussian functions representing these waves. In other words, when any of the two Gaussian functions representing P and T waves in the dynamical model extends to three times its spread (equal to 99% confidence bound), it is considered as the onset point. The same is true for the offset point. The block diagram for the Gaussian mixture model based delineation method is shown in Fig. 1.28.

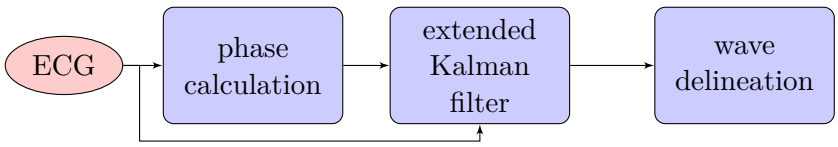


Figure 1.28: Block diagram for the Gaussian mixed model-based delineation method.

To conclude the methods of this kind, the advantages are their sequential property which is suitable for on-line ECG processing and their low computational load. Nevertheless, this class of methods requires the number of Gaussian kernels known a priori. If the assumption that the waves are symmetric does not hold, it may have difficulties on handling abnormal morphologies and rhythms.

Other P and T wave delineation techniques

Finally, we would like to mention other delineation strategies based on length transformation [GG89], uniform thresholding [CM08], approximating function theory [YLL98], characterization of TU complexes [VGP+00] and pattern recognition [TS90]. More recently, Cabasson proposed in [Cab08] to consider the P and T wave delineation as a time delay estimation problem. The Woody's method [AMRL67], which is often used on the identification and analysis of the delayed signals, was studied to deal with the ECG wave delineation. Note that some of these methods can only be used to obtain a subset of P and T-wave characteristic points and most of the methods mentioned here do not allow the estimation of the waveform. Note also that all the approaches mentioned above are single-lead based methods. Concerning the multi-lead ECG processing, one solution is to include post-processing decision rules to deal with multi-lead files by choosing global marks based on the single-lead based sets of locations. In [AMRL09], a multi-lead methodology regarding boundaries location is proposed and validated. The proposed approach departs from the single-lead WT based system and attends

to the spatial characteristics of the different available leads, aiming to achieve a more robust delineation. In this work, we will limit ourself on single-lead ECG processing.

Chapter 2

P and T wave analysis using window based Bayesian models and MCMC methods

Contents

2.1 Introduction

In this chapter, we introduce a window based Bayesian model which simultaneously solves the P and T wave delineation task and the waveform estimation problem. This model takes into account appropriate prior distributions for the unknown parameters (wave locations and amplitudes, waveform coefficients). The prior distributions and the likelihood of the observed

data are combined to provide the posterior distribution of the unknown parameters. This posterior distribution depends on hyperparameters that can be fixed a priori or estimated from the observed data. This work will consider both possibilities depending on the available information regarding the hyperparameters. To alleviate numerical problems related to the posterior associated to the P and T wave delineation, we propose to resort to Markov chain Monte-Carlo (MCMC) methods [RC04]. MCMC methods are powerful numerical tools, appropriate to solve complex Bayesian inference problems. Section 2.3.2 concentrates on a particular MCMC method referred to as partially collapsed Gibbs sampler (PCGS) whose convergence properties have been studied in [VP08]. The PCGS has shown interesting properties for electromyography (EMG) [GCIF11] and optical coherence tomography (OCT) [KNHH09], [KTHD10]. The ECG state sequence composed of the P wave, QRS complex and T wave parameters is assumed to be Markovian, since the current state (P-wave, QRS complex, and T wave) only depends on the previous state. This property motivates our study of a PCGS imposing a strong local dependency on the wave locations. The local dependency improves the convergence and the computational efficiency of the Gibbs sampler. The generated samples are used to estimate the unknown model parameters. The wave detection and delineation criteria based on the posterior distributions are presented in Section 2.3.4. Simulation results performed on the standard annotated QT database [LMGM97] as well as a comparison among the PCGS, the block Gibbs sampler and other algorithms are given in Section 2.3.5.

In Section 2.4, the proposed Bayesian model is modified to consider the baseline within each non-QRS component which was assumed to be filtered out in Section 2.2.2, and to represent P and T waves by their respective dimensionality-reducing expansion according to Hermite basis functions. The local dependency of the ECG signal is otherwise expressed by a block constraint, and a block Gibbs sampler is proposed to estimate the unknown parameters associated to the modified Bayesian model. The block Gibbs sampler further improves convergence and computational efficiency. Comparison results with the previously proposed PCGS algorithm on the QT database are presented in Section 2.4.5. Finally, conclusions are reported in Section 2.5.

2.2 Window based Bayesian model

It is common to view ECGs as the combination of two elements, namely, QRS complexes and non-QRS components. The interval between each successive pair of QRS offset and the subsequent QRS onset constitutes a non-QRS component. Due to the nonstationary nature of ECGs, detection and estimation must involve a limited set of consecutive beats. In the window based Bayesian model, we first detect QRS complexes that are the most prominent components of the ECG signal, and then we shift a non-overlapping D-beat processing window to cover the whole signal. In the processing window, detected QRS complexes become a reference for detecting P and T-waves. We define a T search interval and a P search interval for each beat, relative to the QRS complex boundaries depending on a recursively computed RR interval. The T and P search regions within the processing window are then extracted individually to form T and P wave search blocks.

2.2.1 Preprocessing

QRS detection

QRS complexes are detected using Pan and Tompkins algorithm [PT85] based on digital analysis of the slope, amplitude and width. A brief introduction of this technique is introduced in Section 1.4.2. The filtering that is done prior to this algorithm is found to be satisfactory. Thus no additional filtering is required before the delineation of P and T waves. Note that other QRS detection algorithm mentioned in Section 1.4.2 could be used in this preprocessing step.

Removal of baseline drift

In the proposed algorithm, waveform coefficients are estimated simultaneously with the wave locations. Baseline drift causes inaccurate waveform estimation results. For this reason, we employ the median filtering method proposed in [CM07] to remove both the global baseline drift and the local baseline in each RR interval (see Section 1.4.1 for more details). Note that other baseline removal technique discussed in Section 1.4.1 could be used in this preprocessing step.

Construction of P and T wave search blocks

As shown in Fig. 2.1, in the D-beat processing window, D right-hand neighborhoods of successive QRS offsets can be extracted to form T wave search windows. Suppose that k_{off} denotes a QRS offset location, then a T wave indicator can only appear in the right-hand neighborhood of k_{off} which can be denoted as $J_T(k_{\text{off}}) = (k_{\text{off}} + 1, \ldots, k_{\text{off}} + N_T)$, where N_T denotes the T wave search region width ($N_T = \text{RRI}/2$, where RRI is the length of the current estimated RR interval). Similarly, D left-hand neighborhoods of D successive QRS onsets can be extracted

to form P wave search windows. The length of each search interval within a search window is fixed according to the current estimated RR interval length RRI. Thus it will differ from one beat to another.

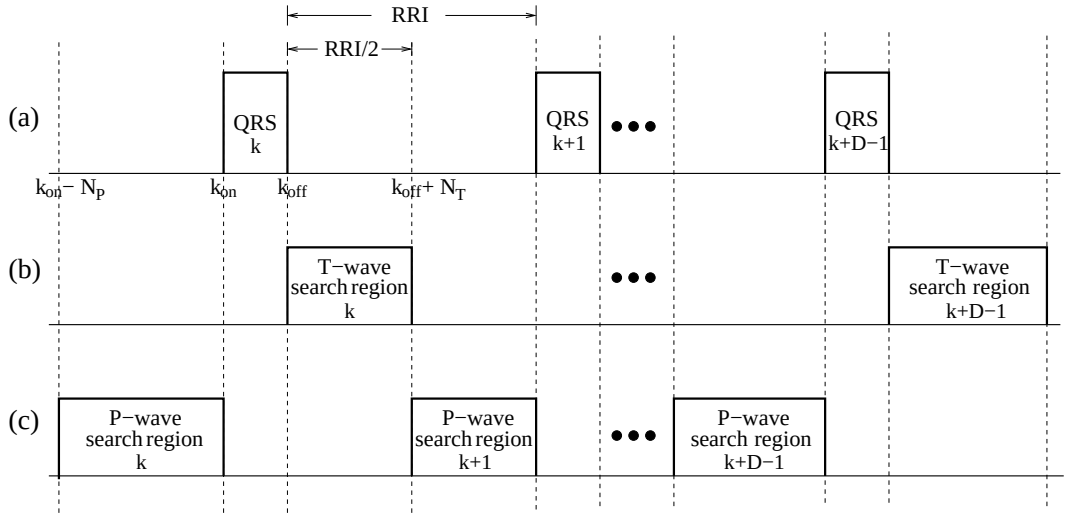


Figure 2.1: Preprocessing procedure within the D-beat processing window: (a) Located QRS complexes. (b) T wave search window. (c) P wave search window. RRI is the interval from a QRS offset to the next QRS onset, and RRI/2 is half of RRI, here we set $N_T = N_P$ =RRI/2.

2.2.2 Signal model for a D-beat processing window

For a D-beat processing window, the proposed algorithm processes T wave and P wave search windows individually using the same Bayesian inference. The following paragraphs introduce the Bayesian model applied to T wave search windows, while it should be noted that this model is compatible with minor modifications for P wave search windows as well.

Deconvolution models have been widely used in many signal processing applications including signal segmentation [DTD07], layer detection [KNHH09], and EMG signal decomposition [GCF09]. Following these ideas, we propose to model the P and T waves in ECG signal by using a deconvolution model with deterministic constraints. Note that a similar Bayesian framework has been previously proposed by Kail *et al.* to deal with OCT signals. Directly inspired from all these previous deconvolution models and especially based on the previous work of Kail *et al.* [KNHH09], signals in a T wave search window are assumed to be the convolution of an unknown impulse response $\boldsymbol{h} = [h_{-L}, \ldots, h_L]^T$ with an unknown sparse input sequence $\boldsymbol{u} = [u_1, \ldots, u_M]^T$ such that

$$x_k = \sum_{l=-L}^{L} h_l u_{k-l} + w_k \tag{2.1}$$

with $k \in \{1, \ldots, K\}$, where $K = M + 2L$ is the window length, x_k is the kth sample of the observed signal and w_k denotes additive Gaussian noise with zero mean and variance σ_w^2. Here, we adopt a zero boundary condition, i.e., the unknown sequence u_m is assumed to vanish for

all $m \notin \{1, \ldots, M\}$. In matrix form, (2.1) can be written as

$$\boldsymbol{x} = \boldsymbol{F}\boldsymbol{u} + \boldsymbol{w} \tag{2.2}$$

where $\boldsymbol{x} = [x_1, \ldots, x_K]^T$, $\boldsymbol{w} = [w_1, \ldots, w_K]^T$, $\boldsymbol{F}$ is the Toeplitz matrix of size $K \times M$ with first row $[\boldsymbol{h}_{0:-L}\ \boldsymbol{0}]$ and first column $\left[\boldsymbol{h}_{0:L}^T\ \boldsymbol{0}^T\right]^T$ ($\boldsymbol{0}$ is a vector of zeros with suitable length).

The sequence $\boldsymbol{u}$ can be further decomposed by using a binary indicator $b_m \in \{0, 1\}, m \in \{1, \ldots, M\}$ indicating the T wave existence ($b_m = 1$) or absence ($b_m = 0$) multiplied by weights defining the T wave amplitudes contained in the vector $\boldsymbol{a} = [a_1, \ldots, a_M]^T$, as illustrated in Fig. 2.2. Note that the impulse response $\boldsymbol{h}$ is supposed to be identical for D consecutive T wave search intervals within the processing window whereas the amplitudes in $\boldsymbol{a}$ vary from one interval to another. Consequently, (2.2) can be written as

$$\boldsymbol{x} = \boldsymbol{F}\boldsymbol{B}\boldsymbol{a} + \boldsymbol{w} \tag{2.3}$$

where $\boldsymbol{B} = \text{diag}(\boldsymbol{b})$ denotes the $M \times M$ diagonal matrix whose diagonal elements are the components of $\boldsymbol{b} = [b_1, \ldots, b_M]^T$.

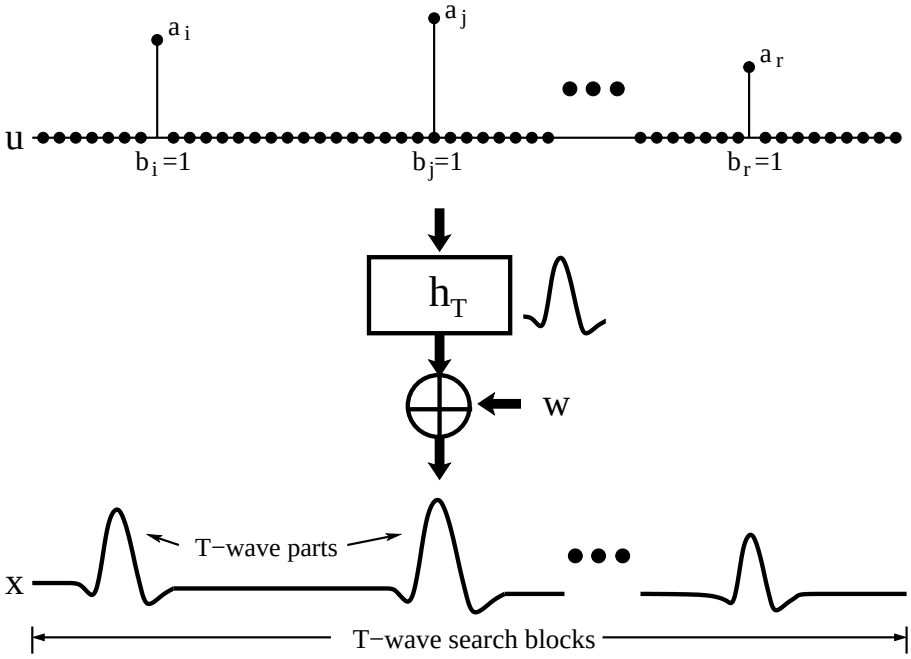


Figure 2.2: Modeling of T wave parts within the T wave search blocks.

2.2.3 Bayesian inference

The unknown parameter vector resulting from the window based signal model presented in Section 2.2.2 is denoted as $\boldsymbol{\theta} = \left[\boldsymbol{b}^T, \boldsymbol{a}^T, \boldsymbol{h}^T, \sigma_w^2\right]^T$. This work proposes to estimate $\boldsymbol{\theta}$ by using

Bayesian estimation theory. Bayesian inference on $\boldsymbol{\theta}$ is based on the posterior distribution of $\boldsymbol{\theta}$ which is related to the likelihood of the observations and the prior of the parameters via the classical Bayes rule

$$p\left(\boldsymbol{\theta}|\boldsymbol{x}\right) \propto p\left(\boldsymbol{x}|\boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right) \tag{2.4}$$

where $\propto$ means "proportional to". The likelihood $p\left(\boldsymbol{x}|\boldsymbol{\theta}\right)$ and the prior $p\left(\boldsymbol{\theta}\right)$ for the T wave delineation problem are summarized below.

Likelihood

Using (2.3) and the fact that w_k is white Gaussian, the likelihood of the observed data vector $\boldsymbol{x}$ can be expressed as

$$p\left(\boldsymbol{x}|\boldsymbol{\theta}\right) = \frac{1}{\left(2\pi\right)^{\frac{K}{2}} \sigma_w^K} \exp\left[-\frac{1}{2\sigma_w^2} \left\|\boldsymbol{x} - \boldsymbol{FBa}\right\|^2\right] \tag{2.5}$$

where $\|\boldsymbol{x}\| = \left(\boldsymbol{x}^T\boldsymbol{x}\right)^{\frac{1}{2}}$ denotes the Euclidean norm.

Prior distributions

The indicator vector $\boldsymbol{b}$ is a deterministic function of the amplitude vector $\boldsymbol{a}$. Thus, the prior probability density function (pdf) of the sparse sequence $\boldsymbol{u}$ can be decomposed as $p(\boldsymbol{b}, \boldsymbol{a}) = p\left(\boldsymbol{a}|\boldsymbol{b}\right) p\left(\boldsymbol{b}\right)$. Since there are no known relations between the sparse sequence, the pulse shape and the noise variance, the unknown parameters $(\boldsymbol{b}, \boldsymbol{a})$, $\boldsymbol{h}$ and σ_w^2 are assigned a priori statistically independent of each other such that

$$p\left(\boldsymbol{\theta}\right) = p\left(\boldsymbol{b}, \boldsymbol{a}\right) p\left(\boldsymbol{h}\right) p\left(\sigma_w^2\right) = p\left(\boldsymbol{a}|\boldsymbol{b}\right) p\left(\boldsymbol{b}\right) p\left(\boldsymbol{h}\right) p\left(\sigma_w^2\right). \tag{2.6}$$

We will now discuss the prior distributions of each of these parameters. As a classical choice to sparsity modeling, the binary indicator b_k is modeled as a Bernoulli sequence

$$b_k \sim \mathcal{B}e\left(\lambda\right) \tag{2.7}$$

where $\mathcal{B}e\left(\lambda\right)$ is the Bernoulli distribution with parameter λ such that $P\left[b_k = 1\right] = \lambda$ and $P\left[b_k = 0\right] = 1 - \lambda$. Because of the physiological constraint of ECGs, each ECG cycle should contain at most one T wave and P wave. Successive T waves can only appear in search intervals located in the right-hand neighborhoods of each QRS offset (whereas P waves are located in the left-hand neighborhoods of each QRS onset). Thus the T wave indicator vector $\boldsymbol{b}$ cannot have two elements $b_k = 1$ and $b_{k'} = 1$ closer than a *minimum-distance* d, where d depends on the RR interval length (d is set to RRI/2 in our simulations). The prior of $\boldsymbol{b}$ can then be defined as the product of the likelihood of independent Bernoulli random variables and a

minimum-distance constraint indicator function $I_C(\mathbf{b})$

$$p(\boldsymbol{b}) \propto \left[\prod_{k=1}^{K} p(b_k)\right] I_C(\mathbf{b}) = \lambda^{\|\boldsymbol{b}\|^2}(1-\lambda)^{K-\|\boldsymbol{b}\|^2} I_C(\mathbf{b}) \tag{2.8}$$

where C represents the minimum-distance constraint, i.e., $I_C(\boldsymbol{b}) = 1$ if $\boldsymbol{b} \in C$ and $I_C(\boldsymbol{b}) = 0$ if $\boldsymbol{b} \notin C$.

Similar to the indicators, amplitudes at different k are also modeled as statistically independent

$$p(\boldsymbol{a}|\boldsymbol{b}) = \prod_{k=1}^{K} p(a_k|b_k) \tag{2.9}$$

For the T wave amplitudes a_k at those k where $b_k = 1$ (note that $b_k = 0$ implies $a_k = 0$), we choose a zero-mean Gaussian prior, i.e.

$$p(a_k|b_k) = \begin{cases} \delta(a_k) & \text{if } b_k = 0 \\ \mathcal{N}(0, \sigma_a^2) & \text{if } b_k = 1 \end{cases} \tag{2.10}$$

where $\delta(\cdot)$ is the Dirac delta function. Choosing a zero-mean Gaussian prior allows for both positive and negative amplitudes. It follows that the sparse sequence $u_k = b_k a_k$ is a Bernoulli-Gaussian (BG) sequence with minimum-distance constraints. Consequently, the T wave detection problem can be seen as a BG blind deconvolution with deterministic local constraints as in [KNHH09, KTHD10].

The impulse response vector $\boldsymbol{h}$ is assigned a Gaussian prior

$$p(\boldsymbol{h}) = \mathcal{N}\left(\mathbf{0}, \sigma_h^2 \boldsymbol{I}_{2L+1}\right) \tag{2.11}$$

where $\boldsymbol{I}_{2L+1}$ denotes the identity matrix of size $(2L+1) \times (2L+1)$. Choosing conjugate Gaussian prior for $\boldsymbol{h}$ simplifies the algorithm since the resulting conditional distribution is also Gaussian. Here, σ_a^2 and σ_h^2 are fixed hyperparameters. The impulse response is normalized to avoid scale ambiguity (different values of amplitude and impulse response could provide the same convolution results) such that $\sigma_h^2 = 1$. Moreover, the proposed algorithm normalizes the ECG signals with different amplitude resolutions by their maximum R-peak amplitude such that $\sigma_a^2 = 1$ can cover all possible amplitude values.

The noise variance σ_w^2 is assigned an inverse gamma prior

$$p(\sigma_w^2) = \mathcal{IG}(\xi, \eta) = \frac{\eta^\xi}{\Gamma(\xi)} \frac{1}{(\sigma_w^2)^{\xi+1}} \exp\left(-\frac{\eta}{\sigma_w^2}\right) I_{\mathbb{R}^+}(\sigma_w^2) \tag{2.12}$$

where $\Gamma(\xi)$ is the gamma function, $I_{\mathbb{R}^+}(x)$ is the unit step function, and ξ and η are fixed hyperparameters providing a vague prior. The inverse gamma distribution is convenient (and commonly used in similar contexts) because it is the conjugate prior of the Gaussian likelihood function [DTD07].

Posterior distribution

The posterior distribution of the unknown parameter vector $\boldsymbol{\theta}$ can be computed as follows

$$\begin{aligned} p(\boldsymbol{\theta}|\boldsymbol{x}) &\propto p(\boldsymbol{x}|\boldsymbol{\theta})\, p(\boldsymbol{a}|\boldsymbol{b})\, p(\boldsymbol{b})\, p(\boldsymbol{h})\, p\left(\sigma_w^2\right) \\ &\propto \exp\left[-\frac{\|\boldsymbol{x}-\boldsymbol{F}\boldsymbol{B}\boldsymbol{a}\|^2+\eta}{\sigma_w^2}-\frac{\|\boldsymbol{h}\|^2}{\sigma_h^2}-\frac{\|\boldsymbol{a}\|^2}{\sigma_a^2}\right]\lambda^{\|\boldsymbol{b}\|^2}(1-\lambda)^{K-\|\boldsymbol{b}\|^2} I_C(\mathbf{b}) I_{\mathbb{R}^+}(\sigma_w^2). \quad (2.13) \end{aligned}$$

The usual Bayesian estimators related to this posterior are the minimum mean square error (MMSE) estimator and the maximum a posteriori (MAP) estimator [Van68]. Due to the complexity of the posterior distribution given by (2.13), it is difficult to obtain closed-form expressions for these estimators. Markov chain Monte Carlo (MCMC) methods are often used when the analytic expression of a detector or estimator is too complex to be calculated directly. The MMSE or MAP estimators of the unknown parameters are then approximated by a sample-based scheme (as, e.g., in Section 2.3.3), where a sample is generated by means of an ergodic Markov chain whose stationary distribution is the target distribution from which the sample realizations are to be drawn. The main principles of MCMC methods can be found in [RC04]. In this work, we concentrate on a particular MCMC method known as the Gibbs sampler. An extension of this method referred to as partially collapsed Gibbs sampler (PCGS) was studied in [VP08]. The idea of the PCGS is to introduce three modifications to the classical Gibbs sampler (marginalization, trimming and permutation) without changing the stationary distribution of the Markov chain. The PCGS's flexibility regarding the choice of the sampling distributions makes it applicable to many cases in which the posteriors required by the Gibbs sampler cannot be calculated analytically. In the next section, we first review the Gibbs sampler and the PCGS principles. Then, we propose a PCGS that exhibits fast convergence in the presence of a minimum distance constraint to generate samples distributed according to (2.13).

2.3 Partially collapsed Gibbs sampler

2.3.1 A review of Gibbs sampler and PCGS

MCMC methods can be used to sample from probability distributions by constructing a Markov chain that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the desired distribution. The quality of the sample improves as a function of the number of steps. Typical use of MCMC sampling is to approximate the target distribution. The most common application of these algorithms is numerically calculating multi-dimensional integrals. In these methods, an ensemble of "walkers" moves around randomly. At each point where the walker steps, the integrand value at that point is counted toward the integral. The walker then may make a number of tentative steps around the area, looking for a place with reasonably high contribution to the integral to move into next. In Bayesian statistics, multi-dimensional integrals often arises. As a consequence, MCMC methods are widely used to resolve complex Bayesian inference problems [RC04].

Gibbs sampler

The Gibbs sampler is an example of MCMC methods. The algorithm was described by Stuart and Donald Geman in 1984 [GG84]. Gibbs sampler is applicable when the joint distribution is not known explicitly or is difficult to sample from directly, but the conditional distribution of each variable is known and is easy to sample from. The Gibbs sampler generates an instance from the distribution of each variable in turn, conditional on the current values of the other variables. It can be shown that the sequence of samples constitutes a Markov chain, and the stationary distribution of that Markov chain is just the sought-after joint distribution. The Gibbs sampler is particularly well-adapted to sampling the posterior distribution of a Bayesian network, since Bayesian networks are typically specified as a collection of conditional distributions.

The point of Gibbs sampling is that given a multivariate distribution it is simpler to sample from a conditional distribution than to marginalize by integrating over a joint distribution. Consider a random vector $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_L]^T$, and let $\boldsymbol{\theta}_{\sim l}$ denote $\boldsymbol{\theta}$ without the lth entry θ_l. To obtain realizations (samples) from the joint distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$, the Gibbs sampler iteratively samples each θ_l from $p(\theta_l|\boldsymbol{\theta}_{\sim l}, \boldsymbol{x})$ in an arbitrary order. That is, sample each variable from the distribution of that variable conditioned on all other variables, making use of the most recent values and updating the variable with its new value as soon as it has been sampled (see Algorithm 1). At the tth iteration of the Gibbs sampler, L such sampling steps produce a new realization $\boldsymbol{\theta}^{(t)}$ from $p(\boldsymbol{\theta}|\boldsymbol{x})$. This strategy is known to converge to the target distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$, which is the stationary distribution of the underlying Markov chain. Since the initialization may strongly influence the first few realizations, only the realizations after a certain "burn-in period" are used in the sample. The main advantages of the Gibbs sampler are the generality of its formulation and the fact that it circumvents the "curse of dimensionality." However, a known weakness is that statistical dependencies between (some of) the θ_l tend to result in slow convergence of the Markov chain to its stationary distribution [GIC11].

Algorithm 1 Prototype Gibbs sampler.

{Initialization:}
Set appropriate initial values $\boldsymbol{\theta}^{(0)}$.
{Iterations:}
for $t = 1, 2, \ldots, N$ **do**
  for $l = 1, 2, \ldots, L$ **do**
    {Step l of the Gibbs sampler:}
    Sample $\theta_l^{(t)}$ from $p(\theta_l^{(t)}|\theta_1^{(t)}, \ldots, \theta_{l-1}^{(t)}, \theta_{l+1}^{(t-1)}, \ldots, \theta_L^{(t-1)}, \boldsymbol{x})$
  end for
end for

PCGS

The PCGS is an extension of the Gibbs sampler. Here we present three basic tools for constructing a PCGS. They are designed to ensure that the resulting PCGS converges quickly to the target distribution.

Marginalization. The first step in constructing a PCGS is to marginalize some components of $\boldsymbol{\theta}$ out of some steps of the sampler. Instead of sampling only the entry θ_l in step l, some other entries may be sampled along with θ_l instead of being conditioned upon. Let $\mathcal{J}(l) \subseteq \{1, \ldots, L\}$, and let the vectors $\boldsymbol{\theta}_{\mathcal{J}(l)}$ and $\boldsymbol{\theta}_{\sim\mathcal{J}(l)}$ contain the entries of $\boldsymbol{\theta}$ indexed by the subset $\mathcal{J}(l)$ and by its complement $\{1, \ldots, L\} \setminus \mathcal{J}(l)$, respectively. Then step l in Algorithm 1 may sample from $p(\boldsymbol{\theta}_{\mathcal{J}(l)}|\boldsymbol{\theta}_{\sim\mathcal{J}(l)}, \boldsymbol{x})$ instead of $p(\theta_l|\theta_{\sim l}, \boldsymbol{x})$. This can improve the convergence rate significantly, especially when there are strong dependencies among certain subsets of $\boldsymbol{\theta}$.

Trimming. Trimming here means discarding a subset of the components that were to be sampled in one or more steps of a Gibbs sampler. With the marginalization, some $\mathcal{J}(l)$ for different l may overlap. That is within one entire PCGS iteration, some θ_l are sampled several times. If a θ_l is sampled several times in consecutive steps, only the last value is relevant, since the other values are never used. Such unused entries can thus be dropped from the respective sampling distribution. This can formulated as follows: for any $l \in 1, \ldots, L-1$, let the vector $\boldsymbol{\theta}_{\mathcal{J}'(l)}$ contain those entries of $\boldsymbol{\theta}_{\mathcal{J}(l)}$ that are not contained in $\boldsymbol{\theta}_{\mathcal{J}(l+1)}$, i.e., $\mathcal{J}'(l) = \mathcal{J}(l) \setminus \mathcal{J}(l+1)$. Then step l may sample from $p(\boldsymbol{\theta}_{\mathcal{J}'(l)}|\boldsymbol{\theta}_{\sim\mathcal{J}(l)}, \boldsymbol{x})$ instead of $p(\boldsymbol{\theta}_{\mathcal{J}(l)}|\boldsymbol{\theta}_{\sim\mathcal{J}(l)}, \boldsymbol{x})$, which may reduce the complexity of the sampling steps. Note that the distributions used for sampling are generally no longer conditional distributions associated with the full joint distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$, but conditional distributions associated with certain marginal distributions of $p(\boldsymbol{\theta}|\boldsymbol{x})$.

Permutation. In the case of a L-step Gibbs sampler (as, e.g., in Algorithm 1), the steps can be reordered into $L!$ possible permutations. Our goal in permuting the steps is to arrange them so that as many of the marginalized components as possible are intermediate quantities that are not conditioned on in subsequent steps. After trimming, permutations are only allowed if they preserve the justification of the trimming already applied.

These modifications do not change the stationary distribution of the Markov chain [VP08].

The PCGS's flexibility regarding the choice of the sampling distributions makes it applicable to many cases in which the posteriors required by the Gibbs sampler cannot be calculated analytically. The following paragraphs propose a PCGS to obtain a sample $\mathcal{S} \triangleq \left\{\boldsymbol{b}^{(t)}, \boldsymbol{a}^{(t)}, \boldsymbol{h}^{(t)}, \sigma_n^{2(t)}\right\}_{t=1}^{N}$ from the posterior distribution (2.13). From this sample, the unknown parameters $\boldsymbol{b}$, $\boldsymbol{a}$, $\boldsymbol{h}$, and σ_n^2 can be detected or estimated.

2.3.2 Proposed PCGS for P and T wave analysis

Reference sampler

To obtain samples from $p\left(\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}, \sigma_w^2|\boldsymbol{x}\right)$, the Gibbs sampler that does not take into account the minimum-distance constraint iteratively generates samples from $p\left(b_k|\boldsymbol{b}_{\sim k}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$, $p\left(\boldsymbol{a}|\boldsymbol{b}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$, $p\left(\boldsymbol{h}|\boldsymbol{a}, \boldsymbol{b}, \sigma_w^2, \boldsymbol{x}\right)$ and $p\left(\sigma_w^2|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{h}, \boldsymbol{x}\right)$. This can be summarized in Algorithm 2.

Algorithm 2 One reference sampler iteration:

for $k = 1, \ldots, K$ **do**
  sample b_k from $p\left(b_k|\boldsymbol{b}_{\sim k}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$
end for
sample $\boldsymbol{a}$ from $p\left(\boldsymbol{a}|\boldsymbol{b}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$
sample $\boldsymbol{h}$ from $p\left(\boldsymbol{h}|\boldsymbol{b}, \boldsymbol{a}, \sigma_w^2, \boldsymbol{x}\right)$
sample σ_w^2 from $p\left(\sigma_w^2|\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}, \boldsymbol{x}\right)$

The reference sampler here is a PCGS, not a classical Gibbs sampler, because the sampling distribution for b_k is not a conditional distribution associated with the full joint posterior (since it is not possible to sample according to $p\left(b_k|\boldsymbol{b}_{\sim k}, \boldsymbol{a}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$, the parameter vector $\boldsymbol{a}$ is marginalized out). However, as pointed out in [KTHD10], this reference sampler is poorly suited to problems with local constraints because a constraint that excludes parts of the hypothesis space may even inhibit convergence to $p(\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}, \sigma_w^2|\boldsymbol{x})$ altogether.

Proposed PCGS

In order to improve convergence of the sampler, we propose to split the unknown parameter vector $\boldsymbol{\theta}$ into two parts, i.e., $(\boldsymbol{b}, \boldsymbol{a})$ that contains the constrained parameters and $(\boldsymbol{h}, \sigma_w^2)$ that contains the unconstrained parameters. To accelerate the convergence of the Gibbs sampler, we propose a PCGS that takes into consideration the local constraints affecting $\boldsymbol{b}$ and $\boldsymbol{a}$. More precisely, denote as $J_d\left(k\right)$ a right-hand neighborhood of k with length d, i.e., $J_d\left(k\right) = \{k, \ldots, k+d-1\}$. We recall that d is the minimum-distance constraint: no consecutive non-zero wave indicators can be nearer than d. This neighborhood divides the wave indicator vector $\boldsymbol{b}$ and the wave amplitude vector $\boldsymbol{a}$ into two parts $\boldsymbol{b}_{J_d(k)} = [b_k, \ldots, b_{k+d-1}]^T$, $\boldsymbol{b}_{\sim J_d(k)} = [b_1, \ldots, b_{k-1}, b_{k+d}, \ldots, b_K]^T$ and $\boldsymbol{a}_{J_d(k)} = [a_k, \ldots, a_{k+d-1}]^T$, $\boldsymbol{a}_{\sim J_d(k)} = [a_1, \ldots, a_{k-1}, a_{k+d}, \ldots, a_K]^T$. The proposed PCGS iteratively generates b_k and a_k according to the conditional distributions $p\left(b_k|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$ and $p\left(a_k|b_k, \boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$, which is equivalent to

jointly sampling (b_k, a_k) from $p\left(b_k, a_k|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$.

The proposed PCGS is different from the reference sampler in Algorithm 2 because the sampling distribution $p\left(b_k, a_k|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$ is not a conditional distribution associated with the joint posterior $p\left(\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}, \sigma_w^2|\boldsymbol{x}\right)$. Rather, it is a conditional distribution associated with $p\left(\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}, \sigma_w^2|\boldsymbol{x}\right)$ marginalized with respect to all parameters in the neighborhood $J_d(k)$ without b_k and a_k, i.e., marginalized with respect to $J_d\left(k\right) \setminus k = \{k+1, \ldots, k+d-1\}$. Consequently, $\boldsymbol{b}_{J_d(k)}$ and $\boldsymbol{a}_{J_d(k)}$ are not contained in the condition for (b_k, a_k). This difference from the Gibbs sampler is essential, since it gives the proposed PCGS the freedom required to explore the restricted hypothesis space efficiently.

To see that the proposed algorithm is a valid PCGS, consider a sampler that samples $(\boldsymbol{b}_{J_d(k)}, \boldsymbol{a}_{J_d(k)})$ from $p\left(\boldsymbol{b}_{J_d(k)}, \boldsymbol{a}_{J_d(k)}|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$. All the sampling distributions in this sampler are conditional distributions associated with the full joint posterior, and the proposed PCGS can be seen as a trimmed version of it. The trimming can be justified because all elements of $J_d\left(k\right)$ except k itself are also contained in $J_d\left(k+1\right)$. Thus, they will be sampled again after the current step before being conditioned upon, which makes them eligible for trimming. This also explains why it is unnecessary to sample parameters indexed by $\sim J_d\left(k\right)$.

The resulting algorithm is summarized in Algorithm 3 whereas the different conditional distributions are derived in the Appendix A and detailed below. Note that a similar PCGS was previously proposed in [KTHD10] for optical coherence tomography and its convergence was investigated in [VP08].

Algorithm 3 One proposed PCGS iteration:

```
set k = 1
while k ≤ K do
  sample b_k from p(b_k | b_{~J_d(k)}, a_{~J_d(k)}, h, σ_w^2, x)
  if b_k = 1 then
    sample a_k from p(a_k | b_k = 1, b_{~J_d(k)}, a_{~J_d(k)}, h, σ_w^2, x)
    set b_{J_d(k)\k} = 0
    set k = k + d − 1
  end if
  set k = k + 1
end while
sample h from p(h | b, a, σ_w^2, x)
sample σ_w^2 from p(σ_w^2 | b, a, h, x)
```

Indicators. As explained previously, the sampling distribution for b_k is a conditional distribution associated with the joint posterior $p\left(\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}, \sigma_w^2|\boldsymbol{x}\right)$ marginalized with respect to the remaining parameters in the neighborhood $J_d\left(k\right) \setminus k$. While the marginalization with respect to $\boldsymbol{a}_{J_d(k)}$ is easily done in closed form, the marginalization of the joint posterior cannot be analytically obtained with respect to the discrete-valued $\boldsymbol{b}_{J_d(k)\setminus k} = [b_{k+1}, \ldots, b_{k+d-1}]^T$.

Indeed, this conditional distribution can be calculated as

$$p\left(b_k|\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_w^2,\boldsymbol{x}\right)=\sum_{\mathbf{b}_{J_d(k)\setminus k}} p\left(\boldsymbol{b}_{J_d(k)}|\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_w^2,\boldsymbol{x}\right). \tag{2.14}$$

Thus, we propose to sample $\mathbf{b}_{J_d(k)}$ from $p\left(\boldsymbol{b}_{J_d(k)}|\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_w^2,\boldsymbol{x}\right)$ and then use the b_k contained in the sample. Therefore, the sampling distribution for wave indicators is

$$p\left(\boldsymbol{b}_{J_d(k)}|\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_w^2,\boldsymbol{x}\right)\propto\sigma_1\exp\left(\frac{|\mu_1|^2}{2\sigma_1^2}\right)p\left(\boldsymbol{b}\right)$$

with

$$\begin{aligned}\sigma_1^2 &= \left(\frac{\left\|\boldsymbol{F}_{J_d(k)}\boldsymbol{b}_{J_d(k)}\right\|^2}{\sigma_w^2}+\frac{1}{\sigma_a^2}\right)^{-1}\\ \mu_1 &= \frac{\sigma_1^2\boldsymbol{b}_{J_d(k)}^T\boldsymbol{F}_{J_d(k)}^T\left(\boldsymbol{x}-\boldsymbol{F}_{\sim J_d(k)}\boldsymbol{B}_{\sim J_d(k)}\boldsymbol{a}_{\sim J_d(k)}\right)}{\sigma_w^2}\end{aligned}$$

where $\boldsymbol{F}_{J_d(k)}$ denotes the columns of $\boldsymbol{F}$ indexed by $J_d\left(k\right)$, $\mathbf{F}_{\sim J_d(k)}$ denotes $\boldsymbol{F}$ without those columns, and $\boldsymbol{B}_{\sim J_d(k)}$ denotes the diagonal matrix diag $\left(\boldsymbol{b}_{\sim J_d(k)}\right)$ (see Appendix A for computational details). Note that, in the case where we sample $b_k=1$ at position k, this $b_k=1$ will force the next $d-1$ indicators to be zero. Therefore, after generating a 1 for b_k, the subsequent indicators $b_{k+1},\ldots,b_{k+d-1}$ can be set to zero and the corresponding $d-1$ sampling steps can be skipped. This is the consequence of the minimum distance constraint which ensures that there can not be two successive T waves (or P waves) closer than a certain distance. By exploiting this property, the proposed PCGS improves the convergence rate of the sampler. This will be confirmed in the simulation section (see a convergence comparison with the reference sampler Fig. 2.11).

Amplitudes. Using the fact that $p\left(\boldsymbol{a}\right)$ is a conjugate prior, we obtain

$$p\left(a_k|b_k=1,\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_w^2,\boldsymbol{x}\right)=\mathcal{N}\left(\mu_1,\sigma_1^2\right) \tag{2.15}$$

Note that the amplitude a_k is sampled only when $b_k=1$, i.e., when a wave has been detected.

Waveform coefficients. Because $\boldsymbol{h}$ is a priori Gaussian, straightforward computations lead to

$$p\left(\boldsymbol{h}|\boldsymbol{b},\boldsymbol{a},\sigma_w^2,\boldsymbol{x}\right)=\mathcal{N}\left(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2\right) \tag{2.16}$$

with

$$\boldsymbol{\mu}_2=\frac{\boldsymbol{\sigma}_2^2\boldsymbol{U}^T\boldsymbol{x}}{\sigma_w^2},\quad\boldsymbol{\Sigma}_2=\left(\frac{\boldsymbol{U}^T\boldsymbol{U}}{\sigma_w^2}+\frac{\boldsymbol{I}_{2L+1}}{\sigma_h^2}\right)^{-1}$$

where $\boldsymbol{U}$ is the Toeplitz matrix of size $K\times(2L+1)$ with first row $[\boldsymbol{u}_{L+1:1}\ \mathbf{0}]$ and first column

$\left[\boldsymbol{u}_{L+1:K}^T \; \mathbf{0}^T\right]^T$. Note that $\boldsymbol{Uh} = \boldsymbol{FBa}$. Thus (2.3) can be represented as $\boldsymbol{x} = \boldsymbol{Uh} + \boldsymbol{w}$. As mentioned in 2.2.3, scale ambiguity inherent to the convolution model is resolved by normalizing $\boldsymbol{h}$ at every iteration.

Noise variance. The conditional distribution of the noise variance is the following inverse gamma distribution

$$p\left(\sigma_w^2|\boldsymbol{b},\boldsymbol{a},\boldsymbol{h},\boldsymbol{x}\right) = IG\left(\xi + \frac{K}{2}, \eta + \frac{1}{2}\left\|\boldsymbol{x} - \boldsymbol{FBa}\right\|^2\right). \tag{2.17}$$

An Alternative PCGS

An alternative PCGS, which marginalizes out entirely the parameter vector $\boldsymbol{a}$ in the sampling substeps for the b_k, can be formulated in Algorithm 4. Note that this alternative PCGS has been proposed previously in [KTHD10].

Algorithm 4 One alternative PCGS iteration:

set $k = 1$
while $k \leq K$ **do**
  sample b_k from $p\left(b_k|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$
  if $b_k = 1$ **then**
    sample a_k from $p\left(a_k|b_k = 1, \boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right)$
    set $\boldsymbol{b}_{J_d(k)\setminus k} = \mathbf{0}$
    set $k = k + d - 1$
  end if
  set $k = k + 1$
end while
sample $\boldsymbol{h}$ from $p\left(\boldsymbol{h}|\boldsymbol{b},\boldsymbol{a},\sigma_w^2,\boldsymbol{x}\right)$
sample σ_w^2 from $p\left(\sigma_w^2|\boldsymbol{b},\boldsymbol{a},\boldsymbol{h},\boldsymbol{x}\right)$

The sampling distribution associated to this alternative PCGS is as follows (the sampling distributions for the waveform coefficients and the noise variance parameters are the same as the previous proposed PCGS)

$$\begin{aligned} p\left(\boldsymbol{b}_{J_d(k)}|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right) &\propto |\boldsymbol{\Sigma}_3| \exp\left(\boldsymbol{\mu}_3^T \boldsymbol{\Sigma}_3^{-1} \boldsymbol{\mu}_3\right) p\left(\boldsymbol{b}\right) \\ p\left(a_k|b_k = 1, \boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_w^2, \boldsymbol{x}\right) &= \mathcal{N}\left(\boldsymbol{\mu}_3, \boldsymbol{\Sigma}_3\right) \end{aligned} \tag{2.18}$$

with

$$\begin{aligned} \boldsymbol{\Sigma}_3 &= \left(\frac{1}{\sigma_w^2}\boldsymbol{B}^T\boldsymbol{F}^T\boldsymbol{FB} + \frac{1}{\sigma_a^2}\boldsymbol{I}\right)^{-1} \\ \boldsymbol{\mu}_3 &= \frac{1}{\sigma_w^2}\boldsymbol{\Sigma}_3\boldsymbol{B}^T\boldsymbol{F}^T\boldsymbol{x}. \end{aligned}$$

Compared to the proposed PCGS, the alternative PCGS is "more collapsed" since its sampling distributions are conditioned on less parameters. Therefore, theoretically, the gain in convergence rate is slightly bigger than for the proposed PCGS. Nevertheless, the alternative PCGS is significantly more complex than the proposed PCGS because it requires inversion of $\boldsymbol{\Sigma}_3$, which is a $K \times K$ matrix.

2.3.3 Parameter estimation with the PCGS

P and T wave detection and estimation are based on the estimated joint posterior distribution of wave indicators, wave amplitudes and waveform coefficients. This posterior is computed from histograms of the samples generated by the PCGS. Unlike most of the approaches found in the literature, no rigid amplitude threshold is used to determine whether waves are representative or not.

The posterior distribution of wave indicators carries information regarding the probability of having a P or T wave at a given location. The indicator posterior probability at position k can be approximated as follows (the probability of having $b_k = 1$ is equal to $\mathrm{E}[b_k]$, where $\mathrm{E}[\cdot]$ is the mathematical expectation since b_k is a binary random variable)

$$\hat{b}_{k,\mathrm{MMSE}} = \frac{1}{N_r} \sum_{t=1}^{N_r} b_k^{(N_{\mathrm{bi}}+t)} \tag{2.19}$$

where $b_k^{(t)}$ denotes the indicator at position k generated at iteration t, while N_r is the number of iterations used for the estimation and N_{bi} is the number of burn-in iterations. Note that (2.19) is the MCMC approximation to the MMSE estimator. The detection results $\hat{\mathbf{b}}$ can be obtained with various degrees of certainty by using a local MAP strategy: since there are at most one T wave (P-wave) in each T-searching neighborhood (P-searching neighborhood), the proposed algorithm compares the highest estimated posterior probability of each neighborhood to a given probability threshold (γ_P for P-wave and γ_T for T wave) in order to decide whether it is representative or not. The value of γ_P and γ_T can be determined by studying the receiver operating characteristic (ROC) curve (see Section 2.3.5 for details). If the local MAP is higher than the threshold, the corresponding indicator location can be seen as the estimate wave location in this searching neighborhood.

For estimating the wave amplitude a_k corresponding to a position k where a P or T wave has been detected, we use the approximated MMSE estimator of a_k conditionally upon $b_k = 1$

$$\hat{a}_{k,\mathrm{MMSE}} = \frac{1}{|\mathcal{T}_k|} \sum_{t \in \mathcal{T}_k} a_k^{(N_{\mathrm{bi}}+t)} \tag{2.20}$$

where $a_k^{(t)}$ denotes the k^{th} entry of the amplitude vector $\boldsymbol{a}$ generated by the Markov chain at iteration t, $\mathcal{T}_k$ is the set of indices t of all iterations satisfying $b_k^{(t)} = 1$ excluding burn-in iterations. Note that $\hat{a}_{k,\mathrm{MMSE}}$ is calculated only when a P or T wave has been detected.

The approximated MMSE estimators of the waveform coefficients $\boldsymbol{h}$ and the noise variance

σ_n^2 are defined as

$$\hat{\boldsymbol{h}}_{\text{MMSE}} = \frac{1}{N_r} \sum_{t=1}^{N_r} \boldsymbol{h}^{(N_{\text{bi}}+t)} \tag{2.21}$$

$$\hat{\sigma}^2_{w,\text{MMSE}} = \frac{1}{N_r} \sum_{t=1}^{N_r} \left(\sigma_w^2\right)^{(N_{\text{bi}}+t)} \tag{2.22}$$

where $\boldsymbol{h}^{(t)}$ and $\left(\sigma_w^2\right)^{(t)}$ denote respectively $\boldsymbol{h}$ and σ_w^2 generated at iteration t.

2.3.4 P and T wave delineation criteria

This section introduces the strategy to determine the wave fiducial points (peak, onset and end) based on the estimates of the wave parameters. The wave peak position are obtained from the wave indicator estimates whereas the wave boundaries are computed from the estimated peak position and waveform estimates.

P and T wave peak location

One issue with the blind deconvolution model (2.1) is the so-called time-shift ambiguity [LI06]. This problem is due to the existence of several solutions of (2.1) differing by time-shift operations. Recall that the signal model in (2.1) can be written as

$$\boldsymbol{x} = \boldsymbol{h} \star \boldsymbol{u} + \boldsymbol{w} \tag{2.23}$$

where $\star$ denotes the convolution operator and $\boldsymbol{u}$ is an unknown Bernoulli-Gaussian sequence satisfying a minimum distance constraint which characterizes the wave location and amplitude, and $\boldsymbol{h}$ is the unknown impulse response associated with the waveform. The time-shift ambiguity results from the following equation

$$\boldsymbol{h} \star \boldsymbol{u} = (d_\tau \star \boldsymbol{h}) \star (d_{-\tau} \star \boldsymbol{u}), \quad \forall \tau \in \mathbb{Z} \tag{2.24}$$

where d_τ is the time delay filter of τ samples.

Following [CCL96], we propose to compensate the time-shift ambiguity by including a time-shift step after generating the waveform samples in the Gibbs sampler to ensure that the maximum of the waveform is always at the middle of the alloted time window $\boldsymbol{h} = [h_{-L}, \ldots, h_L]^T$, i.e.,

$$|h_0| = \max_i |h_i| \,. \tag{2.25}$$

Enforcing condition (2.25) ensures the identifiability of $\boldsymbol{h}$ and $\boldsymbol{u}$. Note that an hybrid Gibbs sampling solution has been proposed in [LI06] to solve the time-shift ambiguity. Instead of enforcing constraints on the sampler, the method investigated in [LI06] removes time ambiguities by possibly shifting the samples $\boldsymbol{h}^{(t)}, (\boldsymbol{b}^{(t)}, \boldsymbol{a}^{(t)})$ w.r.t. the time index, prior to computing averages. A Metropolis-Hastings step is included in the Gibbs sampler to accept or reject the

time-shift. Note that in our application, the simple constraint enforcing has been proved to be sufficient to solve the ambiguity problem. One of the convenient consequences of this shift is that the wave indicator position indicates directly the wave peak location.

P and T wave boundaries

Since the estimated waveform $\hat{\boldsymbol{h}}_{\text{MMSE}}$ carries information regarding the wave morphology, we propose a delineation criterion which is based on the waveform estimate of each processing window. The waveform onset and end are estimated as either the first value of $\hat{\boldsymbol{h}}_{\text{MMSE}}$ below appropriate thresholds denoted as ζ_{Pon} and ζ_{Poff} for P waves and ζ_{Ton} and ζ_{Toff} for T waves, (see left part of Fig. 2.3) or the first local minimum of $\hat{\boldsymbol{h}}_{\text{MMSE}}$ (see right part of Fig. 2.3).

Since there is no universal rule to locate onsets and ends of waves, the delineation thresholds have been obtained by minimizing the error between estimates and published annotations. The following results have been obtained for the QT database

$$\begin{aligned} \zeta_{\text{Ton}} &= 0.02\text{max}\left(\hat{\boldsymbol{h}}_T\right), \quad \zeta_{\text{Toff}} = 0.1\text{max}\left(\hat{\boldsymbol{h}}_T\right) \\ \zeta_{\text{Pon}} &= 0.05\text{max}\left(\hat{\boldsymbol{h}}_P\right), \quad \zeta_{\text{Poff}} = 0.1\text{max}\left(\hat{\boldsymbol{h}}_P\right). \end{aligned}$$

The general flowchart for the proposed algorithm including preprocessing, PCGS and wave delineation is shown in Fig. 2.4.

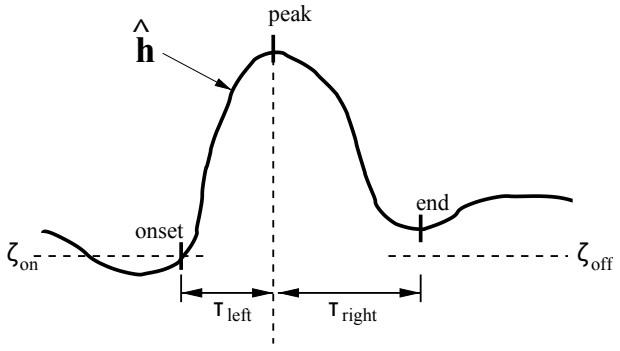


Figure 2.3: Parameters of the wave delineation method.

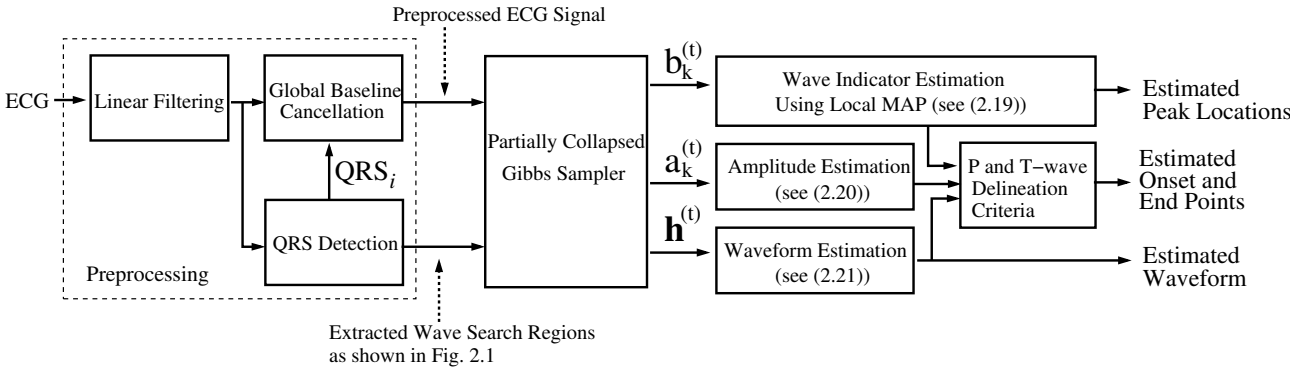


Figure 2.4: Block diagram for the PCGS P and T wave delineation algorithm.

2.3.5 Simulation results

Many simulations have been conducted to validate the proposed algorithm. First, we show some posterior distributions and estimation results for one typical example. Then, graphical evaluations and analytical results on an entire database are provided. Usually, the validation of the ECG wave detector or delineator is done using a manually annotated database. In this thesis, we use one of the easily-available standard databases, namely the QT database (QTDB) [LMGM97]. The QTDB includes 105 records from the widely used MIT-BIH Arrhythmia database (MITBIH), the European ST-T database and some other well known databases. This database was developed with the purpose of providing a wave limit validation reference.

One typical example

The first simulations have been obtained by applying the proposed algorithm on the dataset "sele0136" of QTDB. This example has been chosen because signals from this data set present rhythm changes with obvious amplitude variations. The processing window length D has been set to 10 beats to cope with the pseudo-stationary nature of the ECG as in [MO04]. For each P or T wave search block, we have generated $N_{\mathrm{r}} = 100$ realizations according to the priors given in Section 2.2.3 with $\sigma_a^2 = 1$, $\sigma_h^2 = 1$, $\xi = 11$ and $\eta = 0.5$ (these are fixed hyperparameters that provide a noninformative prior). The value of λ in (2.7) has been fixed by dividing the number of R peaks within the processing window by the window length K. We considered $N_{\mathrm{bi}} = 40$ burn-in iterations and 60 iterations to compute the estimates[1]. Note that running 100 iterations of the proposed algorithm for a 10-beat ECG block sampled at $Fs = 250$Hz (i.e., ECG signals lasting about 10 seconds) takes approximately 11 seconds for a MATLAB® implementation on a 3.0-GHz Pentium IV. However, these codes can be further optimized and converted to low-level languages for clinical use.

As mentioned before, the estimates of the unknown parameters are derived from their posterior distributions. Fig. 2.5 (b) shows the posterior distributions of P and T wave indicator

[1]To determine the values of N_{bi} and N_{r}, we have implemented the multivariate potential scale reduction factor [BG98] (see Fig. 2.11 for details).

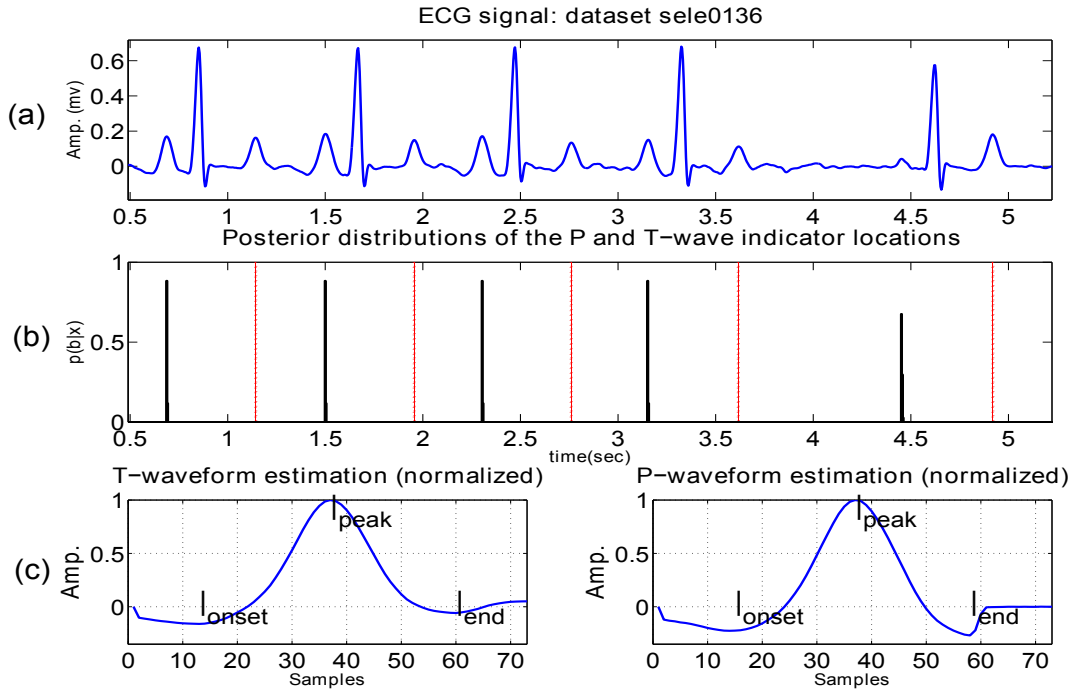


Figure 2.5: (a) an ECG signal portion from the QTDB "sele0136"; (b) posterior distributions of the P and T wave indicator locations $p(\boldsymbol{b}_{\mathrm{P}})$ (in black) and $p(\boldsymbol{b}_{\mathrm{T}})$ (in dotted red); (c) estimated P and T-waveforms.

locations estimated using the last 60 Markov chain iterations. The posterior probability is very high for most of the actual P and T wave locations except for P-wave indicators around time instant 4.45. Indeed, the algorithm seems less confident to locate P-wave indicators around this location. If we employ a simple rigid threshold on the entire block, this wave indicator could be missed in the estimation. However, with the local maximum posterior strategy explained in Section 2.3.4, a relatively low value of γ_P can be used to ensure the detection of low magnitude waves without increasing false positives. Fig. 2.7 shows that the P wave at time instant 4.45 is properly estimated.

Once we have obtained the P and T wave locations, the corresponding wave amplitudes can be estimated by using (2.20). Fig. 2.6 shows the posterior distributions $p(\mathbf{a}|\mathbf{x})$ and the estimates $\hat{\mathbf{a}}$ of P-wave amplitudes at time 1.500s, 2.304s, 3.152s and 4.452s. The estimated P and T-waveforms determined using (2.21) are presented in Fig. 2.5 (c). As explained previously, P and T wave delineation is based on these estimated waveform coefficients. The estimated P and T-waves and the delineation results of "sele0136" are illustrated in Fig. 2.7 (top) and Fig. 2.7 (bottom), respectively. The P and T wave estimation and delineation results are very satisfactory for this example.

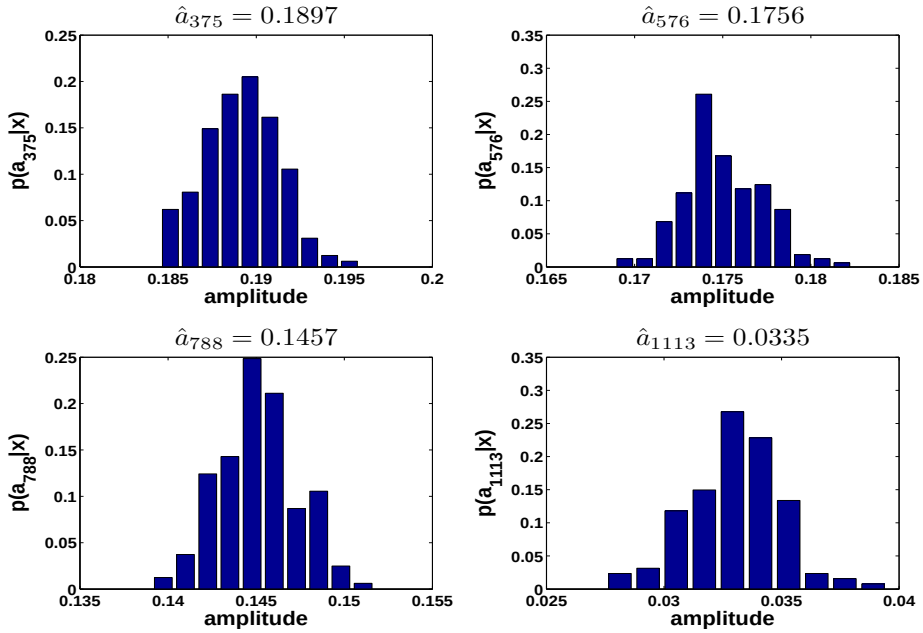


Figure 2.6: Posterior distributions of the P-wave amplitudes $p(a_k|\boldsymbol{x})$ at time instant 1.500, 2.304, 3.152, and 4.452s.

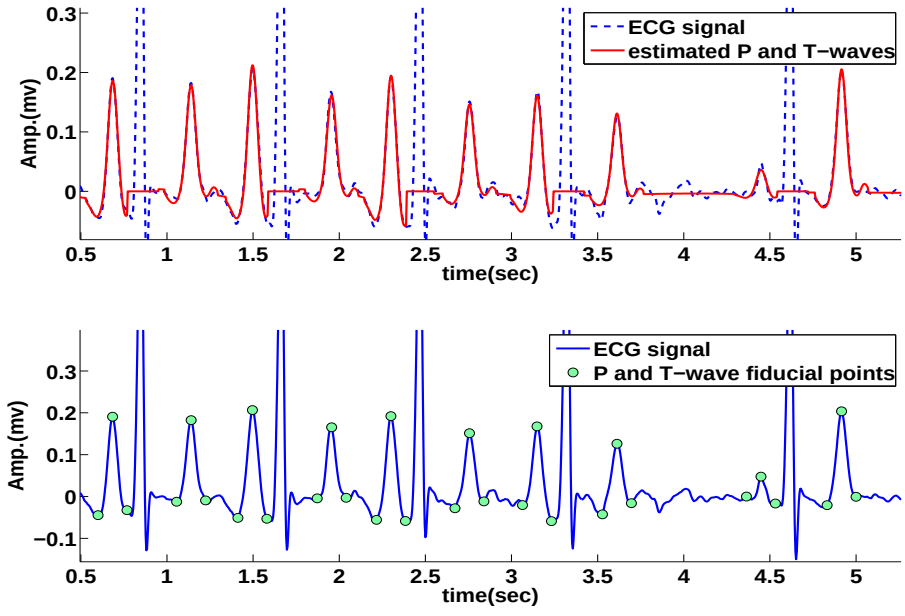


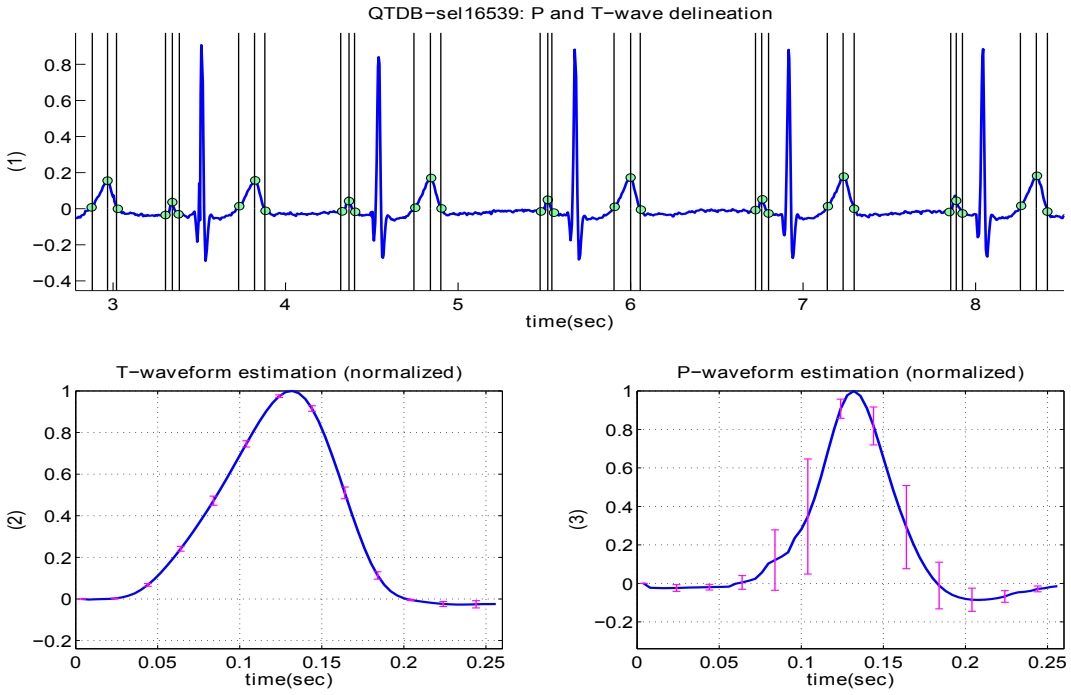
Figure 2.7: On the top, real ECG signal from the QTDB "sele0136" (dashed blue) and estimated P and T-waves (red); on the bottom, P and T wave delineation results.

P and T wave delineation for different wave morphologies

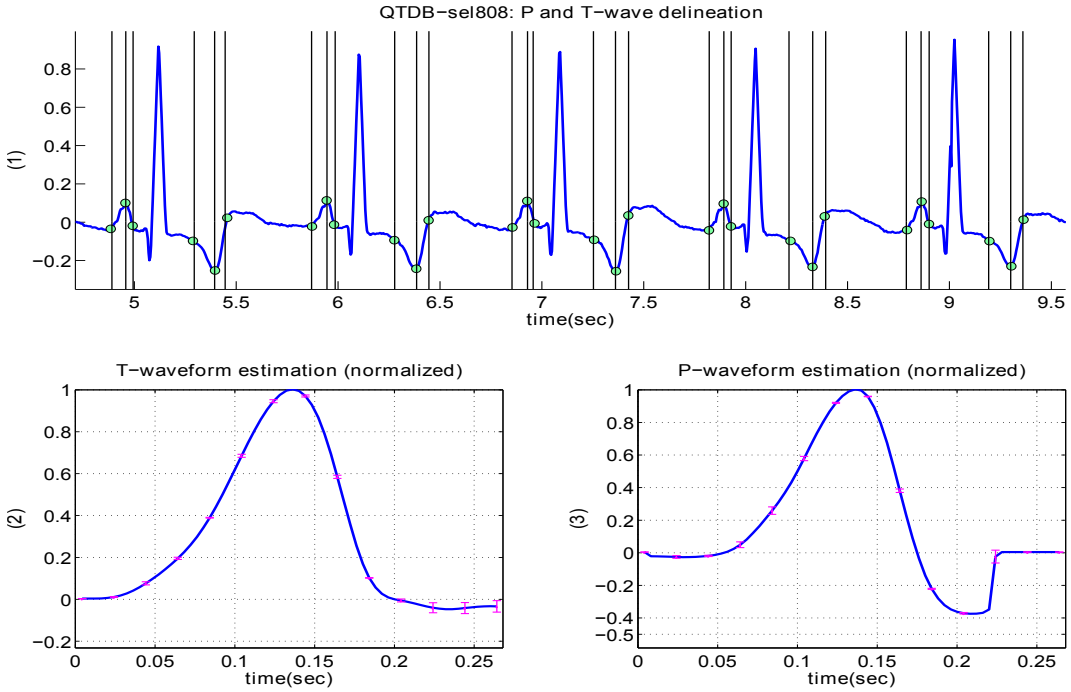
The proposed method estimates the P and T-waveform shapes pointwise for each processing window, and therefore it adapts to various wave morphologies. This section shows some representative results obtained with the proposed method for different ECGs. The first example in Fig. 2.8(a) considers the QTDB dataset "sel16539", where both P and T-waves are associated to normal patients. The delineation results for the P and T-waves are shown in Fig. 2.8(a)(1), whereas the estimated waveform of P and T wave for a processing window are presented in Fig. 2.8(a)(2) and Fig. 2.8(a)(3). All kinds of slope, magnitude and polarity for the P and T-waves are successfully detected and delineated for this example. The second example in Fig. 2.8(b) considers the QTDB dataset "sel808", where T-waves are inverted. The third example considers noisy feeble P-waves and ascending T-waves from the QTDB dataset "sele0607". The results presented in Fig. 2.9(a) show that the proposed method provides a good waveform estimation for noisy feeble waves. Fig. 2.9(b) shows an example of signals that contain premature ventricular contractions (PVCs) from the MITBIH dataset "119". As can be seen in the delineation results, the proposed method can handle non-monotonic morphological abnormalities. Note in particular that the estimated T wave of the sixth beat has been superimposed with the estimated P wave of the seventh beat, which is in agreement with the presence of a unique wave in the non QRS region between seconds 8 and 9. Furthermore, with the help of the proposed signal model, the sudden T wave amplitude inversions (the first and the last beats) have been detected, which is a nice property for the PVC detection problem.

ROC for P and T wave detection

Receiver operating characteristics have been studied to select appropriate values of the thresholds γ_P and γ_T for P and T wave detection. The ROC curve for P-wave detection has been computed using three typical datasets and three "no P-wave" datasets available in QT database. The results depicted in Fig. 2.10(a) show the good performance of the proposed detector. The threshold γ_P can be determined from a fixed probability of false alarm (PFA) using this ROC. For the QT database, we have chosen a threshold $\gamma_P = 0.4$ corresponding to a probability of detection PD=1 and PFA=0.05. Similarly, the ROC curve for the T wave detection has been calculated based on three typical datasets from the QT database and three synthesized "no T wave" datasets (since the QT database does not contain signals classified as "no T wave"). A good detection performance can also be confirmed as shown in Fig. 2.10(b). For the QT database, we have chosen $\gamma_T = 0.55$ that corresponds to PD=1 and PFA =0.01.

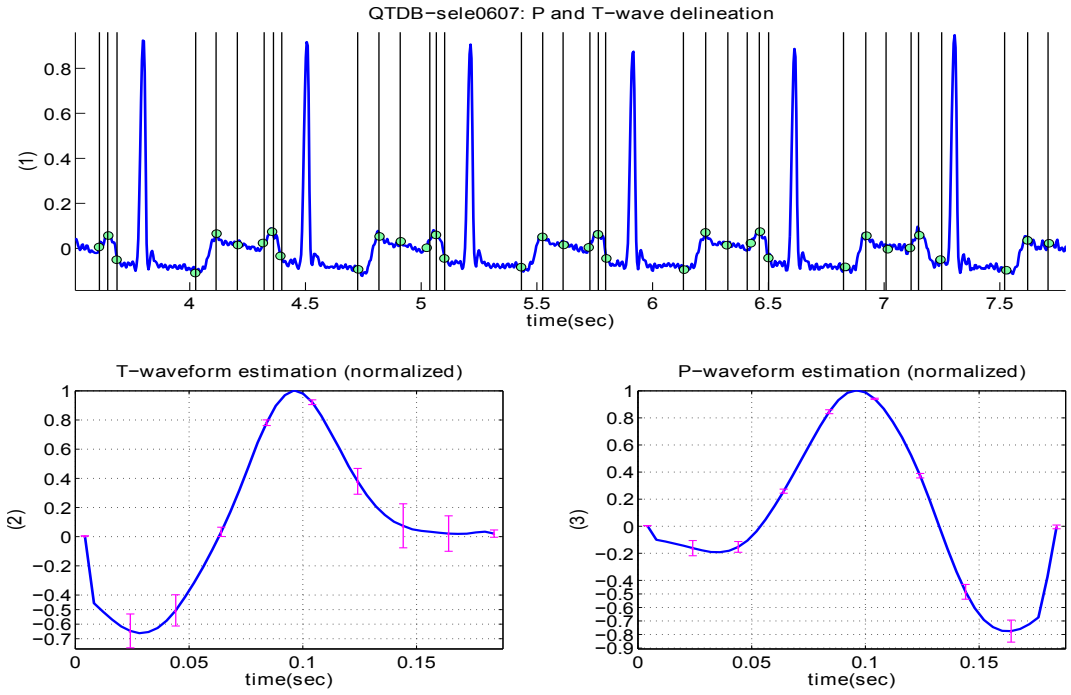


(a) QTDB-sel16539

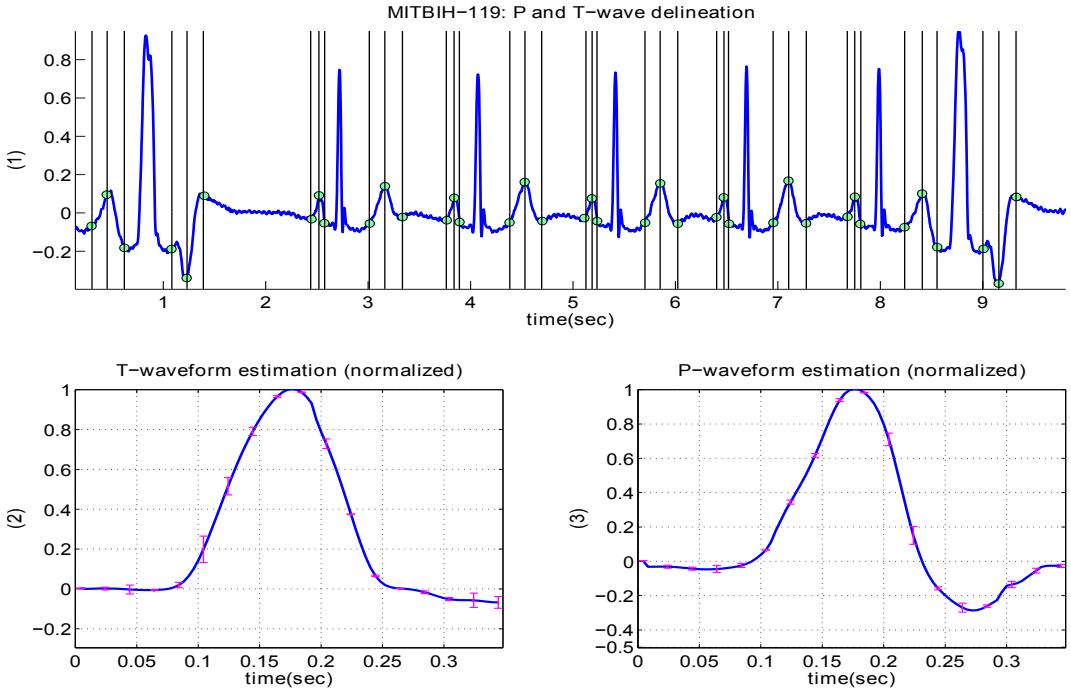


(b) QTDB-sel808

Figure 2.8: Results of processing ECG datasets (a) "QTDB-sel16539" and (b) "QTDB-sel808". Captions for each subfigure: (1) Delineation results: the vertical lines show the manual annotations by an expert and the markers show the results of the proposed algorithm. (2) Mean (in blue) and standard deviations (in pink) of the estimated T-waveform for 1 minute of signal length. (3) Mean (in blue) and standard deviations (in pink) of the estimated P-waveform for 1 minute of signal length.

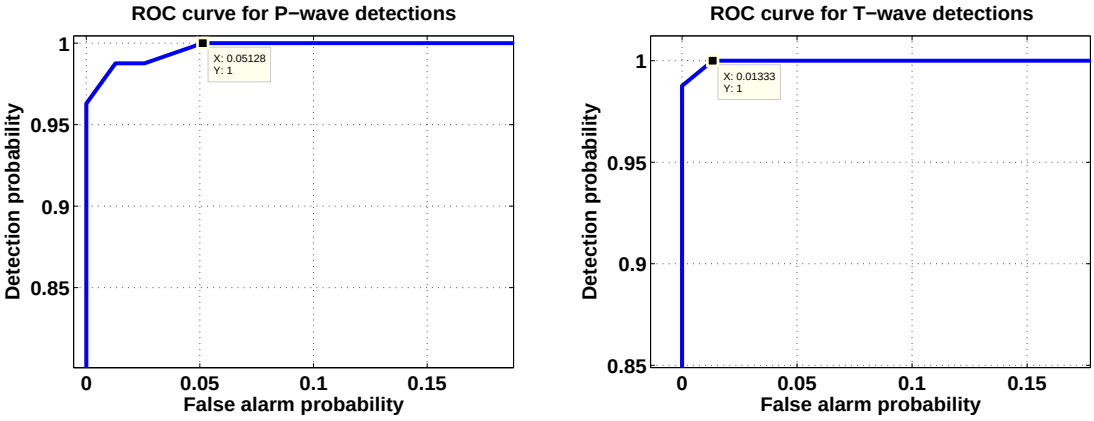


(a) QTDB-sele0607



(b) MITBIH-119

Figure 2.9: Results of processing ECG datasets (a) "QTDB-sele0607" and (b) "MITBIH-119". Captions for each subfigure: (1) Delineation results: the vertical lines show the manual annotations by an expert and the markers show the results of the proposed algorithm. (2) Mean (in blue) and standard deviations (in pink) of the estimated T-waveform for 1 minute of signal length. (3) Mean (in blue) and standard deviations (in pink) of the estimated P-waveform for 1 minute of signal length.



(a) ROC curve for P-wave detection

(b) ROC curve for T wave detection

Figure 2.10: ROC analysis for P and T wave detection.

Convergence diagnostic

A crucial issue when using MCMC methods is convergence assessment, which can help us to determine appropriate values of the numbers of burn-in iterations N_{bi} and computation iterations N_r. To monitor the convergence of the proposed PCGS, we have implemented the multivariate potential scale reduction factor (MPSRF) criterion proposed by Brooks *et al.* in [BG98]. This diagnostic is based on the comparison between estimates resulting from p parallel Monte Carlo chains of length q as follows

$$\text{MPSRF} = \frac{p-1}{p} + \frac{q+1}{q}\text{eig}\left(\boldsymbol{V}_{\text{intra}}^{-1}\boldsymbol{V}_{\text{inter}}\right) \tag{2.26}$$

where the inter-chain and intra-chain covariance matrices are defined as follows

$$\boldsymbol{V}_{\text{intra}} = \frac{1}{p\,(q-1)}\sum_{j=1}^{p}\sum_{t=1}^{q}\left(\boldsymbol{\psi}_{jt} - \overline{\boldsymbol{\psi}}_{j\cdot}\right)\left(\boldsymbol{\psi}_{jt} - \overline{\boldsymbol{\psi}}_{j\cdot}\right)^{T} \tag{2.27}$$

$$\boldsymbol{V}_{\text{inter}} = \frac{1}{p-1}\sum_{j=1}^{p}\left(\boldsymbol{\psi}_{j\cdot} - \overline{\boldsymbol{\psi}}_{\cdot\cdot}\right)\left(\boldsymbol{\psi}_{j\cdot} - \overline{\boldsymbol{\psi}}_{\cdot\cdot}\right)^{T}. \tag{2.28}$$

Here, $\left\{\boldsymbol{\psi}_{jt}^{(i)}, j = 1,\ldots,p; t = 1,\ldots,q\right\}$ denotes the ith element of the parameter vector $\boldsymbol{\psi}$ in chain j at time t, $\overline{\boldsymbol{\psi}}_{j\cdot}$ denotes the local mean of the jth chain, $\overline{\boldsymbol{\psi}}_{\cdot\cdot}$ denotes the global mean of all chains and eig $(\boldsymbol{V})$ is the largest eigenvalue of the positive-definite matrix $\boldsymbol{V}$. As an example, Fig. 2.11 shows the MPSRF criterion for signals belonging to the QTDB dataset "sele0136" with $p = 10$ independent chains of the proposed PCGS and the classical Gibbs sampler for $\boldsymbol{\psi} = [\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}]$. Since a value of MPSRF below 1.2 is recommended in [BG98], the criterion confirms a good convergence of the proposed sampler with $N_{\text{bi}} = 40$ burn-in iterations. Note that the PCGS outperforms significantly the reference GS in terms of convergence speed.

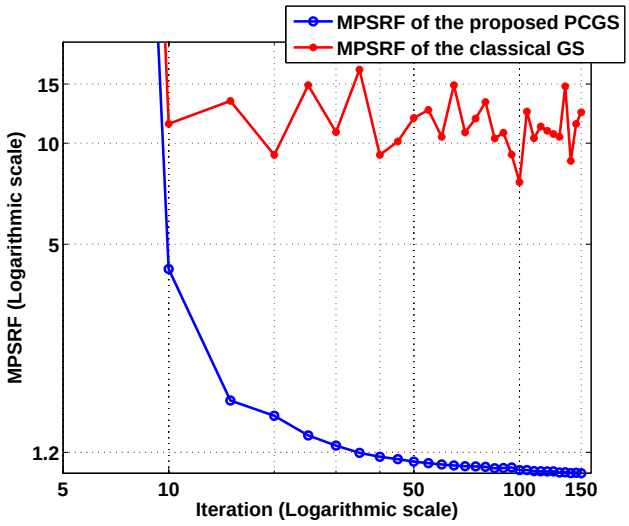
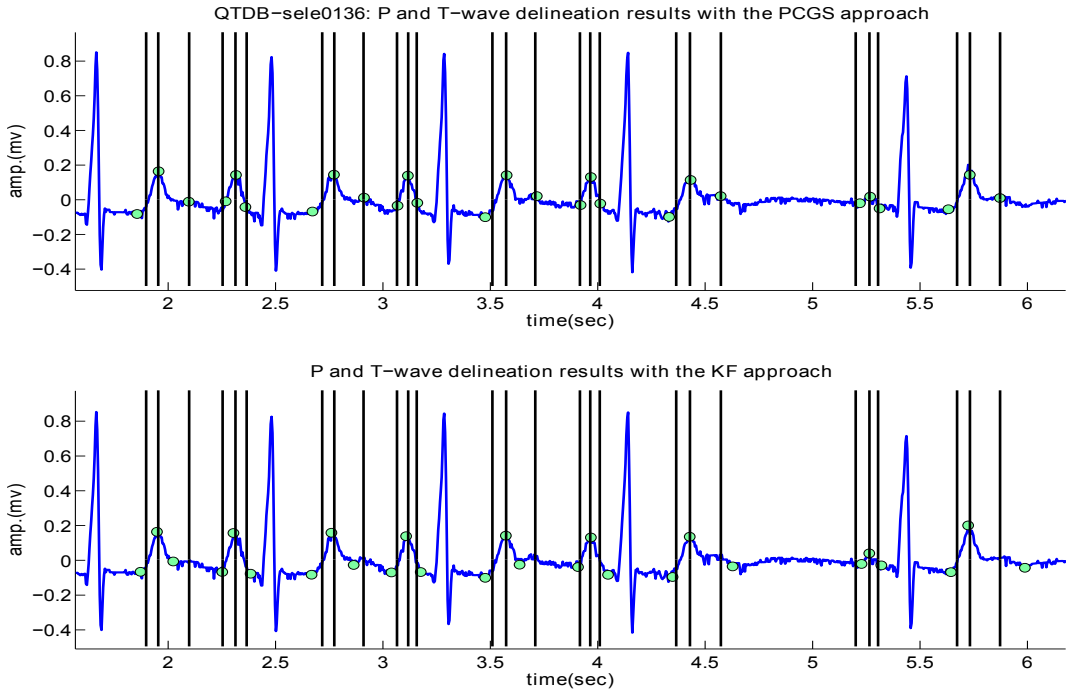


Figure 2.11: Evolution of MPSRF criterion on QTDB dataset "sele0136" for the unknown parameters ($\boldsymbol{\psi} = [\hat{\boldsymbol{b}}, \hat{\boldsymbol{a}}, \mathbf{h}]$), the proposed PCGS in blue and the reference GS in red.

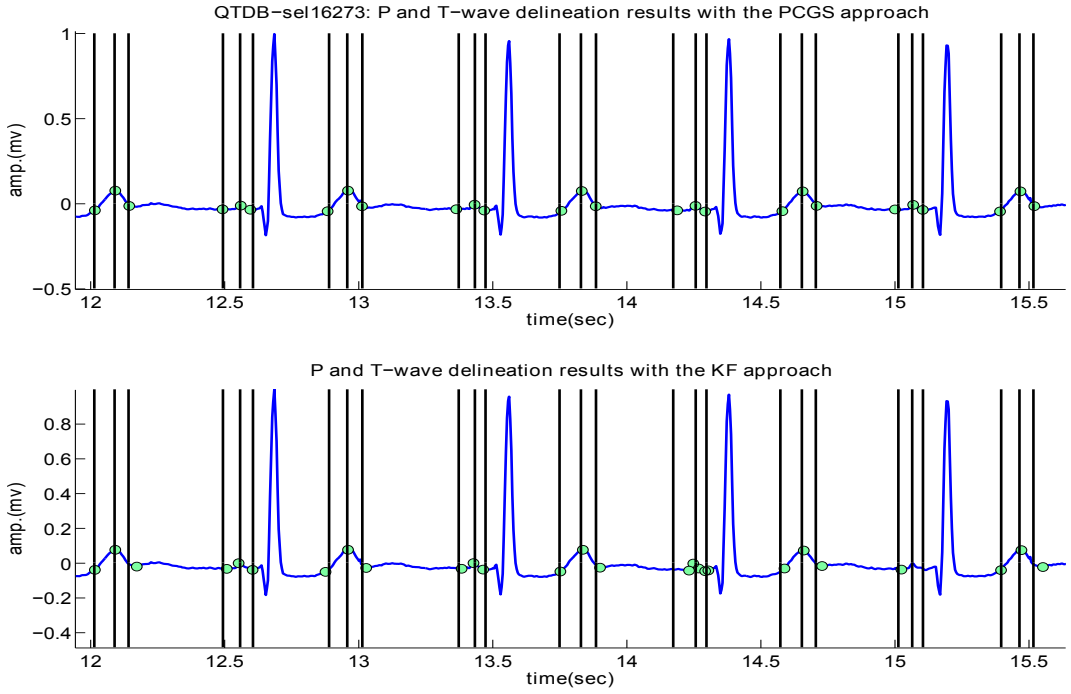
Qualitative comparisons with a Gaussian mixture model approach

As introduced in Section 1.4.3, Sayadi *et al.* have recently proposed an ECG segmentation approach based on a Gaussian mixture model and Kalman filters (KF) [SS09]. With the help of O. Sayadi, we have carried out a comparison of the PCGS based method with the KF based method. Some qualitative comparisons of the two methods on several representative datasets from the QT database are presented in Fig. 2.12-2.13. It can be seen that the PCGS estimates are closer to the manual annotations (depicted by vertical black lines) than the estimates resulting from [SS09]. Meanwhile, Fig. 2.14 shows the absolute errors of T_{peak}, T_{end}, P_{on}, P_{peak} and P_{end} for the delineation results of the two methods (the absolute error of a given parameter vector is defined as the norm of the difference between the actual value of the parameter vector and its estimate). These results have been obtained for 118 representative signals from the QT database. It appears that the proposed PCGS provides smaller errors than the KF based method of [SS09], especially when considering pathological ECG signals (the Gaussian mixture model studied in [SS09] is more appropriate to normal ECG signals).

It should also be noted that 1) the performance of the KF method depends on its initialization which can be difficult to adjust, 2) the Gaussian mixture model and the KF method of [SS09] are not really appropriate to ECG signals with abnormal rhythms (the KF method always determines a fixed number of Gaussian kernels to fit the data) contrary to the PCGS that estimates the whole P and T wave shapes. Moreover, the algorithm of [SS09] is not able to handle the absence of T or P wave in some pathological ECGs, contrary to the PCGS method. As shown in Fig. 2.13(b), the KF method detect P waves in dataset "sel50" from the QT database (see green circles before the QRS) while the cardiologists have classified all signals from this dataset as "no P wave" pathologies. Conversely, the PCGS method does not detect the P waves for these signals, as desired. Finally, it is important to note that the price to pay for the good performance of the proposed PCGS is its computational complexity which is

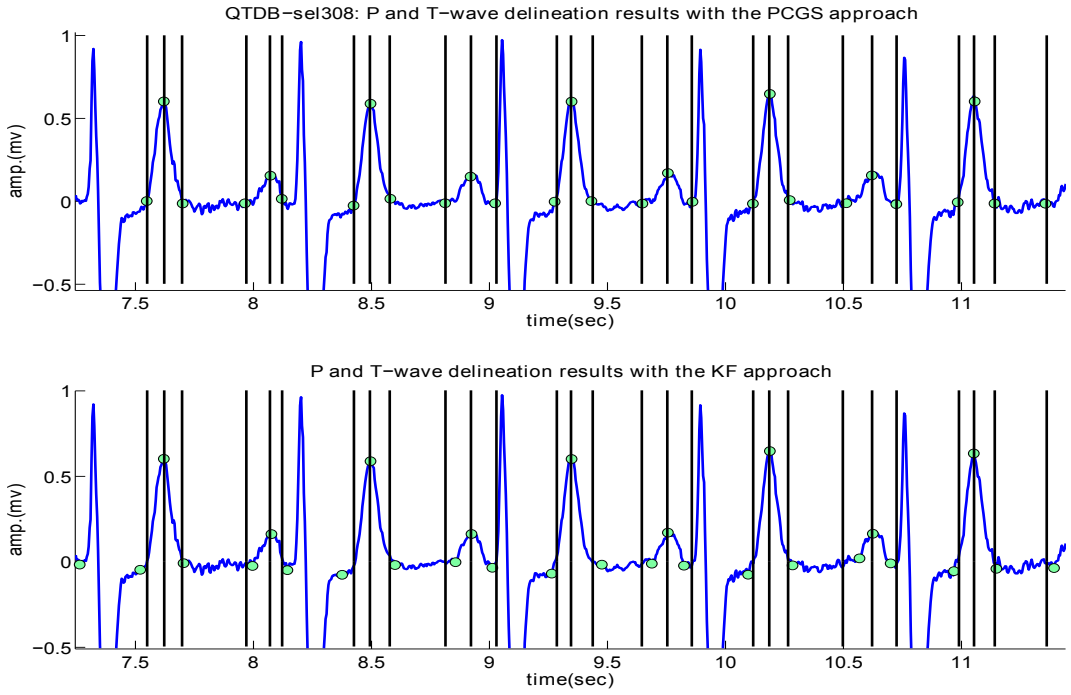


(a) QTDB-sele0136

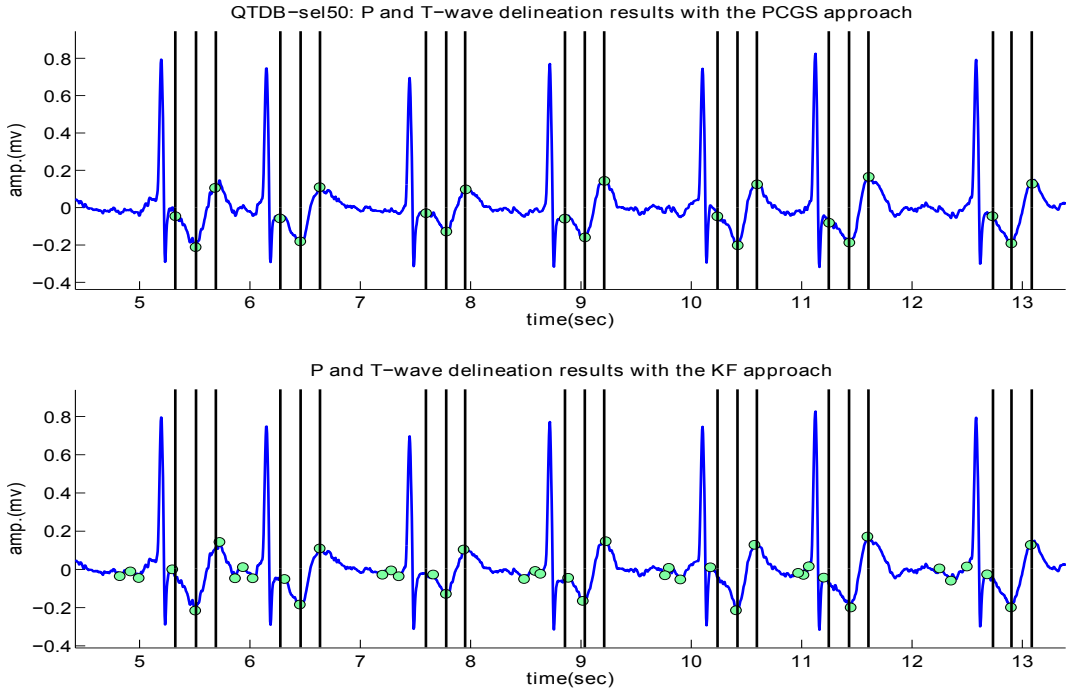


(b) QTDB-sel16273

Figure 2.12: Delineation results for QT database (a) "sele0136", which presents some rhythm changes with different amplitudes between P and T-waves; (b) "sel16273", where the P-waves have low slopes. Captions for each subfigure: The upper plot shows the results of the proposed PCGS method, and the lower plot shows the results of the method proposed in [SS09]. The vertical lines in both of two plots show the manual detection results provided by cardiologists.



(a) QTDB-sel308



(b) QTDB-sel50

Figure 2.13: Delineation results for QT database (a) "sel308", which has giant T-waves; (b) "sel50", which is classified as "no P wave" pathology. Captions for each subfigure: The upper plot shows the results of the proposed PCGS method, and the lower plot shows the results of the method proposed in [SS09]. The vertical lines in both of two plots show the manual detection results provided by cardiologists.

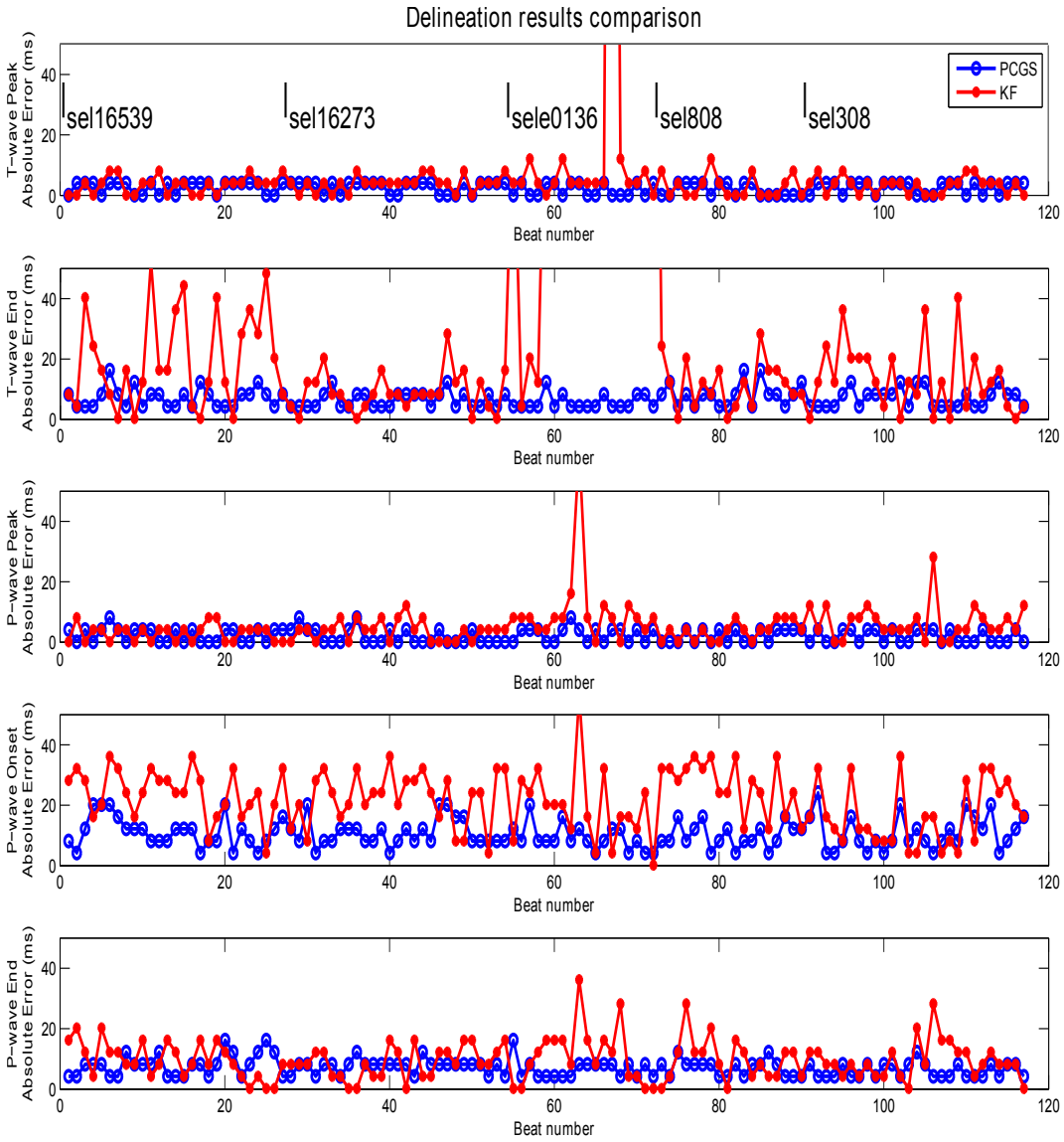


Figure 2.14: Absolute errors between the estimated values of T_{peak}, T_{end}, P_{on}, P_{peak} and P_{end} and manual annotations for representative signals from the QT datasets "sel16539", "sel16273", "sele0136", "sel808" and "sel308" (blue empty circles correspond to the PCGS method whereas red full circles correspond to the results provided by O. Sayadi associated to the method of [SS09].

significantly larger than the KF method of [SS09].

Quantitative comparisons with other classical methods

The analytical evaluation of the P and T wave detection can be performed by calculating the sensitivity (also referred to as detection rate) $Se = TP/(TP + FN)$ and positive predictivity $P^+ = TP/(TP + FP)$, 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 stands for the number of false positive (wave was not present but was detected). The performance of wave delineation is measured by the average of errors m, which stands for the time differences between cardiologist annotations and results of the proposed method. The average of the intra-recording standard deviations denoted as s was also computed.

The validation results obtained with the PCGS-based delineator and the three other methods of [LJC94, MAO+04, VGP+00] on the QTDB are given in Table 3.1. All ECG signals used in this paper have been preprocessed by Pan and Tompkins' QRS detection algorithm, with an overall QRS detection result of $Se = 99.7\%$ and $P^+ = 99.6\%$. The beats where QRS complexes were not well detected were excluded from the P and T wave evaluation.

The detection results for the QTDB show that the proposed method can detect with high sensitivity the P and T-waves annotated by cardiologists in the ECG signals. We obtained a sensitivity of $Se = 98.93\%$ for the P waves and a sensitivity of $Se = 99.81\%$ for the T waves. These results are slightly better than the ones obtained with the other methods. As for the positive predictivity, we have obtained very good results since $P^+ = 97.40\%$ for the P waves and $P^+ = 98.97\%$ for the T waves, which clearly outperforms the other algorithms evaluated in Table 3.1. This is partly because the minimum-distance constraint in Bayesian detection reduces the probability of false positives. The delineation performance is also presented in Table 3.1. The proposed algorithm can delineate the annotated P and T waves with mean errors m that do not exceed two samples (8 ms). The standard deviations s are around four samples for the P wave and five samples for the T wave, which is quite satisfactory.

Histograms of deviations between the results of the proposed algorithm compared to the "gold standard" of the manually measured annotations for TP interval ($\mathrm{TP_{int}} = \mathrm{T_{peak}} - \mathrm{P_{peak}}$), P wave duration ($\mathrm{P_{dur}} = \mathrm{P_{end}} - \mathrm{P_{onset}}$), ST interval ($\mathrm{ST_{int}} = \mathrm{S_{peak}} - \mathrm{T_{end}}$) and QT^p interval ($\mathrm{QT}^p_{\mathrm{int}} = \mathrm{Q_{peak}} - \mathrm{T_{peak}}$) are presented in Fig. 2.15. These histograms are in agreement with the results of [MKS06] which indicate that typical standard deviations for QT interval measurement are 20-30 ms. The deviations of the PCGS-based method are also similar to those obtained with the more recent technique studied in [SS09]. Smaller deviations (mostly below 8 ms) have been obtained for detections which rely on peak points, i.e., $\mathrm{TP_{int}}$ and $\mathrm{QT}^p_{\mathrm{int}}$. For those detections which rely on peak boundaries, the deviations are also in the acceptable range (mostly below 20 ms). Note that the proposed method focuses on P and T wave analysis. Thus deviations of QRS locations are not considered in the validation.

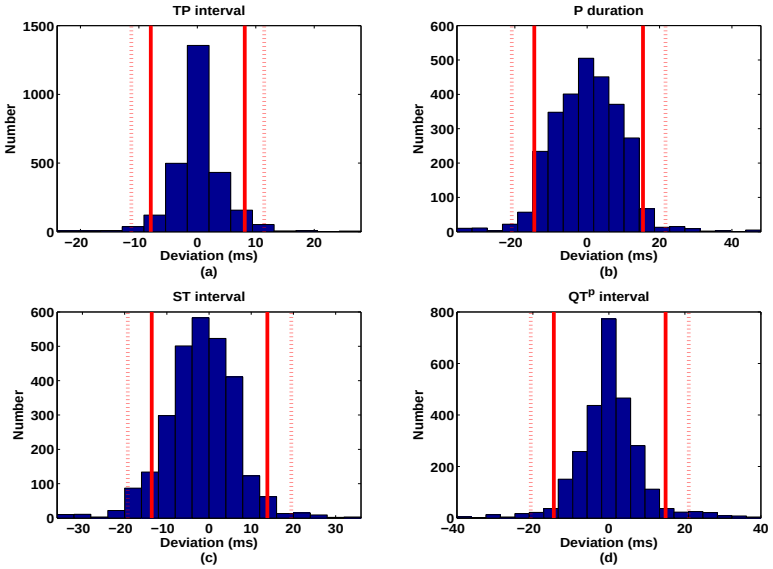


Figure 2.15: Histograms of deviations between the results of the proposed automatic algorithm compared to the "gold standard" of manually measured annotations including (a) TP interval, (b) P wave duration, (c) ST interval and (d) QT^p interval. The vertical lines indicate the 95% (continuous red) and 99% (dashed red) confidence intervals.

Table 2.1: Delineation and detection performance comparison of the window based partially collapsed Gibbs sampler method of [LMT10] (PCGS), the Wavelet transform based method of [MAO+04] (WT), the low-pass differentiation based method of [LJC94] (LPD), the action potential model of [VGP+00] and the delineation error tolerance of [TCWP85]. (N/A: not available)

Method	Parameters	P_{on}	P_{peak}	P_{end}	T_{on}	T_{peak}	T_{end}
multi-beat window based PCGS (proposed)	annotations	3176	3176	3176	1345	3403	3403
	$Se\,(\%)$	98.93	98.93	98.93	99.01	99.81	99.81
	$P^+\,(\%)$	97.40	97.40	97.40	96.07	98.97	98.97
	$m \pm s\,(ms)$	3.7±17.3	4.1±8.6	-3.1±15.1	7.1±18.5	1.3±10.5	4.3±20.8
WT [MAO+04]	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 [LJC94]	$Se\,(\%)$	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 model [VGP+00]	$Se\,(\%)$	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	N/A	-12.0±23.4	0.8±30.3
Tolerance	$2s_{\text{CSE}}$ (ms)	10.2	N/A	12.7	N/A	N/A	30.6

2.4 Modified Bayesian model and a block Gibbs sampler

In Section 2.3, the local dependency of the ECG was expressed by a minimum-distance constraint: a T wave (P wave) indicator vector cannot have two non-zero elements closer than a certain distance, where this distance depends on the RR interval length. This is because successive T waves (P waves) can only appear in search intervals located in the right-hand (left-hand) neighborhoods of each QRS offset (onset). In this section, we propose to exploit this local dependency in a different way.

It is common to view ECGs as the combination of two elements, namely, QRS complexes and non-QRS intervals. Non-QRS intervals are located between a QRS end and the subsequent QRS onset. The locations of non-QRS intervals are provided by a preliminary QRS detection step, e.g., resulting from Pan and Tompkins algorithm [PT85]. The non-QRS interval $\mathcal{J}_n$ associated with the nth heartbeat consists of a T wave search interval $\mathcal{J}_{\mathrm{T},n}$, which may contain a T wave, and a P wave interval $\mathcal{J}_{\mathrm{P},n}$, which may contain a P wave. Different from the preprocessing procedure defined in Section 2.2.1, the search intervals of this section are only defined to identify wave indicator constraints, i.e., only the locations of the wave peaks are constrained to lie within their respective search intervals. Rather than extracting left- and right-hand neighborhoods of QRS complexes to form two signal portions (one for T waves and the other for P waves) and processing them individually using the same Bayesian inference, the whole non-QRS signal component will be considered together in the processing so that the border situation (where a T or P wave has part of its waveform across the border between $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$) can be properly handled.

The temporal lengths of the intervals $\mathcal{J}_n$, $\mathcal{J}_{\mathrm{T},n}$, and $\mathcal{J}_{\mathrm{P},n}$ are denoted as N_n, $N_{\mathrm{T},n}$, and $N_{\mathrm{P},n}$, respectively. Here, $N_{\mathrm{T},n}$ and $N_{\mathrm{P},n}$ can be determined by a cardiologist or simply as fixed percentages of N_n. Note that $N_{\mathrm{T},n} + N_{\mathrm{P},n} = N_n$. Our goal is to estimate the waveforms and amplitudes of the P and T waves and their locations within their respective intervals. Due to the pseudo-stationary nature of ECGs, we still adopt the D-beat processing window strategy to perform detection and estimation only for a limited set of consecutive beats at once. More specifically, we will consider the beats $n \in \{1, \ldots, D\}$ located within a D-beat processing window of length M (see Fig. 2.16(a)). As shown in Fig. 2.16(b), the signal in each non-QRS interval can be approximated by two pulses representing the P and T waves plus a local baseline. The T waveforms within a window are assumed to be equal, whereas the amplitudes and locations of the T pulses vary with the beat index n. The fact that the whole non-QRS component is considered together allows us to include the residual local baseline (see the blue dotted lines in Fig. 2.16(b)) in the signal model and to be estimated simultaneously with other wave parameters.

Thus, the T waves within a D-beat processing window can be modeled by the convolution of the unknown T waveform $\boldsymbol{h}_{\mathrm{T}} = [h_{\mathrm{T},-L} \cdots h_{\mathrm{T},L}]^T$ with an unknown "impulse" sequence $\boldsymbol{u}_{\mathrm{T}} = [u_{\mathrm{T},1} \cdots u_{\mathrm{T},M}]^T$ indicating the T wave locations and amplitudes (see Fig. 2.16(c)). Note that at most D entries of $\boldsymbol{u}_{\mathrm{T}}$ are nonzero; each of these "impulses" corresponds to one T wave. Similarly, the P waves within a D-beat processing window are modeled by the convolution of $\boldsymbol{h}_{\mathrm{P}} = [h_{\mathrm{P},-L} \cdots h_{\mathrm{P},L}]^T$ with $\boldsymbol{u}_{\mathrm{P}} = [u_{\mathrm{P},1} \cdots u_{\mathrm{P},M}]^T$. Let $\mathcal{J}$ denote the union of all T and P wave intervals $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$ within the window considered, and let $K \triangleq |\mathcal{J}|$ denote the corresponding signal length. The non-QRS signal component can then be written as

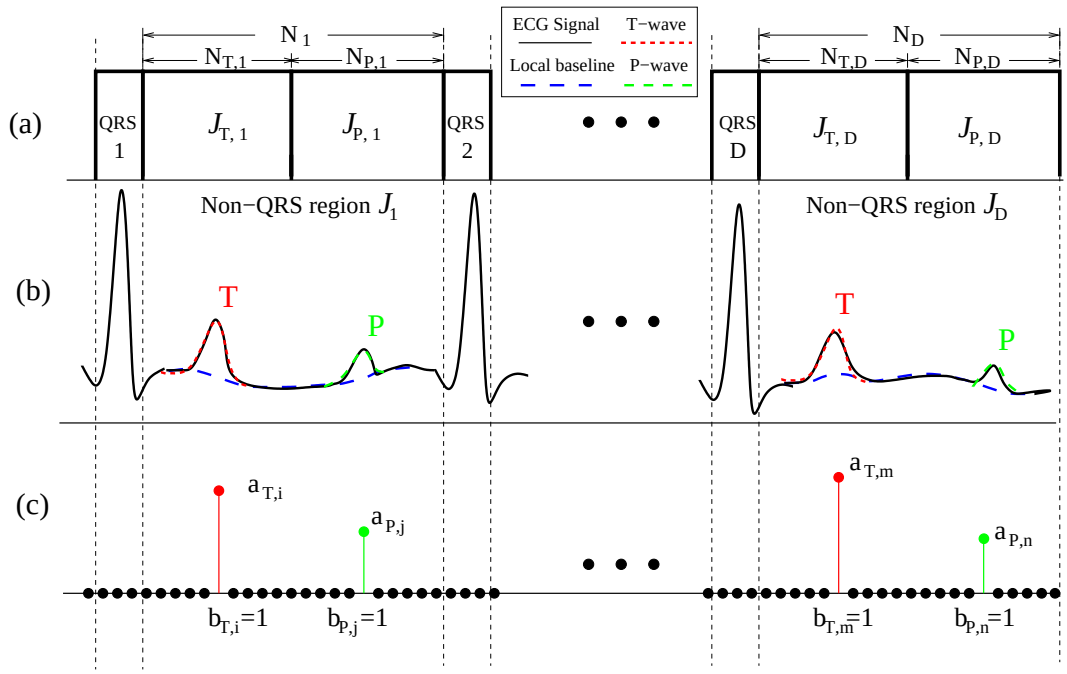


Figure 2.16: ECG signal within a D-beat window: (a) QRS and non-QRS intervals (in this example, $N_{\mathrm{T},n} = N_{\mathrm{P},n} = N_n/2$), (b) signal model for the non-QRS intervals, (c) impulse sequences $u_{\mathrm{T},k}$ and $u_{\mathrm{P},k}$.

$$x_k = \sum_{l=-L}^{L} h_{\mathrm{T},l} u_{\mathrm{T},k-l} + \sum_{l=-L}^{L} h_{\mathrm{P},l} u_{\mathrm{P},k-l} + c_k + w_k\,, \quad k \in \mathcal{J}, \tag{2.29}$$

where c_k denotes the baseline sequence and w_k denotes white Gaussian noise with unknown variance σ_w^2. Note that we assume that $u_{\mathrm{T},k} = u_{\mathrm{P},k} = 0$ for $k \notin \mathcal{J}$.

2.4.1 Hermite basis decomposition and local baseline

Instead of using a pointwise waveform as in Section 2.2.2, we propose in this modified model to represent the P and T waveforms by a basis expansion using discrete-time versions of the first G Hermite functions [HB97]. The Hermite basis was first used for ECG processing by Sörnmo *et al.* [SBNP81] for evaluating QRS shape features. It has also shown interesting properties for ECG data compression [JOLC93]. Thus, the waveform vectors can be written as

$$\boldsymbol{h}_{\mathrm{T}} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\,, \quad \boldsymbol{h}_{\mathrm{P}} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}\,, \tag{2.30}$$

where $\boldsymbol{H}$ is an $(2L+1) \times G$ matrix whose columns are the first G Hermite functions (with $G \leq 2L+1$), suitably sampled and truncated to length $2L+1$, and $\boldsymbol{\alpha}_{\mathrm{T}}$ and $\boldsymbol{\alpha}_{\mathrm{P}}$ are unknown coefficient vectors of length G. By using these expansions, the number of unknown parameters can be reduced from $2L+1$ to G for each waveform (e.g., we have chosen $G = 20$ and $2L+1$ is

around 83 for ECGs with heart rate around 60 beats per minute and sampled at 250Hz) and the out-of-band noise can be properly rejected.

Baseline removal is generally recognized as an important step that is beneficial to ECG delineation, since most traditional techniques (such as the widely used third-order spline interpolation) only remove baseline fluctuations at frequencies lower than heart rate [BZB06]. Here, we propose to model the residual local baseline within the nth non-QRS interval $\mathcal{J}_n$ by a 4th-degree polynomial, i.e.,

$$c_{n,k} = \sum_{i=1}^{5} \gamma_{n,i} k^{i-1}, \quad k = 1, \ldots, N_n \,, \tag{2.31}$$

for each $n \in \{1, \ldots, D\}$. This local baseline model extends that of [CM07], which assumes that the local baseline is constant in $\mathcal{J}_n$ (i.e., $\gamma_{n,i} = 0$ for $i \geq 2$). In vector-matrix form, (2.31) reads as $\boldsymbol{c}_n = \boldsymbol{M}_n \boldsymbol{\gamma}_n$, with the known $N_n \times 5$ Vandermonde matrix $\boldsymbol{M}_n$ and the unknown coefficient vector $\boldsymbol{\gamma}_n = (\gamma_{n,1} \cdots \gamma_{n,5})^T$. The baseline sequence for the entire D-beat window can then be written as

$$\boldsymbol{c} = \boldsymbol{M}\boldsymbol{\gamma} \,, \tag{2.32}$$

where $\boldsymbol{c}$, $\boldsymbol{M}$, and $\boldsymbol{\gamma}$ are obtained by suitably stacking the $\boldsymbol{c}_n$, $\boldsymbol{M}_n$, and $\boldsymbol{\gamma}_n$ as follows

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{0} & \cdots & \cdots & \cdots \\ \boldsymbol{M}_1 & \boldsymbol{0} & \cdots & \cdots \\ \boldsymbol{0} & \cdots & \cdots & \cdots \\ \vdots & \boldsymbol{M}_2 & \vdots & \vdots \\ \vdots & \boldsymbol{0} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \boldsymbol{M}_D \end{bmatrix}_{K \times (5*D)} . \tag{2.33}$$

Using (2.30), (2.32) and (2.33), we obtain the following vector representation of the non-QRS signal components of a D-beat processing window in (2.29)

$$\boldsymbol{x} = \boldsymbol{F}_{\mathrm{T}} \boldsymbol{u}_{\mathrm{T}} + \boldsymbol{F}_{\mathrm{P}} \boldsymbol{u}_{\mathrm{P}} + \boldsymbol{M}\boldsymbol{\gamma} + \boldsymbol{w} \,. \tag{2.34}$$

Here, $\boldsymbol{F}_{\mathrm{T}}$ is the $K \times M$ Toeplitz matrix with first row $\begin{bmatrix} (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}})_0 & \cdots & (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}})_{-L} & \boldsymbol{0} \end{bmatrix}$, and with first column $\begin{bmatrix} (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}})_0^T & \cdots & (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}})_L^T & \boldsymbol{0}^T \end{bmatrix}^T$. $\boldsymbol{F}_{\mathrm{P}}$ is defined similarly, with $\boldsymbol{\alpha}_{\mathrm{T}}$ replaced by $\boldsymbol{\alpha}_{\mathrm{P}}$.

Similar to the signal model in Section 2.2.2, the proposed detection-estimation method is based on writing the impulse sequences as products $u_{\mathrm{T},k} = b_{\mathrm{T},k} a_{\mathrm{T},k}$ ($u_{\mathrm{P},k} = b_{\mathrm{P},k} a_{\mathrm{P},k}$) of binary indicator sequences $b_{\mathrm{T},k} \in \{0, 1\}$ ($b_{\mathrm{P},k} \in \{0, 1\}$) and amplitude factors $a_{\mathrm{T},k} \in \mathbb{R}$ ($a_{\mathrm{P},k} \in \mathbb{R}$). Each $b_{\mathrm{T},k} = 1$ ($b_{\mathrm{P},k} = 1$) indicates the location of a T wave (P wave), and the corresponding $a_{\mathrm{T},k}$ ($a_{\mathrm{P},k}$) is the respective amplitude. Note that the $a_{\mathrm{T},k}$ ($a_{\mathrm{P},k}$) are undefined for all k where $b_{\mathrm{T},k} = 0$ ($b_{\mathrm{P},k} = 0$). Let $\boldsymbol{b}_{\mathrm{T}}$, $\boldsymbol{b}_{\mathrm{P}}$, $\boldsymbol{a}_{\mathrm{T}}$, and $\boldsymbol{a}_{\mathrm{P}}$ denote the length-M vectors corresponding to $b_{\mathrm{T},k}$,

$b_{\mathrm{P},k}$, $a_{\mathrm{T},k}$, and $a_{\mathrm{P},k}$, respectively. Then (2.34) can be rewritten as

$$\boldsymbol{x} = \boldsymbol{F}_{\mathrm{T}}\boldsymbol{B}_{\mathrm{T}}\boldsymbol{a}_{\mathrm{T}} + \boldsymbol{F}_{\mathrm{P}}\boldsymbol{B}_{\mathrm{P}}\boldsymbol{a}_{\mathrm{P}} + \boldsymbol{M}\boldsymbol{\gamma} + \boldsymbol{w}\,, \tag{2.35}$$

with the diagonal $M \times M$ matrices $\boldsymbol{B}_{\mathrm{T}} \triangleq \mathrm{diag}(\boldsymbol{b}_{\mathrm{T}})$ and $\boldsymbol{B}_{\mathrm{P}} \triangleq \mathrm{diag}(\boldsymbol{b}_{\mathrm{P}})$.

2.4.2 Bayesian inference

The unknown parameter vector resulting from the above parametrization is $\boldsymbol{\theta} = \left[\boldsymbol{\theta}_{\mathrm{T}}^T\, \boldsymbol{\theta}_{\mathrm{P}}^T\, \boldsymbol{\theta}_{\mathrm{cw}}^T\right]^T$, where $\boldsymbol{\theta}_{\mathrm{T}} \triangleq \left[\boldsymbol{b}_{\mathrm{T}}^T\, \boldsymbol{a}_{\mathrm{T}}^T\, \boldsymbol{\alpha}_{\mathrm{T}}^T\right]^T$ and $\boldsymbol{\theta}_{\mathrm{P}} \triangleq \left[\boldsymbol{b}_{\mathrm{P}}^T\, \boldsymbol{a}_{\mathrm{P}}^T\, \boldsymbol{\alpha}_{\mathrm{P}}^T\right]^T$ are related to the T and P waves, respectively and $\boldsymbol{\theta}_{\mathrm{cw}} \triangleq \left[\boldsymbol{\gamma}^T\, \sigma_w^2\right]^T$ is related to the local baseline and noise. As expressed by (2.4), Bayesian detection/estimation relies on the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$ which is related to the likelihood of the observations $p(\boldsymbol{x}|\boldsymbol{\theta})$ and the prior of the parameters $p(\boldsymbol{\theta})$. The likelihood $p(\boldsymbol{x}|\boldsymbol{\theta})$ and the prior $p(\boldsymbol{\theta})$ for the modified signal model in (2.35) are summarized below.

Likelihood

Using our model (2.35) and the fact that w_k is white Gaussian noise, the likelihood function is obtained as

$$p(\boldsymbol{x}|\boldsymbol{\theta}) \propto \frac{1}{\sigma_w^K} \exp\!\left(-\frac{1}{2\sigma_w^2}\|\boldsymbol{x} - \boldsymbol{F}_{\mathrm{T}}\boldsymbol{B}_{\mathrm{T}}\boldsymbol{a}_{\mathrm{T}} - \boldsymbol{F}_{\mathrm{P}}\boldsymbol{B}_{\mathrm{P}}\boldsymbol{a}_{\mathrm{P}} - \boldsymbol{M}\boldsymbol{\gamma}\|^2\right), \tag{2.36}$$

where $\|\cdot\|$ is the ℓ_2 norm, i.e., $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$.

Prior distributions

Since there are no known relations between the impulse parameters $(\boldsymbol{b}_{\mathrm{T}}, \boldsymbol{a}_{\mathrm{T}})$, $(\boldsymbol{b}_{\mathrm{P}}, \boldsymbol{a}_{\mathrm{P}})$, the waveform parameters $\boldsymbol{\alpha}_{\mathrm{T}}$, $\boldsymbol{\alpha}_{\mathrm{P}}$, the baseline coefficients $\boldsymbol{\gamma}$, and the noise variance σ_w^2, all these sets of parameters are modeled as *a priori* statistically independent of each other such that

$$\begin{aligned} p\left(\boldsymbol{\theta}\right) &= p\left(\boldsymbol{\theta}_{\mathrm{T}}\right) p\left(\boldsymbol{\theta}_{\mathrm{P}}\right) p\left(\boldsymbol{\theta}_{\mathrm{cw}}\right) \\ &= p\left(\boldsymbol{b}_{\mathrm{T}}\right) p\left(\boldsymbol{a}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}}\right) p\left(\boldsymbol{\alpha}_{\mathrm{T}}\right) p\left(\boldsymbol{b}_{\mathrm{P}}\right) p\left(\boldsymbol{a}_{\mathrm{P}}|\boldsymbol{b}_{\mathrm{P}}\right) p\left(\boldsymbol{\alpha}_{\mathrm{P}}\right) p\left(\boldsymbol{\gamma}\right) p\left(\sigma_n^2\right). \end{aligned} \tag{2.37}$$

We will now discuss the prior distributions of each of these parameters. Let $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}$, $n \in \{1, \ldots, D\}$ comprise all entries of the T wave indicator vector $\boldsymbol{b}_{\mathrm{T}}$ that are indexed by the T wave interval $\mathcal{J}_{\mathrm{T},n}$. The indicators are subject to a *block constraint*: within $\mathcal{J}_{\mathrm{T},n}$, there is one T wave indicator (thus, $\|\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}\| = 1$) or none (thus, $\|\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}\| = 0$), the latter case being very

unlikely. Therefore, we define the prior of $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}$ as

$$p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}) = \begin{cases} p_0 & \text{if } \|\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}\| = 0 \\ p_1 & \text{if } \|\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}\| = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{2.38}$$

where $p_1 = (1-p_0)/N_{\mathrm{T},n}$ and p_0 is chosen very small. Note that there are $N_{\mathrm{T},n}$ vectors satisfying $\|\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}\| = 1$ whereas the zero vector is the only vector satisfying $\|\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}\| = 0$. Thus the probabilities in (2.38) sum to one. The $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}$ are statistically independent of each other, and all remaining entries of the total vector $\boldsymbol{b}_{\mathrm{T}}$ (i.e., entries outside the T wave intervals $\mathcal{J}_{\mathrm{T},n}$) are zero. Thus, the prior of $\boldsymbol{b}_{\mathrm{T}}$ is given by

$$p(\boldsymbol{b}_{\mathrm{T}}) = \prod_{n=1}^{D} p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}})\,. \tag{2.39}$$

For the T wave amplitudes a_k at those k where $b_k = 1$ (recall that $b_k = 0$ implies $a_k = 0$), the same zero-mean Gaussian prior has been chosen as in Section 2.2.3

$$p(a_k|b_k) = \begin{cases} \delta(a_k) & \text{if } b_k = 0 \\ \mathcal{N}(0, \sigma_a^2) & \text{if } b_k = 1 \end{cases} \tag{2.40}$$

Similar to the indicators, amplitudes at different k are also modeled as statistically independent such that

$$p\left(\boldsymbol{a}|\boldsymbol{b}\right) = \prod_{k=1}^{K} p\left(a_k|b_k\right). \tag{2.41}$$

It follows that $u_{\mathrm{T},k} = b_{\mathrm{T},k}a_{\mathrm{T},k}$ is a Bernoulli-Gaussian (BG) sequence with block constraints. Different from the BG sequence with minimum distance constraint prior in Section 2.2.3, the block constraint considers all the indicators b_k within a search interval $\mathcal{J}_{\mathrm{T},n}$ as an entirety. Doing this allows the Gibbs sampler to generate block by block the indicator vector at each substep instead of position by position (see Section 2.4.3 for more details). Note that the priors of the P wave indicators $b_{\mathrm{P},k}$ and amplitudes $a_{\mathrm{P},k}$ are defined in a fully analogous way, with the same fixed hyperparameters p_0, p_1, and σ_a^2.

The T waveform coefficients are modeled as independent and identically distributed (iid), zero-mean, and Gaussian. The same prior is chosen for the P wave coefficients

$$p(\boldsymbol{\alpha}_{\mathrm{T}}) = \mathcal{N}(\mathbf{0}, \sigma_\alpha^2 \mathbf{I}_{2L+1}) \tag{2.42}$$

$$p(\boldsymbol{\alpha}_{\mathrm{P}}) = \mathcal{N}(\mathbf{0}, \sigma_\alpha^2 \mathbf{I}_{2L+1}). \tag{2.43}$$

where $\mathbf{I}_{2L+1}$ denotes the identity matrix of size $(2L+1) \times (2L+1)$.

The baseline coefficients $\gamma_{n,i}$ are also modeled as iid zero-mean Gaussian

$$p(\boldsymbol{\gamma}) = \mathcal{N}(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}_{5D}). \tag{2.44}$$

The variances σ_a^2, σ_α^2, and σ_γ^2 are fixed hyperparameters as in Section 2.2.3. Finally, the noise variance σ_w^2 is modeled as a random hyperparameter with inverse gamma distribution (cf. (2.17))

$$p(\sigma_w^2) = \mathcal{IG}(\xi, \eta) \tag{2.45}$$

where ξ and η are fixed hyperparameters providing a vague prior. We recall that choosing conjugate priors for $\boldsymbol{\alpha}_\mathrm{T}$, $\boldsymbol{\alpha}_\mathrm{P}$, $\boldsymbol{\gamma}$, and σ_w^2 considerably simplifies our algorithm since the resulting posterior distributions are also Gaussian and inverse gamma, respectively.

Posterior distribution

The posterior distribution of the parameter vector $\boldsymbol{\theta}$ is given by

$$p(\boldsymbol{\theta}|\boldsymbol{x}) \propto p(\boldsymbol{x}|\boldsymbol{\theta}) p(\boldsymbol{a}_\mathrm{T}|\boldsymbol{b}_\mathrm{T}) p(\boldsymbol{b}_\mathrm{T}) p(\boldsymbol{\alpha}_\mathrm{T}) p(\boldsymbol{a}_\mathrm{P}|\boldsymbol{b}_\mathrm{P}) p(\boldsymbol{b}_\mathrm{P}) p(\boldsymbol{\alpha}_\mathrm{P}) p(\boldsymbol{\gamma}) p(\sigma_w^2)\,. \tag{2.46}$$

Taking advantage of the block constraint associated to the wave indicator prior, we propose to use a block Gibbs sampler that generates samples asymptotically distributed according to $p(\boldsymbol{\theta}|\boldsymbol{x})$. From these samples, the discrete parameters $\boldsymbol{b}_\mathrm{T}$ and $\boldsymbol{b}_\mathrm{P}$ are then detected by means of the sample-based maximum a posteriori detector, and the continuous parameters $\boldsymbol{a}_\mathrm{T}$, $\boldsymbol{a}_\mathrm{P}$, $\boldsymbol{\alpha}_\mathrm{T}$, $\boldsymbol{\alpha}_\mathrm{P}$, $\boldsymbol{\gamma}$, and σ_w^2 are estimated by means of the sample-based minimum mean square error estimator.

2.4.3 Block Gibbs sampler for P and T wave analysis

Block Gibbs sampler

A block Gibbs sampler is another variant of the Gibbs sampler. It groups two or more variables together and samples from their joint distribution conditioned on all other variables, rather than sampling from each one individually. Take the prototype Gibbs sampler in Algorithm 1 as an example, instead of iteratively sampling each θ_l from $p(\theta_l|\boldsymbol{\theta}_{\sim l}, \boldsymbol{x})$ ($l \in \{1, \ldots, L\}$) in an arbitrary order, a block Gibbs sampler may firstly sample a subset of the parameters $\boldsymbol{\theta}_\mathcal{J} \subseteq \{\theta_1, \ldots, \theta_L\}$ from their joint distribution conditioned on the rest $\boldsymbol{\theta}_{\sim\mathcal{J}}$, i.e., from $p(\boldsymbol{\theta}_\mathcal{J}|\boldsymbol{\theta}_{\sim\mathcal{J}}, \boldsymbol{x})$. Then, the parameters indexed by $\sim\mathcal{J}$ can be iteratively drawn conditioned on $\boldsymbol{\theta}_\mathcal{J}$, i.e., from $p(\theta_l|\boldsymbol{\theta}_{\sim\mathcal{J}\setminus l}, \boldsymbol{\theta}_\mathcal{J}, \boldsymbol{x})$ with $l \in \{\sim\mathcal{J}\}$. As known, the blocked Gibbs sampler works more efficiently than the ordinary Gibbs sampler if the variables to be inferred are highly dependent.

Proposed block Gibbs sampler for P and T wave analysis

The proposed block Gibbs sampler for our problem is summarized in Algorithm 5. The sampling distributions involved are the conditional distributions of $p(\boldsymbol{\theta}|\boldsymbol{x})$ defined in (2.46) and are detailed below. The sampling distributions of the P wave parameters $\boldsymbol{\theta}_{\mathrm{P}}$ are similar to those of $\boldsymbol{\theta}_{\mathrm{T}}$. We note that $\boldsymbol{b}_{\sim\mathcal{J}_{\mathrm{T},n}}$ denotes $\boldsymbol{b}_{\mathrm{T}}$ without the entries indexed by $\mathcal{J}_{\mathrm{T},n}$, and similarly for $\boldsymbol{a}_{\sim\mathcal{J}_{\mathrm{T},n}}$. To see that the proposed algorithm is a valid Gibbs sampler, note that the sampling steps for $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}$ and $\boldsymbol{a}_{\mathcal{J}_{\mathrm{T},n}}$ are equivalent to jointly sampling $(\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}, \boldsymbol{a}_{\mathcal{J}_{\mathrm{T},n}})$ from $p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}, \boldsymbol{a}_{\mathcal{J}_{\mathrm{T},n}} | \boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{T},n}}, \boldsymbol{\theta}_{\mathrm{P}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x})$, where $\boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{T},n}} \triangleq (\boldsymbol{b}^T_{\sim\mathcal{J}_{\mathrm{T},n}}\, \boldsymbol{a}^T_{\sim\mathcal{J}_{\mathrm{T},n}}\, \boldsymbol{\alpha}^T_{\mathrm{T}})^T$.

Algorithm 5 Block Gibbs sampler for P and T wave analysis.

for $n = 1, \ldots, D$ **do**
  Sample the block $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}} \sim p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}} | \boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{T},n}}, \boldsymbol{\theta}_{\mathrm{P}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x})$ (see (2.47))
  for $k \in \mathcal{J}_{\mathrm{T},n}$ **do**
    if $b_{\mathrm{T},k} = 1$ **then**
      Sample $a_{\mathrm{T},k} \sim p(a_{\mathrm{T},k} | b_{\mathrm{T},k} = 1, \boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{T},n}}, \boldsymbol{\theta}_{\mathrm{P}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x})$ (see (2.48))
    end if
  end for
  Sample the block $\boldsymbol{b}_{\mathcal{J}_{\mathrm{P},n}}$ from $p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{P},n}} | \boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{P},n}}, \boldsymbol{\theta}_{\mathrm{T}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x})$
  for $k \in \mathcal{J}_{\mathrm{P},n}$ **do**
    if $b_{\mathrm{P},k} = 1$ **then**
      Sample $a_{\mathrm{P},k} \sim p(a_{\mathrm{P},k} | b_{\mathrm{P},k} = 1, \boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{P},n}}, \boldsymbol{\theta}_{\mathrm{T}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x})$
    end if
  end for
end for
Sample $\boldsymbol{\alpha}_{\mathrm{T}}$ from $p(\boldsymbol{\alpha}_{\mathrm{T}} | \boldsymbol{b}_{\mathrm{T}}, \boldsymbol{a}_{\mathrm{T}}, \boldsymbol{\theta}_{\mathrm{P}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x})$ (see (2.49))
Sample $\boldsymbol{\alpha}_{\mathrm{P}}$ from $p(\boldsymbol{\alpha}_{\mathrm{P}} | \boldsymbol{b}_{\mathrm{P}}, \boldsymbol{a}_{\mathrm{P}}, \boldsymbol{\theta}_{\mathrm{T}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x})$
Sample $\boldsymbol{\gamma}$ from $p(\boldsymbol{\gamma} | \boldsymbol{\theta}_{\mathrm{T}}, \boldsymbol{\theta}_{\mathrm{P}}, \sigma_w^2, \boldsymbol{x})$ (see (2.50))
Sample σ_w^2 from $p(\sigma_w^2 | \boldsymbol{\theta}_{\mathrm{T}}, \boldsymbol{\theta}_{\mathrm{P}}, \boldsymbol{\gamma}, \boldsymbol{x})$ (see (2.51))

T wave indicators. The sampling conditional distribution for the T wave indicator vector $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}$ is

$$p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}} | \boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{T},n}}, \boldsymbol{\theta}_{\mathrm{P}}, \boldsymbol{\theta}_{\mathrm{cw}}, \boldsymbol{x}) \propto \sigma_1 \exp\left(\frac{\mu_1^2}{2\sigma_1^2}\right) p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}), \tag{2.47}$$

with

$$\mu_1 = \frac{\sigma_1^2 \boldsymbol{b}^T_{\mathcal{J}_{\mathrm{T},n}} \boldsymbol{F}^T_{\mathcal{J}_{\mathrm{T},n}}}{\sigma_w^2} (\boldsymbol{x}_{\mathrm{T}} - \boldsymbol{F}_{\sim\mathcal{J}_{\mathrm{T},n}} \boldsymbol{B}_{\sim\mathcal{J}_{\mathrm{T},n}} \boldsymbol{a}_{\sim\mathcal{J}_{\mathrm{T},n}}), \quad \sigma_1^2 = \left(\frac{\left\| \boldsymbol{F}_{\mathcal{J}_{\mathrm{T},n}} \boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}} \right\|^2}{\sigma_w^2} + \frac{1}{\sigma_a^2} \right)^{-1},$$

where $\boldsymbol{F}_{\mathcal{J}_{\mathrm{T},n}}$ contains the columns of $\boldsymbol{F}_{\mathrm{T}}$ indexed by $\mathcal{J}_{\mathrm{T},n}$, $\boldsymbol{F}_{\sim\mathcal{J}_{\mathrm{T},n}}$ is $\boldsymbol{F}_{\mathrm{T}}$ without those columns, $\boldsymbol{x}_{\mathrm{T}} \triangleq \boldsymbol{x} - \boldsymbol{F}_{\mathrm{P}} \boldsymbol{B}_{\mathrm{P}} \boldsymbol{a}_{\mathrm{P}} - \boldsymbol{M}\boldsymbol{\gamma}$, and $\boldsymbol{B}_{\sim\mathcal{J}_{\mathrm{T},n}} \triangleq \mathrm{diag}(\boldsymbol{b}_{\sim\mathcal{J}_{\mathrm{T},n}})$. In each iteration, the sampler evaluates all hypotheses of $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}$ conditioned on the samples of all other parameters. There are $N_{\mathrm{T},n}+1$ such hypotheses, because $\boldsymbol{b}_{\mathcal{J}_{\mathrm{T},n}}$ has either no 1-entry or exactly one 1-entry at $N_{\mathrm{T},n}$ possible locations (cf. the prior in (2.38)).

T wave amplitudes. The sampling distribution for the $a_{\mathrm{T},k}$ is

$$p(a_{\mathrm{T},k}|b_{\mathrm{T},k}=1,\boldsymbol{\theta}_{\sim\mathcal{J}_{\mathrm{T},n}},\boldsymbol{\theta}_{\mathrm{P}},\boldsymbol{\theta}_{\mathrm{cw}},\boldsymbol{x})=\mathcal{N}(\mu_1,\sigma_1^2)\,. \tag{2.48}$$

T waveform coefficients. The sampling distribution for $\boldsymbol{\alpha}_{\mathrm{T}}$ is

$$p(\boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{\theta}_{\mathrm{P}},\boldsymbol{\theta}_{\mathrm{cw}},\boldsymbol{x})=\mathcal{N}(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)\,, \tag{2.49}$$

with

$$\boldsymbol{\mu}_2=\frac{\boldsymbol{\Sigma}_2\boldsymbol{H}^T\boldsymbol{U}_{\mathrm{T}}^T}{\sigma_w^2}\,\boldsymbol{x}_{\mathrm{T}}\,,\quad \boldsymbol{\Sigma}_2=\left(\frac{\boldsymbol{H}^T\boldsymbol{U}_{\mathrm{T}}^T\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}}{\sigma_w^2}+\frac{\mathbf{I}_{L+1}}{\sigma_\alpha^2}\right)^{-1}.$$

Here, $\boldsymbol{U}_{\mathrm{T}}$ is the Toeplitz matrix of size $K\times(2L+1)$ with first row $[\boldsymbol{u}_{\mathrm{T},L+1}\,\cdots\,\boldsymbol{u}_{\mathrm{T},1}\,\mathbf{0}]$ and first column $\left[\boldsymbol{u}_{\mathrm{T},L+1}^T\,\cdots\,\boldsymbol{u}_{\mathrm{T},M}^T\,\mathbf{0}^T\right]^T$.

Local baseline coefficients. The sampling distribution for $\boldsymbol{\gamma}$ is

$$p(\boldsymbol{\gamma}|\boldsymbol{\theta}_{\mathrm{T}},\boldsymbol{\theta}_{\mathrm{P}},\sigma_w^2,\boldsymbol{x})=\mathcal{N}(\boldsymbol{\mu}_3,\boldsymbol{\Sigma}_3)\,, \tag{2.50}$$

with

$$\boldsymbol{\mu}_3=\frac{\boldsymbol{\Sigma}_3\boldsymbol{M}^T}{\sigma_w^2}\,\boldsymbol{x}_{\boldsymbol{\gamma}}\,,\quad \boldsymbol{\Sigma}_3=\left(\frac{\boldsymbol{M}^T\boldsymbol{M}}{\sigma_w^2}+\frac{\mathbf{I}_{5D}}{\sigma_\gamma^2}\right)^{-1},$$

where $\boldsymbol{x}_{\boldsymbol{\gamma}}\triangleq\boldsymbol{x}-\boldsymbol{F}_{\mathrm{T}}\boldsymbol{B}_{\mathrm{T}}\boldsymbol{a}_{\mathrm{T}}-\boldsymbol{F}_{\mathrm{P}}\boldsymbol{B}_{\mathrm{P}}\boldsymbol{a}_{\mathrm{P}}$.

Noise variance. The sampling distribution for σ_w^2 is

$$p(\sigma_w^2|\boldsymbol{\theta}_{\mathrm{T}},\boldsymbol{\theta}_{\mathrm{P}},\boldsymbol{\gamma},\boldsymbol{x})=\mathcal{IG}(\xi',\eta')\,, \tag{2.51}$$

with $\xi'=\xi+\frac{K}{2}$ and $\eta'=\eta+\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{F}_{\mathrm{T}}\boldsymbol{B}_{\mathrm{T}}\boldsymbol{a}_{\mathrm{T}}-\boldsymbol{F}_{\mathrm{P}}\boldsymbol{B}_{\mathrm{P}}\boldsymbol{a}_{\mathrm{P}}-\boldsymbol{M}\boldsymbol{\gamma}\|^2$.

2.4.4 Parameter estimation with the block Gibbs sampler

Similarly to the PCGS approach, P and T wave detection and estimation are based on the estimated joint posterior distribution of wave indicators, wave amplitudes and waveform coefficients. This posterior is computed from the last samples drawn from $p(\boldsymbol{\theta}|\boldsymbol{x})$ by the block Gibbs sampler. The posterior distribution of wave indicators carries information regarding the probability of having a P or T wave at a given location. Thus the wave detection results (wave locations) can be obtained by using a sample-based MAP estimator. Since we expect at most one T wave (P wave) indicator in the T search interval $\mathcal{J}_{\mathrm{T},n}$ (P search interval $\mathcal{J}_{\mathrm{P},n}$ with $n\in 1,\dots D$), the sample-based blockwise MAP detector of the wave indicator vector $\hat{\boldsymbol{b}}$ is given

by

$$\hat{b}_k = \begin{cases} \arg \max\limits_{b_k \in \{0,1\}} p_{\mathcal{S}}\left(b_k\right) , \ k \in \mathcal{J}_{\mathrm{T},n} \\ \arg \max\limits_{b_k \in \{0,1\}} p_{\mathcal{S}}\left(b_k\right) , \ k \in \mathcal{J}_{\mathrm{P},n}. \end{cases} . \tag{2.52}$$

Here, $p_{\mathcal{S}}\left(b_k\right)$ is a sample-based approximation of the posterior probability $p(b_k|\boldsymbol{x})$ of $b_k \in \{0,1\}$ defined as the number of realizations that contain the respective value of b_k normalized by the total number of generated samples N_{s}. Note that compared to the local MAP estimator for the wave indicator parameter in Section 2.3.3, the advantage here is that no probability threshold is needed to determine whether the wave is representative or not.

Concerning the other parameters $\boldsymbol{\theta}_{\mathrm{d}} = \left[\boldsymbol{a}_{\mathrm{T}}^T \boldsymbol{\alpha}_{\mathrm{T}}^T \boldsymbol{a}_{\mathrm{P}}^T \boldsymbol{\alpha}_{\mathrm{P}}^T \boldsymbol{\gamma}^T \sigma_w^2\right]^T$, the sample-based MMSE estimation can be given as

$$\hat{\boldsymbol{\theta}}_{\mathrm{d}} = \frac{1}{N_{\mathrm{d}}} \sum_{i=N_{\mathrm{bi}}+1}^{N_{\mathrm{bi}}+N_{\mathrm{d}}} \boldsymbol{\theta}_{\mathrm{d}}^{(i)} \tag{2.53}$$

We recall that N_{bi} is the number of burn-in iterations, N_{d} denotes the number of generated samples corresponding to the estimated wave indicator vector $\hat{\boldsymbol{b}}$ in (2.52), $\boldsymbol{\theta}_{\mathrm{d}}^{(i)}$ denotes the parameter vector generated at iteration i, and the subscript "d" in $\boldsymbol{\theta}_{\mathrm{d}}$ indicates that the parameters correspond to the detection indicated by $\hat{\boldsymbol{b}}$.

2.4.5 Simulation results

Simulations have been conducted on the QTDB to validate the performance of the proposed algorithm. Some representative results produced by the modified window based Bayesian model and the block Gibbs sampler are illustrated as follows. Fig. 2.17(a) shows an ECG signal segment from the QTDB dataset "sele0607", which contains overlapping P and T waves. Corresponding estimates of the marginal posterior probabilities of having a P or T wave at a given location are shown in Fig. 2.17(b). We note that the sample-based MAP detector used in our algorithm detects a wave at a given location if the respective posterior probability estimate is above 1/2. Fig. 2.17(c) shows the estimated T and P waveforms for the 10-beat window considered (i.e., $D = 10$). The wave onsets and ends were determined by using the delineation criteria presented in Section 2.3.4. Fig. 2.18(a) shows the estimates of the P and T waves (in red) and of the local baseline (in black), while Fig. 2.18(b) shows the P and T wave delineation results (i.e., estimated onset, peak, and end of each detected P and T wave). The close agreement of these estimation and delineation results with the ECG signal shows that the proposed algorithm can properly handle the stress test ECGs.

An example of signals that contains premature ventricular contractions (PVCs) is also studied by processing the QTDB dataset "sel803". As shown in Fig. 2.19, the proposed method can handle these non-monotonic morphological abnormalities. Note in particular that the posterior probabilities of having a P wave around time instant 7.5 is very low, which is in agreement with the presence of a unique wave in the non QRS interval. Furthermore, with the help of the proposed signal model, Fig. 2.20 shows that the sudden T wave amplitude inversion has been detected, which is a nice property for the PVC detection problem.

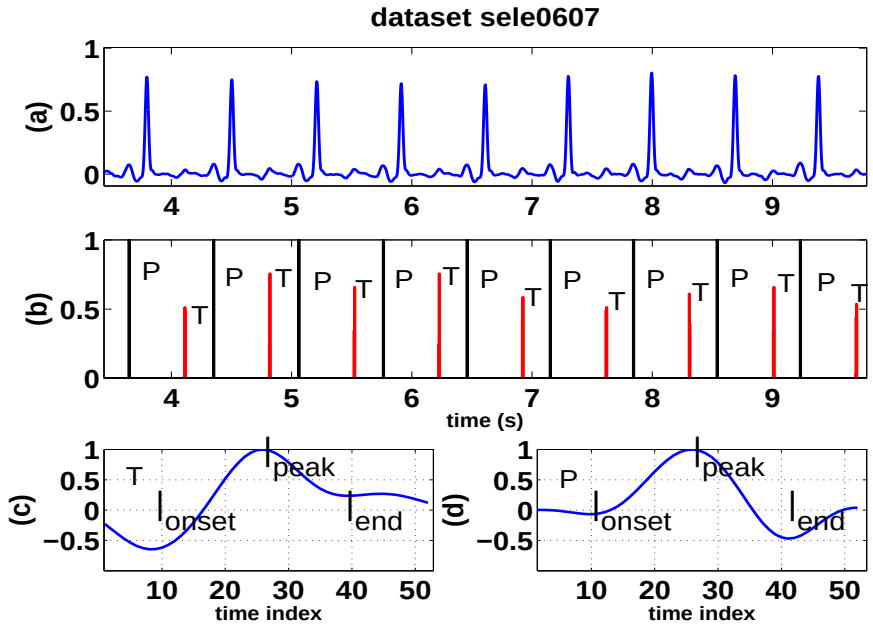


Figure 2.17: (a) an ECG signal portion from the QTDB "sele0607", (b)posterior distributions of the P and T wave indicator locations $p(\boldsymbol{b}_{\mathrm{P}})$ (in black) and $p(\boldsymbol{b}_{\mathrm{T}})$ (in dotted red), (c) estimated P and T-waveforms.

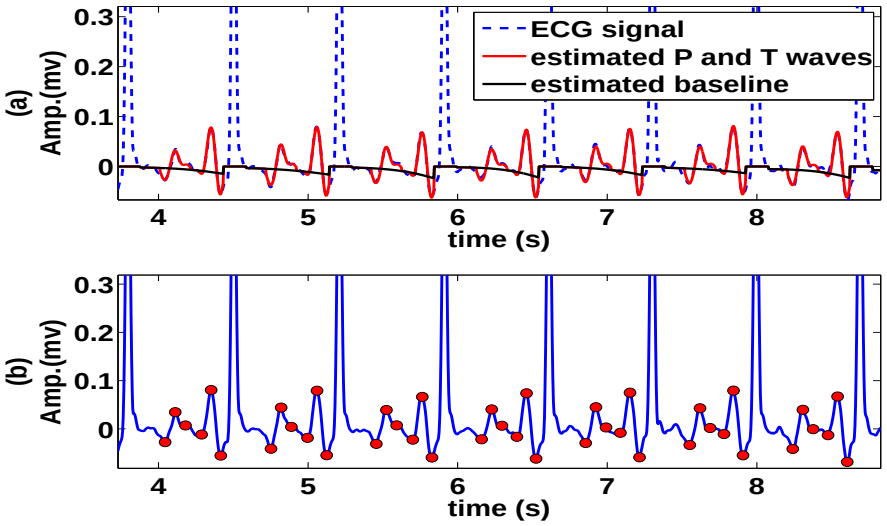


Figure 2.18: (a) ECG signal "sele0607" (dashed blue), estimated local baseline (black), and estimated P and T waves (red). (b) Results of P and T wave delineation.

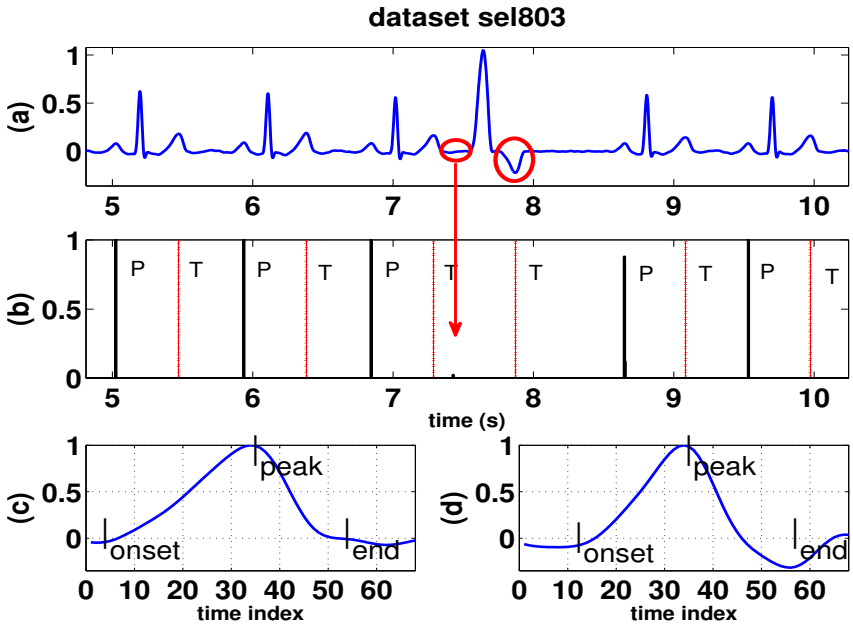


Figure 2.19: (a) an ECG signal portion from the QTDB "sel803", (b) posterior distributions of the P and T wave indicator locations $p(\boldsymbol{b}_{\mathrm{P}})$ (black) and $p(\boldsymbol{b}_{\mathrm{T}})$ (dotted red), (c) estimated P and T-waveforms.

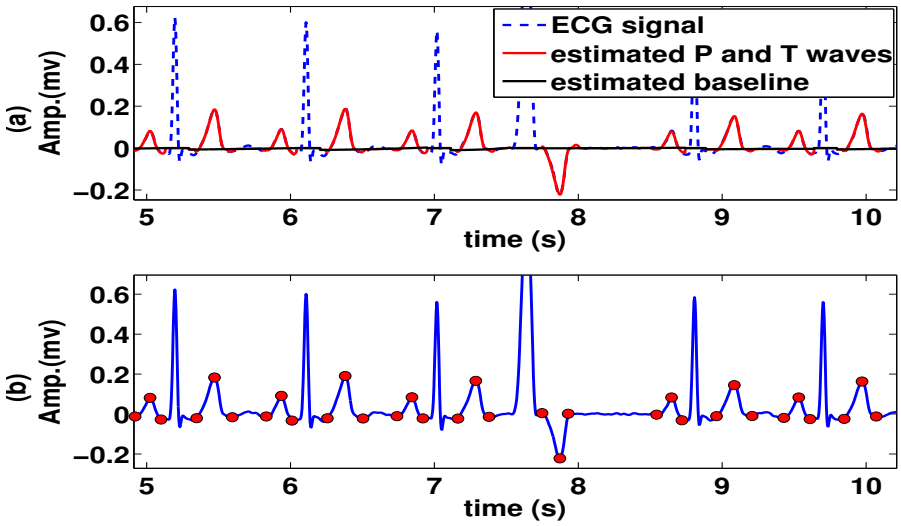


Figure 2.20: (a) ECG signal "sele0607" (dashed blue), estimated local baseline (black), and estimated P and T waves (red). (b) Results of P and T wave delineation.

Table 2.2: Block Gibbs sampler delineation performance compared to the PCGS.

Method	Parameters	P_{on}	P_{peak}	P_{end}	T_{on}	T_{peak}	T_{end}
Block Gibbs sampler	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
PCGS	annotations	3176	3176	3176	1345	3403	3403
	$Se\,(\%)$	98.93	98.93	98.93	99.01	99.81	99.81
	$P^{+}\,(\%)$	97.40	97.40	97.40	96.07	98.97	98.97
	$m \pm s\,(ms)$	3.7±17.3	4.1±8.6	-3.1±15.1	7.1±18.5	1.3±10.5	4.3±20.8

For a quantitative analysis, Table 2.2 presents the detection sensitivity (Se), the positive predictivity (P^{+}) and the means (μ) and standard deviations (σ) of the differences between the automated delineation results and the manual annotations, for the proposed block Gibbs sampler algorithm and for the previously introduced PCGS algorithm. It is seen that the block Gibbs sampler method outperforms the PCGS in terms of both detection sensitivity and delineation accuracy. This is partially due to the fact that the modified Bayesian model, which considers the local baseline, is more suitable to the ECG signal and thus can lead to better waveform estimations.

Based on the comparison of the two Gibbs sampling methods, it also should be noted that the complexity of the block Gibbs sampler is lower than that of the PCGS. The complexity of the sampler depends strongly on the number of hypotheses for which the probabilities have to be evaluated in one iteration. For the block Gibbs sampler, the T wave indicator of the nth search interval $\mathcal{J}_{\text{T},n}$ is sampled by evaluating $N_{\text{T},n}+1$ hypothesis (with $N_{\text{T},n}$ the length of the interval). For a D-beat processing window, the overall complexity per iteration equals the sum of the each individual T wave search interval length, i.e, $\sum_{n=1}^{D} N_{\text{T},n}$. While for the PCGS under the same condition, the overall average complexity per iteration is about K, which is the length of the whole D-beat processing window (please refer to [KTHD12] for detailed explication). Since the search intervals are defined as a fixed percentage of the heart cycles, we always have

$$\sum_{n=1}^{D} N_{\text{T},n} \leq K. \tag{2.54}$$

This can also be confirmed by the simulation. Running 100 iterations of the block Gibbs sampler algorithm for a 10-beat ECG block sampled at $F_s = 250$Hz takes approximately 4 seconds for a MATLAB® implementation on a 3.0-GHz Pentium IV, which is more than two times faster than the PCGS under the same condition.

2.5 Conclusion

This chapter studied a window based Bayesian model [LMT10] and its modified version [LKT+11, LTM+11] performing joint P and T wave delineation and waveform estimation. These models take into account appropriate prior distributions for the unknown parameters (wave locations and amplitudes, waveform coefficients). The prior distributions and the likelihood of the observed data are combined to provide the posterior distribution of the unknown parameters. The ECG state sequence composed of the P wave, QRS complex and T wave parameters is Markovian, since the current state (P wave, QRS complex, and T wave) only depends on the previous state, which means that there are at most one T wave and one P wave within each cardiac cycle. This property motivates our study of various Bayesian sampling algorithms imposing a strong local dependency on the wave locations. In our work, the local dependency is first exploited by a minimum-distance constraint prior on the wave indicator vector, and a PCGS is proposed to resolve the unknown parameters of this Bayesian model. Note that a similar framework has been proposed to OCT signal processing by Kail *et al.* [KNHH09]. Then, a block constraint prior is proposed to consider simultaneously the whole non-QRS component, and a block Gibbs sampler is studied to further improve the convergence and the computational efficiency of the Gibbs sampler. Concerning the parameter estimation, instead of deploying rigid detection and delineation criterion for all ECG time series, we used a local detection strategy (sample-based MMSE and blockwise MAP estimators) and a flexible delineation criteria based on the estimation of P and T waveforms in consecutive D-beat processing windows. The main contributions of this chapter are:

1. PCGS [LMT10]

   - The introduction of a window based Bayesian model for P and T wave delineation. This model is based on a modified Bernoulli-Gaussian sequence with minimum distance constraint for the wave locations and appropriate priors for the amplitudes, wave impulse responses and noise variance.
   - The derivation of a PCGS allowing for generation of samples distributed according to the posterior distribution associated to the Bayesian model. The proposed PCGS which takes into account the strong local dependency of the ECG waves overcomes the slow convergence problem encountered with the classical Gibbs sampler.
   - The proposed Bayesian model allows for simultaneous estimation of the P and T wave fiducial points and the P and T waveforms, which is rarely done by other ECG delineation methods.
   - The Bayesian simulation methods allow for determination of confidence intervals which provide reliability information about the estimates (e.g., see error bars in Fig. 2.8). This could be useful for medical diagnosis.

2. Block Gibbs sampler [LKT+11, LTM+11]

   - The window based Bayesian model is modified to consider simultaneously the whole non-QRS component. The local dependency of the ECG signal is otherwise expressed by a block constraint on the wave locations.

- A block Gibbs sampler is proposed to resolve the unknown parameters of the modified Bayesian model. Compared to the PCGS, it further improves the computational efficiency.
- Benefiting from the fact that the whole non-QRS component is considered together, the residual local baseline within each non-QRS component is modeled by a polynomial and its coefficients are estimated jointly with other wave parameters by the proposed block Gibbs sampler.
- Accurate yet parsimonious models are used for the P and T waveforms (modeled by Hermite expansions).
- A sample-based blockwise MAP estimator is used to estimate the wave indicator parameter. Compared to the MMSE wave indicator estimator for the PCGS, no probability threshold is needed to determine whether the wave is representative or not.

The resulting algorithms were validated using the entire annotated QT database. A comparison with other benchmark methods showed that the proposed method provides a reliable detection and an accurate delineation for a wide variety of wave morphologies. The most significant improvement was found in the P and T wave detection rate and the positive predictivity. In addition, the proposed method can provide accurate waveform estimation. Consequently, it allows for observation of the waveform evolution among processing windows. If we extract the T wave search region on every other beat rather than successively, the proposed method can be directly used to perform T wave alternans (TWA) analysis, which can be a pathological sign of a cardiac problem. Indeed, the wave amplitude can be used to decide the presence or absence of TWA, while the waveform estimation can reflect the characterization of TWA waveform. The application of the window based Bayesian model to TWA detection in clinical research background is presented in detail in Chapter 4.

Note however that the trade off with the good detection and estimation performance is a higher computational cost when compared to other more classical methods. It is also important to mention that the proposed Bayesian model relies on a nonoverlapped multiple-beat processing window to estimate the waveforms. More precisely, the waveforms of all P or T waves within a multiple-beat processing window were assumed to be unique, whereas their amplitudes and locations varied from one beat to another. Due to the pseudo-cyclostationary nature of the ECG signal, the P and T waveforms in a given beat are usually similar but not exactly equal to those of the adjacent beats. Therefore, the performance of P and T wave delineation can be expected to improve if the waveforms are estimated in a beat-to-beat manner that allows for temporal variations of waveform morphology across the beats. Moreover, it is interesting to process data on-line, to reduce storage costs and to ensure rapid adaptation to changing signal characteristics. Following this initiative, a beat-to-beat based Bayesian model will be discussed in the next chapter, and associated Bayesian methods are proposed to estimate the unknown parameters.

Chapter 3

Sequential beat-to-beat analysis of P and T waves using Bayesian algorithms and MCMC methods

Contents

3.1 Introduction

In this chapter, we study Bayesian methods that enable simultaneous P and T wave delineation and waveform estimation on a beat-to-beat basis. First, a beat-to-beat Bayesian model is proposed by introducing dependencies among waveform coefficients in the multiple-beat-window-based model studied in Chapter 2. Instead of assigning a white Gaussian prior to the temporal sequence of waveform coefficients, we use a prior "with memory" that depends on the estimates of the previous beat. A Gibbs sampler with a block constraint is then used for estimating the parameters of the resulting beat-to-beat model. Simulation results show that the proposed sequential model and processing improve the convergence of the samplers proposed in the previous chapter as well as the accuracy of estimating the locations, amplitudes, and

shapes of the P and T waves. The improved convergence can be explained by a considerable reduction of the parameter dimension, since only one beat is processed instead of multiple beats.

In the spirit of taking into account all the information contained in the past of the beat to be processed rather than considering only one previous beat, a sequential Monte Carlo (SMC) method is studied in the second part of this chapter. Following the SMC analysis principle, the sequential nature of the ECG signal is exploited by using a dynamic model under the Bayesian framework. Particle filters (PFs) are then investigated to estimate the unknown parameters of the dynamic model. A classical PF is first studied. The key idea of this filter is to represent the required posterior density function by a set of random samples (particles) with associated weights and to compute estimates based on these samples and weights. In a second step, a Rao-Blackwellized PF (also referred to as marginalized PF) is proposed to exploit the linear Gaussian sub-structure available in the state parameters of the dynamic model to obtain better estimates and to reduce the algorithm computational complexity. Simulation results show that the proposed dynamic model and the PFs are promising to resolve the sequential P and T wave analysis problem in the ECG signals.

3.2 Beat-to-beat Bayesian model and block Gibbs sampler

3.2.1 Signal model for one non-QRS interval

Similar to the window based approaches presented in Section 2.4, the non-QRS intervals which are located between the end of a QRS complex and the subsequent QRS onset are provided by a preliminary QRS detection step, using, e.g., the Pan-Tompkins algorithm [PT85]. The difference here is that only one non-QRS interval is processed at a given time instant.

As shown in Fig. 3.1, the non-QRS interval $\mathcal{J}_n$ associated with the nth heartbeat consists of two complementary subintervals: a T search interval $\mathcal{J}_{\mathrm{T},n}$, which may contain a T wave, and a P search interval $\mathcal{J}_{\mathrm{P},n}$, which may contain a P wave. The temporal lengths of the intervals $\mathcal{J}_n$, $\mathcal{J}_{\mathrm{T},n}$, and $\mathcal{J}_{\mathrm{P},n}$ will be denoted by N_n, $N_{\mathrm{T},n}$, and $N_{\mathrm{P},n}$, respectively. Note that $N_{\mathrm{T},n} + N_{\mathrm{P},n} = N_n$. The lengths $N_{\mathrm{T},n}$ and $N_{\mathrm{P},n}$ can be determined by a cardiologist or simply as fixed percentages of N_n. Our goal is to estimate the locations, amplitudes, and shapes (waveforms) of the T and P waves within their respective search intervals $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$. Note that only the locations of the wave peaks are constrained to lie within their respective search intervals. The whole non-QRS signal component will be considered together in the processing so that the border situation (where a T or P wave has part of its waveform across the border between $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$) can be properly handled.

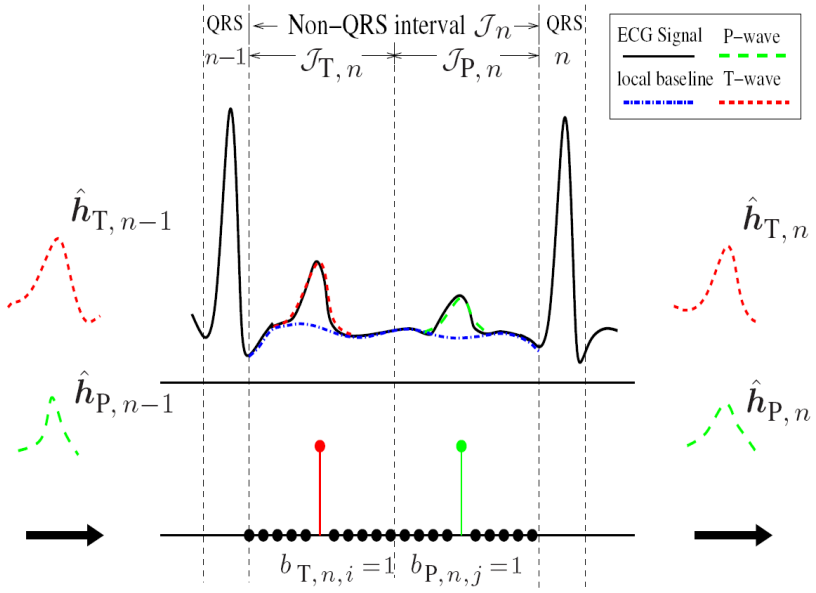


Figure 3.1: Signal model for the beat-to-beat processing scheme.

The signal in the non-QRS interval $\mathcal{J}_n$ can be approximated by two pulses representing the T and P waves plus a local baseline (see Fig. 3.1). Similar to the blind deconvolution problem in the multi-beat case, the T and P waves within $\mathcal{J}_n$ are modeled by the convolution of an unknown binary "indicator sequence" $\boldsymbol{b}_n = (b_{n,1} \cdots b_{n,N_n})^T$ indicating the wave locations ($b_{n,i} = 1$ if there is a wave at location i, $b_{n,i} = 0$ otherwise) with unknown T and P waveforms

$\boldsymbol{h}_{\mathrm{T},n} = (h_{\mathrm{T},n,-L} \cdots h_{\mathrm{T},n,L})^T$ and $\boldsymbol{h}_{\mathrm{P},n} = (h_{\mathrm{P},n,-L} \cdots h_{\mathrm{P},n,L})^T$. 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 T and P waves. The indicators associated with the T search interval $\mathcal{J}_{\mathrm{T},n}$ and the P search interval $\mathcal{J}_{\mathrm{P},n}$ are $\{b_{n,k}\}_{k=1}^{N_{\mathrm{T},n}}$ and $\{b_{n,k}\}_{k=N_{\mathrm{T},n}+1}^{N_n}$, respectively; in each case, at most one indicator is nonzero because at most one wave may be present in any given search interval. Note that contrary to the signal models proposed for multi-beat window based approaches (see Section 2.2.2 and Section 2.4), the amplitude vector associated to the sparse sequence is omitted in the beat-to-beat model. This is due to the fact that waveform coefficients are estimated in a beat-to-beat basis which makes the amplitude parameter redundant.

From the above discussion, it follows that the nth non-QRS signal component can be expressed by the convolution relation

$$x_{n,k} = \sum_{j=1}^{N_{\mathrm{T},n}} h_{\mathrm{T},n,k-j} b_{n,j} + \sum_{j=N_{\mathrm{T},n}+1}^{N_n} h_{\mathrm{P},n,k-j} b_{n,j} + c_{n,k} + w_{n,k}\,, \quad k \in \mathcal{J}_n. \tag{3.1}$$

Here, $c_{n,k}$ denotes the baseline sequence and $w_{n,k}$ is a white Gaussian noise with unknown variance $\sigma_{w,n}^2$. The indicators $b_{n,k}$ are constrained as discussed above. Furthermore, we have set $h_{\mathrm{T},n,k} = h_{\mathrm{P},n,k} = 0$ for $k \notin \{-L, \ldots, L\}$.

Following the idea of Section 2.4.1, the T and P waveforms are represented by a basis expansion using discrete-time versions of Hermite functions such that

$$\boldsymbol{h}_{\mathrm{T},n} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n}\,, \quad \boldsymbol{h}_{\mathrm{P},n} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P},n} \tag{3.2}$$

where $\boldsymbol{H}$ is a $(2L+1) \times G$ matrix whose columns are the first G Hermite functions (with $G \leq 2L+1$), suitably sampled and truncated to length $2L+1$, and $\boldsymbol{\alpha}_{\mathrm{T},n}$ and $\boldsymbol{\alpha}_{\mathrm{P},n}$ are unknown coefficient vectors of length G. Similarly, the residual local baseline within the non-QRS interval is modeled by using a polynomial of degree 4, i.e.,

$$c_{n,k} = \sum_{i=1}^{5} \gamma_{n,i} k^{i-1}, \quad k \in \mathcal{J}_n\,. \tag{3.3}$$

In vector-matrix form, (3.3) reads as $\boldsymbol{c}_n = \boldsymbol{M}\boldsymbol{\gamma}_n$, with the known $N_n \times 5$ Vandermonde matrix $\boldsymbol{M}$ and the unknown coefficient vector $\boldsymbol{\gamma}_n = (\gamma_{n,1} \cdots \gamma_{n,5})^T$.

Using (3.2) and (3.3), we obtain the following vector representation of the non-QRS signal in (3.1)

$$\boldsymbol{x}_n = \boldsymbol{B}_{\mathrm{T},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n} + \boldsymbol{B}_{\mathrm{P},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P},n} + \boldsymbol{M}\boldsymbol{\gamma}_n + \boldsymbol{w}_n \tag{3.4}$$

where $\boldsymbol{x}_n = (x_{n,1} \cdots x_{n,N_n})^T$, $\boldsymbol{w}_n = (w_{n,1} \cdots w_{n,N_n})^T$, $\boldsymbol{B}_{\mathrm{T},n}$ is the $N_n \times (2L+1)$ Toeplitz matrix with first row $(b_{n,L+1} \cdots b_{n,1}\ 0 \cdots 0)$ and first column $(b_{n,L+1} \cdots b_{n,N_{\mathrm{T},n}}\ 0 \cdots 0)^T$, and $\boldsymbol{B}_{\mathrm{P},n}$ is the $N_n \times (2L+1)$ Toeplitz matrix with last row $(0 \cdots 0\ b_{n,N_n} \cdots b_{n,N_n-L})$ and last column $(0 \cdots 0\ b_{n,N_{\mathrm{T},n}+1} \cdots\ b_{n,N_n-L})^T$.

3.2.2 Bayesian inference

Due to the parametrization introduced in Section 3.2.1, the unknown parameters for the nth non-QRS interval $\mathcal{J}_n$ are included in the random vector $\boldsymbol{\theta}_n \triangleq (\boldsymbol{b}_n^T\, \boldsymbol{\alpha}_{\mathrm{T},n}^T\, \boldsymbol{\alpha}_{\mathrm{P},n}^T\, \boldsymbol{\gamma}_n^T\, \sigma_{w,n}^2)^T$. Note, in particular, that the noise variance $\sigma_{w,n}^2$ is modeled as a random parameter. Bayesian detection/estimation relies on the posterior distribution

$$p(\boldsymbol{\theta}_n|\boldsymbol{x}_n) \propto p(\boldsymbol{x}_n|\boldsymbol{\theta}_n)p(\boldsymbol{\theta}_n) \tag{3.5}$$

where $\propto$ means "equal up to a multiplicative 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.

Likelihood Function

Using (3.4) and the fact that $w_{n,k}$ is white and Gaussian with variance $\sigma_{w,n}^2$, the likelihood function is obtained as

$$p(\boldsymbol{x}_n|\boldsymbol{\theta}_n) = \mathcal{N}(\boldsymbol{B}_{\mathrm{T},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n} + \boldsymbol{B}_{\mathrm{P},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P},n} + \boldsymbol{M}\boldsymbol{\gamma}_n,\, \sigma_{w,n}^2\mathbf{I}_{N_n}) \tag{3.6}$$

where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$ denotes the multivariate Gaussian probability density function with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}$ and $\|\cdot\|$ denotes the ℓ_2 norm, i.e., $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$.

Prior Distributions

Let the vectors $\boldsymbol{b}_{\mathrm{T},n} \in \{0,1\}^{N_{\mathrm{T},n}}$ and $\boldsymbol{b}_{\mathrm{P},n} \in \{0,1\}^{N_{\mathrm{P},n}}$ contain all entries of the wave indicator vector $\boldsymbol{b}_n$ whose locations are, respectively, in the T search interval $\mathcal{J}_{\mathrm{T},n}$ and in the P search interval $\mathcal{J}_{\mathrm{P},n}$. Note that $\boldsymbol{b}_n = (\boldsymbol{b}_{\mathrm{T},n}^T\, \boldsymbol{b}_{\mathrm{P},n}^T)^T$. The indicators $b_{n,k}$ are subject to a *block constraint*: within $\mathcal{J}_{\mathrm{T},n}$, there is one T wave (thus, $\|\boldsymbol{b}_{\mathrm{T},n}\| = 1$) 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

$$p(\boldsymbol{b}_{\mathrm{T},n}) = \begin{cases} p_0 & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 0 \\ p_1 & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

where $p_1 = (1-p_0)/N_{\mathrm{T},n}$ and p_0 is chosen very small. Note that there are $N_{\mathrm{T},n}$ vectors satisfying $\|\boldsymbol{b}_{\mathrm{T},n}\| = 1$ whereas the zero vector is the only vector satisfying $\|\boldsymbol{b}_{\mathrm{T},n}\| = 0$. Thus the probabilities in (3.7) sum to one. 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 (3.7), with the same parameters p_0 and p_1. The vectors $\boldsymbol{b}_{\mathrm{T},n}$ and $\boldsymbol{b}_{\mathrm{P},n}$ are assumed to be statistically independent. The prior of $\boldsymbol{b}_n$ results as

$$p(\boldsymbol{b}_n) = p(\boldsymbol{b}_{\mathrm{T},n})\, p(\boldsymbol{b}_{\mathrm{P},n}). \tag{3.8}$$

The wave indicator vectors $\boldsymbol{b}_n$ for different search intervals (different n) are assumed to be statistically independent.

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 estimates obtained in the $(n-1)$th non-QRS interval $\mathcal{J}_{n-1}$. This is a major difference from the multi-beat processing window Bayesian model in Section 2.4.2. Consider the T wave as an example. The prior of $\boldsymbol{\alpha}_{\mathrm{T},n}$ is defined as

$$p(\boldsymbol{\alpha}_{\mathrm{T},n}|\boldsymbol{b}_{\mathrm{T},n},\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) = \begin{cases} \delta(\boldsymbol{\alpha}_{\mathrm{T},n}-\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\|=0 \\ \mathcal{N}(\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1},\sigma_\alpha^2\mathbf{I}_G) & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\|=1 \end{cases} \tag{3.9}$$

where $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$ is the estimate of the T waveform coefficient vector associated with $\mathcal{J}_{n-1}$ and $\mathbf{I}_G$ is the identity matrix of size $G\times G$. For the variance σ_α^2, we choose a value that allows for 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 $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$, i.e., the waveform coefficients are assigned the same values as those estimated in $\mathcal{J}_{\mathrm{T},n-1}$. The prior of the P waveform coefficient vector $\boldsymbol{\alpha}_{\mathrm{P},n}$ is defined in an analogous way, with $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$ replaced by $\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}$. These definitions of the priors associated with $\boldsymbol{\alpha}_{\mathrm{T},n}$ and $\boldsymbol{\alpha}_{\mathrm{P},n}$ introduce a memory in the statistical model for the T and P waveforms and, in turn, results in a sequential type of processing.

Concerning the baseline coefficients and the noise variance, they are assigned similar priors as in the modified window based Bayesian model in Section 2.4.2. The baseline coefficient vectors $\boldsymbol{\gamma}_n$ for different n are assumed to be independent. Moreover, the baseline coefficients $\gamma_{n,i}$ for a given n are modeled as independent and identically distributed (iid) zero-mean Gaussian, i.e.,

$$p(\boldsymbol{\gamma}_n)=\mathcal{N}(\mathbf{0},\sigma_\gamma^2\mathbf{I}_5), \tag{3.10}$$

with a fixed variance σ_γ^2. The noise variances $\sigma_{w,n}^2$ for different n are modeled as independent and distributed according to the inverse gamma distribution

$$p(\sigma_{w,n}^2)=\mathcal{IG}(\xi,\eta), \tag{3.11}$$

where ξ and η are fixed hyperparameters providing a vague prior.

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

$$\begin{aligned} p(\boldsymbol{\theta}_n|\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1},\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}) &= p(\boldsymbol{\alpha}_{\mathrm{T},n}|\boldsymbol{b}_{\mathrm{T},n},\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1})\,p(\boldsymbol{b}_{\mathrm{T},n}) \\ &\times p(\boldsymbol{\alpha}_{\mathrm{P},n}|\boldsymbol{b}_{\mathrm{P},n},\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1})\,p(\boldsymbol{b}_{\mathrm{P},n})\,p(\boldsymbol{\gamma}_n)\,p(\sigma_{w,n}^2) \end{aligned} \tag{3.12}$$

where (3.8) has been used.

Posterior distribution

The posterior distribution of the parameter vector $\boldsymbol{\theta}_n$ can be derived using Bayes' rule, i.e.,

$$p(\boldsymbol{\theta}_n|\boldsymbol{x}_n, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}) \propto p(\boldsymbol{x}_n|\boldsymbol{\theta}_n)p(\boldsymbol{\theta}_n|\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}) \tag{3.13}$$

where the different densities have been defined in (3.6), (3.12), and the individual prior distribution have been discussed previously.

Due to the complexity of the resulting posterior distribution, we propose a block Gibbs sampler 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 the next section). This sampler has similar structure to the block Gibbs sampler proposed in Section 2.4.3, except that the sampling distributions for certain parameters are different due to the dependencies on the estimates of the previous non-QRS component and the steps to generate wave amplitude samples are omitted. From these samples, the discrete parameters $\boldsymbol{b}_n = (\boldsymbol{b}_{\mathrm{T},n}^T \boldsymbol{b}_{\mathrm{P},n}^T)^T$ are then detected by means of the sample-based maximum a posteriori (MAP) detector, and the continuous parameters $\boldsymbol{\alpha}_{\mathrm{T},n}$, $\boldsymbol{\alpha}_{\mathrm{P},n}$, $\boldsymbol{\gamma}_n$, and $\sigma_{w,n}^2$ are estimated by means of the sample-based minimum mean square error (MMSE) estimator, as described in Section 2.4.4.

3.2.3 Block Gibbs sampler for beat-to-beat wave extraction

The proposed block Gibbs sampler for the nth non-QRS interval $\mathcal{J}_n$ is summarized in Algorithm 6. Note that, the interval index n is omitted from all parameters to simplify the notation, while the index $n-1$ is kept to avoid any ambiguity.

Algorithm 6 Block Gibbs sampler

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

To see that Algorithm 6 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 $p(\boldsymbol{b}_{\mathrm{T}}, \boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x})$, and similarly for $\boldsymbol{b}_{\mathrm{P}}$ and $\boldsymbol{\alpha}_{\mathrm{P}}$. The sampling distributions used in Algorithm 6 are specified in the following and their derivations are provided in the Appendix C.

Wave indicators. The sampling distribution for the T wave indicator vector $\boldsymbol{b}_{\mathrm{T}}$ is

$$p(\boldsymbol{b}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x}) \propto \begin{cases} p_0 & \text{if } \|\boldsymbol{b}_{\mathrm{T}}\| = 0 \\ \sigma_\alpha^{-G}\sqrt{|\boldsymbol{\Sigma}_1|}\exp(\boldsymbol{\mu}_1\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1)\, p_1 & \text{if } \|\boldsymbol{b}_{\mathrm{T}}\| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

with

$$\boldsymbol{\mu}_1 = \boldsymbol{\Sigma}_1\left(\frac{\boldsymbol{H}^T\boldsymbol{B}_{\mathrm{T}}^T\tilde{\boldsymbol{x}}_{\mathrm{T}}}{\sigma_w^2} + \frac{\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}}{\sigma_\alpha^2}\right) \tag{3.15}$$

$$\boldsymbol{\Sigma}_1 = \left(\frac{\boldsymbol{H}^T\boldsymbol{B}_{\mathrm{T}}^T\boldsymbol{B}_{\mathrm{T}}\boldsymbol{H}}{\sigma_w^2} + \frac{\mathbf{I}_G}{\sigma_\alpha^2}\right)^{-1}. \tag{3.16}$$

Here, $\tilde{\boldsymbol{x}}_{\mathrm{T}} \triangleq \boldsymbol{x} - \boldsymbol{B}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}} - \boldsymbol{M}\boldsymbol{\gamma}$. The sampler evaluates all hypotheses for $\boldsymbol{b}_{\mathrm{T}}$ conditioned on the current samples of all other parameters. There are $N_{\mathrm{T}}+1$ such hypotheses, because, according to (3.14), $\boldsymbol{b}_{\mathrm{T}}$ has either no 1-entry or exactly one 1-entry at one of N_{T} possible locations within $\mathcal{J}_{\mathrm{T}}$.

The sampling distribution for the P wave indicator vector, $p(\boldsymbol{b}_{\mathrm{P}}|\boldsymbol{b}_{\mathrm{T}}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}, \boldsymbol{\alpha}_{\mathrm{T}}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x})$, is obtained in an analogous manner, with $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$ replaced by $\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}$, $\boldsymbol{B}_{\mathrm{T}}$ replaced by $\boldsymbol{B}_{\mathrm{P}}$, and $\tilde{\boldsymbol{x}}_{\mathrm{T}}$ replaced by $\tilde{\boldsymbol{x}}_{\mathrm{P}} \triangleq \boldsymbol{x} - \boldsymbol{B}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}} - \boldsymbol{M}\boldsymbol{\gamma}$.

Waveform coefficients. The sampling distribution for $\boldsymbol{\alpha}_{\mathrm{T}}$ is

$$p(\boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x}) \propto \begin{cases} \delta(\boldsymbol{\alpha}_{\mathrm{T}} - \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) & \text{if } \|\boldsymbol{b}_{\mathrm{T}}\| = 0, \\ \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) & \text{if } \|\boldsymbol{b}_{\mathrm{T}}\| = 1 \end{cases} \tag{3.17}$$

with $\boldsymbol{\mu}_1$ and $\boldsymbol{\Sigma}_1$ as defined above. In particular, the samples of $\boldsymbol{\alpha}_{\mathrm{T}}$ are given by $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$ if there is no T wave in the search interval ($\|\boldsymbol{b}_{\mathrm{T}}\| = 0$). The sampling distribution for $\boldsymbol{\alpha}_{\mathrm{P}}$, $p(\boldsymbol{\alpha}_{\mathrm{P}}|\boldsymbol{b}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}, \boldsymbol{\alpha}_{\mathrm{T}}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x})$, is obtained similarly.

Baseline coefficients. The sampling distribution for $\boldsymbol{\gamma}$ is

$$p(\boldsymbol{\gamma}|\boldsymbol{b}, \boldsymbol{\alpha}_{\mathrm{T}}, \boldsymbol{\alpha}_{\mathrm{P}}, \sigma_w^2, \boldsymbol{x}) = \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \tag{3.18}$$

with

$$\boldsymbol{\mu}_2 = \frac{\boldsymbol{\Sigma}_2\boldsymbol{M}^T}{\sigma_w^2}\tilde{\boldsymbol{x}} \tag{3.19}$$

$$\boldsymbol{\Sigma}_2 = \left(\frac{\boldsymbol{M}^T\boldsymbol{M}}{\sigma_w^2} + \frac{\mathbf{I}_5}{\sigma_\gamma^2}\right)^{-1} \tag{3.20}$$

where $\tilde{\boldsymbol{x}} \triangleq \boldsymbol{x} - \boldsymbol{B}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}} - \boldsymbol{B}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}$.

Noise variance. The sampling distribution for σ_w^2 is the following inverse gamma distribution

$$p(\sigma_w^2|\boldsymbol{b}, \boldsymbol{\alpha}_\mathrm{T}, \boldsymbol{\alpha}_\mathrm{P}, \boldsymbol{\gamma}, \boldsymbol{x}) = \mathcal{IG}(\xi', \eta') \tag{3.21}$$

with

$$\xi' = \xi + \frac{N}{2}\,, \qquad \eta' = \eta + \frac{1}{2}\|\tilde{\boldsymbol{x}} - \boldsymbol{M}\boldsymbol{\gamma}\|^2.$$

3.2.4 Wave parameter estimation and threshold-free delineation

Wave detection and parameter estimation

The wave detection and parameter estimation can be obtained similarly to the sample-based estimation strategy developed in Section 2.4.4 for the window based method. We denote by $\mathcal{S} \triangleq \{\boldsymbol{b}_\mathrm{T}^{(i)}, \boldsymbol{b}_\mathrm{P}^{(i)}, \boldsymbol{\alpha}_\mathrm{T}^{(i)}, \boldsymbol{\alpha}_\mathrm{P}^{(i)}, \boldsymbol{\gamma}^{(i)}, \sigma_w^{2(i)}\}_{i=1}^{N_\mathrm{s}}$ the set of samples produced by the beat-to-beat block Gibbs sampler after a burn-in period. The sample-based blockwise MAP detector for the wave indicator vector $\boldsymbol{b}$ is given by

$$\hat{\boldsymbol{b}}_\mathrm{T} = \underset{\boldsymbol{b}_\mathrm{T}^{(i)} \in \mathcal{S}}{\operatorname{argmax}}\, p_\mathcal{S}(\boldsymbol{b}_\mathrm{T}^{(i)})\,, \qquad \hat{\boldsymbol{b}}_\mathrm{P} = \underset{\boldsymbol{b}_\mathrm{P}^{(i)} \in \mathcal{S}}{\operatorname{argmax}}\, p_\mathcal{S}(\boldsymbol{b}_\mathrm{P}^{(i)})\,. \tag{3.22}$$

Here, $p_\mathcal{S}(\boldsymbol{b}_\mathrm{T})$ is a sample-based estimate 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 $\boldsymbol{b}_\mathrm{T}^{(i)}$ in $\mathcal{S}$ that equal the respective value of $\boldsymbol{b}_\mathrm{T}$, normalized by the total number of samples, N_s. Recall that, due to (3.14), each sample $\boldsymbol{b}_\mathrm{T}^{(i)}$ contains at most one non-null value. Thus, the same is true for $\hat{\boldsymbol{b}}_\mathrm{T}$. Analogous considerations apply to $p_\mathcal{S}(\boldsymbol{b}_\mathrm{P})$ and $\hat{\boldsymbol{b}}_\mathrm{P}$.

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

$$\hat{\boldsymbol{\theta}}_{\sim\boldsymbol{b}} = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \boldsymbol{\theta}_{\sim\boldsymbol{b}}^{(i)} \tag{3.23}$$

where $\boldsymbol{\theta}_{\sim\boldsymbol{b}}^{(i)} \triangleq \left(\boldsymbol{\alpha}_\mathrm{T}^{(i)T}\, \boldsymbol{\alpha}_\mathrm{P}^{(i)T}\, \boldsymbol{\gamma}^{(i)T}\, \sigma_w^{2(i)}\right)^T$ and $|\mathcal{I}|$ denotes the number of elements in $\mathcal{I}$. This can be interpreted as a sample-based approximation of the MMSE estimator 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 $\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}$.

Threshold-free wave delineation

For the wave peak location, the detection/estimation results obtained for the wave indicators b_k and the waveforms $h_{\mathrm{T},k}$ and $h_{\mathrm{P},k}$ (as represented by the waveform coefficients $\boldsymbol{\alpha}_\mathrm{T}$ and $\boldsymbol{\alpha}_\mathrm{P}$) are ambiguous with respect to their relative locations, in the sense that a temporal shift of b_k

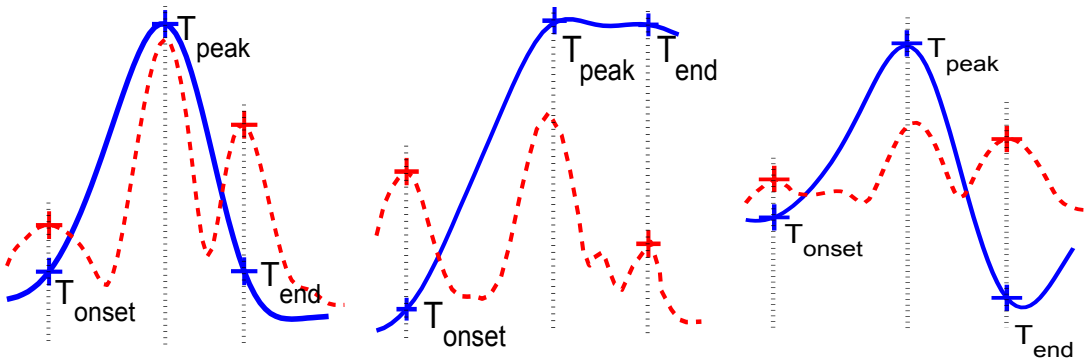
can be compensated for by an inverse temporal shift of $h_{\mathrm{T},k}$ and $h_{\mathrm{P},k}$. The same technique as in Section 2.3.4 is applied to resolve this time-shift ambiguity issue. It ensures that the maximum of the waveform is located at the center $k=0$ of the waveform support interval $\{-L,\ldots,L\}$ and, thus, the position of a nonzero detected indicator $\hat{b}_k = 1$ directly indicates the peak of the respective T or P wave.

The wave delineation consists of determining the boundaries of the detected T and P waves. It is broadly accepted that the significant turning points determined 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 [TS90, CLK06]. The curvature of the estimated T waveform $\hat{h}_{\mathrm{T},k}$ is defined as [TS90]

$$\kappa_{\mathrm{T},k} \triangleq \frac{\hat{h}''_{\mathrm{T},k}}{\left[1+\left(\hat{h}'_{\mathrm{T},k}\right)^2\right]^{3/2}}, \quad k \in \{-L,\ldots,L\} \tag{3.24}$$

where $\hat{h}'_{\mathrm{T},k}$ and $\hat{h}''_{\mathrm{T},k}$ are discrete-time approximations of the first and second derivatives of $\hat{h}_{\mathrm{T},k}$ (e.g., $\hat{h}'_{\mathrm{T},k}$ is defined as the difference $\hat{h}_{\mathrm{T},k} - \hat{h}_{\mathrm{T},k-1}$).

Fig. 3.2 illustrates the method by showing the delineation results obtained for three different T wave morphologies. The largest local maximum of the curvature before and after the wave peak reveals the most significant turning points of the waveform under visual observation. Thus they can be seen as the onset and the end of the wave. Note that due to the good waveform estimation resulting from the Hermite basis decomposition and the proposed beat-to-beat Gibbs sampler, the corresponding curvature properly reflects the wave morphology characteristics. Contrary to most of the state-of-art methods and to the criteria proposed in Section 2.3.4, the delineation method described above avoids the use of rigid detection and delineation thresholds, which is a nice property for robust wave delineation.



(a) (b) (c)

Figure 3.2: 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) [LMGM97] dataset sel17453 channel 1; (b) ascending T wave from QTDB sele0203 channel 1; (c) biphasic T wave from QTDB sele0603 channel 1.

3.2.5 Simulation results

As in the previous studies of this work, we have evaluated the proposed beat-to-beat wave detection/estimation/delineation method on the QT database (QTDB) [LMGM97]. In a preprocessing step, the QRS complexes are detected and the borders of the non-QRS intervals $\mathcal{J}_n$ are determined using the Pan-Tompkins algorithm [PT85]. (The same preprocessing step was performed in Section 2.3.5 and Section 2.4.5 for the multi-beat processing window based approach.) T and P search intervals $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$ are then defined as the first and second half of $\mathcal{J}_n$. The proposed method sequentially processes one non-QRS interval $\mathcal{J}_n$ after another. For each non-QRS interval, the block Gibbs sampler generates 100 samples according to the conditional distributions specified in Section 3.2.3. The first 40 samples constituted the burn-in period, and the remaining 60 are used for detection/estimation (thus, $N_{\mathrm{s}} = 60$). The fixed hyperparameters involved in the prior distributions have been chosen as $p_0 = 0.01$, $\sigma_\gamma^2 = 0.1$, $\sigma_\alpha^2 = 1$, $\xi = 11$, 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 $\sigma_{w,n}^2$. Note that the non-QRS components are normalized by using the corresponding R peak values to handle different amplitude resolutions. For the first non-QRS interval ($n=1$), the previous waveform coefficient estimates $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$ and $\hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}$ have been initialized with a $2L+1$ Hanning window [Har78] whose amplitude is half the R peak.

Qualitative Analysis

In this section, we show the posterior distributions of the indicators b_k as well as estimation and delineation results for some typical examples, and we present a qualitative comparison with the previously proposed window-based Bayesian method of Section 2.4.

Fig. 3.3(a) shows two consecutive heartbeats from the QTDB dataset sele0136. The corresponding sample-based estimates of the marginal posterior probabilities of having a T or P wave at a given location k, $P_{\mathcal{S}}(b_k = 1)$, are depicted in Fig. 3.3(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 position k, and similarly for $k \in \mathcal{J}_{\mathrm{P}}$.) Fig. 3.3(c) shows the T and P waveforms estimated by the proposed beat-to-beat 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 3.2.4). As can be seen, there are noticeable differences between the two consecutive T waveform estimates (at time instants 4.92s and 6.10s), as well as between the two consecutive P waveform estimates (at time instants 5.64s and 6.83s). This confirms the pseudo-stationary 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 T and P waveforms.

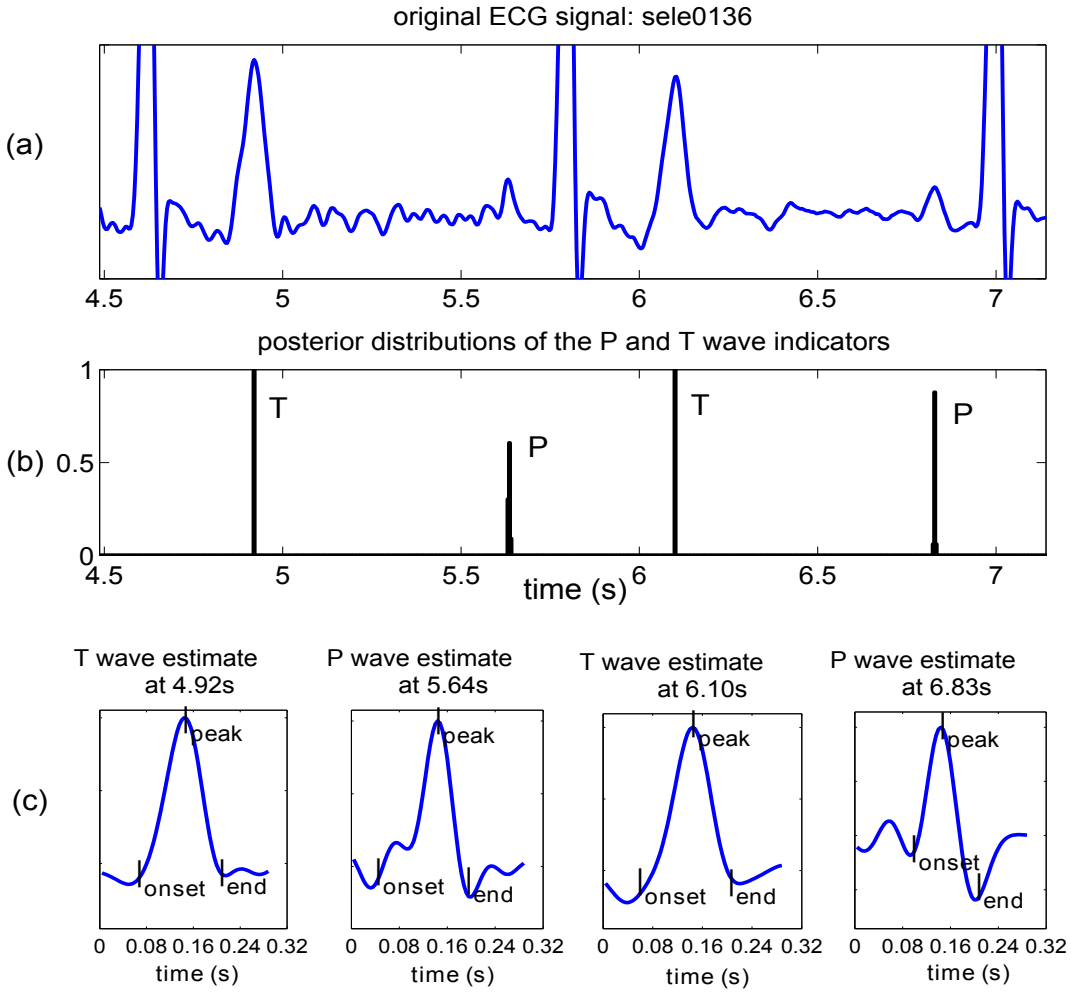


Figure 3.3: (a) Two consecutive heartbeats from QTDB dataset sele0136; (b) estimated marginal posteriors $P_{\mathcal{S}}(b_k=1)$; (c) estimated T and P waveforms and delineation results.

Next, we present a qualitative comparison of the proposed beat-to-beat method with the recently proposed Gaussian mixture model and extended Kalman filter (EKF) method of [SS09] and the window-based Bayesian method of Section 2.4. To evaluate the two methods under real physiological noise conditions, we have added electrode motion (EM) and muscular activity (MA) noise from the MIT-BIH noise stress test database. Since EM noise is known to be easier to handle than MA noise, we only present our results with MA noise. Our results with EM noise were similar or better. Fig. 3.4(a) shows a segment of QTDB dataset sele0136. Fig. 3.4(b) shows the same segment corrupted by MA noise with a signal-to-noise ratio (SNR) of 10dB. Fig. 3.4(c), Fig. 3.4(d) and Fig. 3.4(e) depict the non-QRS signal component obtained from the noisy signal by the method of [SS09], the window based block Gibbs sampler method and by the beat-to-beat block Gibbs sampler method, respectively. The original (noise-free) ECG signal is also shown for comparison. It can be seen that the proposed method provides a closer agreement with the original ECG signal, especially at the onsets and ends of the waves, which is a desirable property for wave delineation.

Fig. 3.5 shows analogous results for a segment of QTDB dataset sel803 that contains premature ventricular contractions (PVCs). The proposed beat-to-beat method is seen to exhibit good performance even in the presence of nonmonotonic morphological abnormalities. Again, the estimates obtained with the proposed method are better than those obtained with the window-based method.

Because the proposed beat-to-beat method processes only one non-QRS interval considered at any given time, both its memory requirements and its computational load are smaller than those of the window-based method in Section 2. For instance, for the proposed method using 100 sampler iterations, the processing time per beat is approximately 0.3s for a nonoptimized MATLAB implementation running on a 3.0-GHz Pentium IV computer, compared to about 2s for the PCGS based method in Section 2.3 .

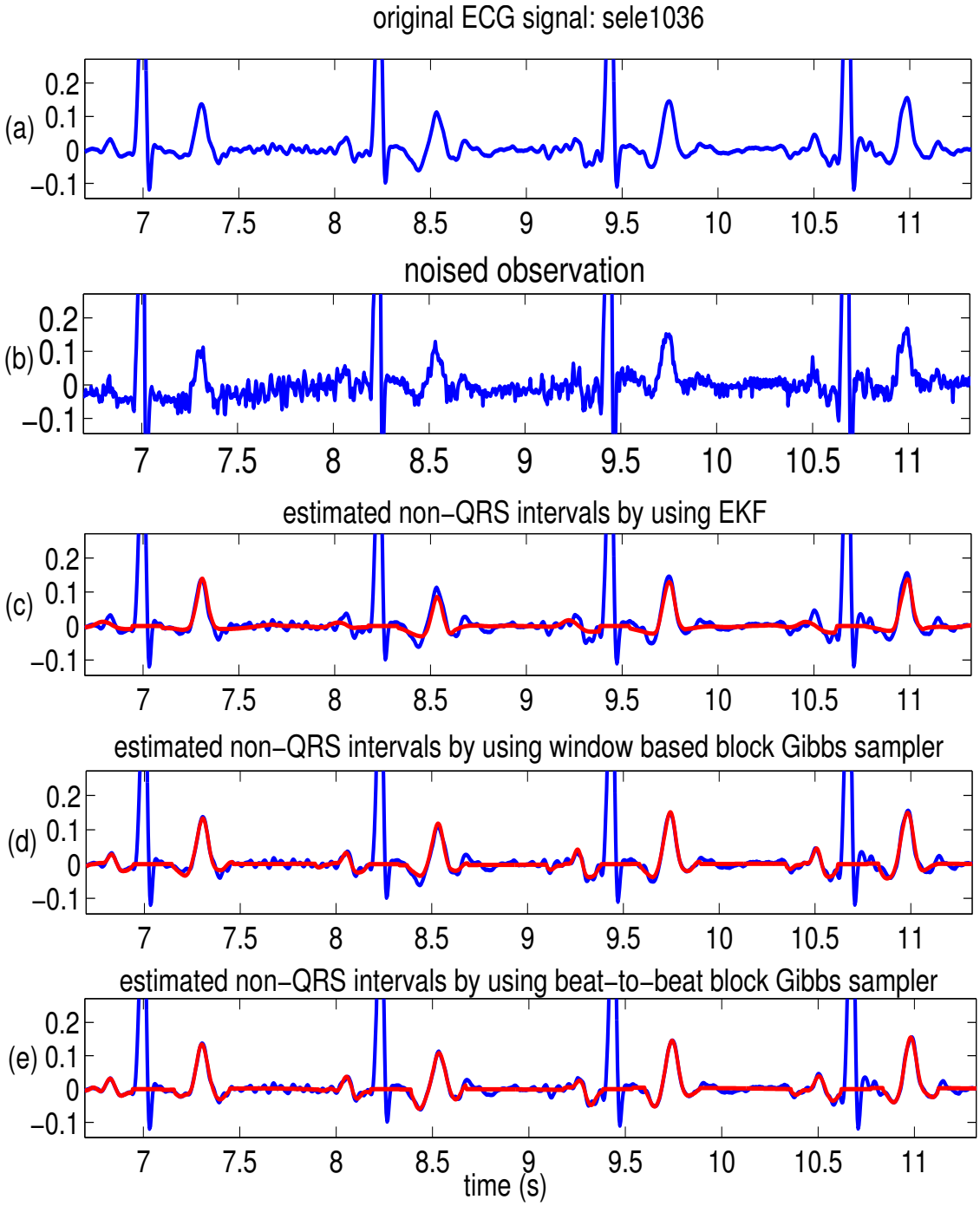


Figure 3.4: (a) Segment from QTDB "sele0136"; (b) noisy version including MA noise with SNR = 10dB; (c) non-QRS signal component estimated by the EKF method of [SS09] (red) and noise-free original signal (blue); (d) non-QRS signal component estimated by the window-based method of Section 2.4 (red) and noise-free original signal (blue); (e) non-QRS signal component estimated by the proposed beat-to-beat method (red) and noise-free original signal (blue).

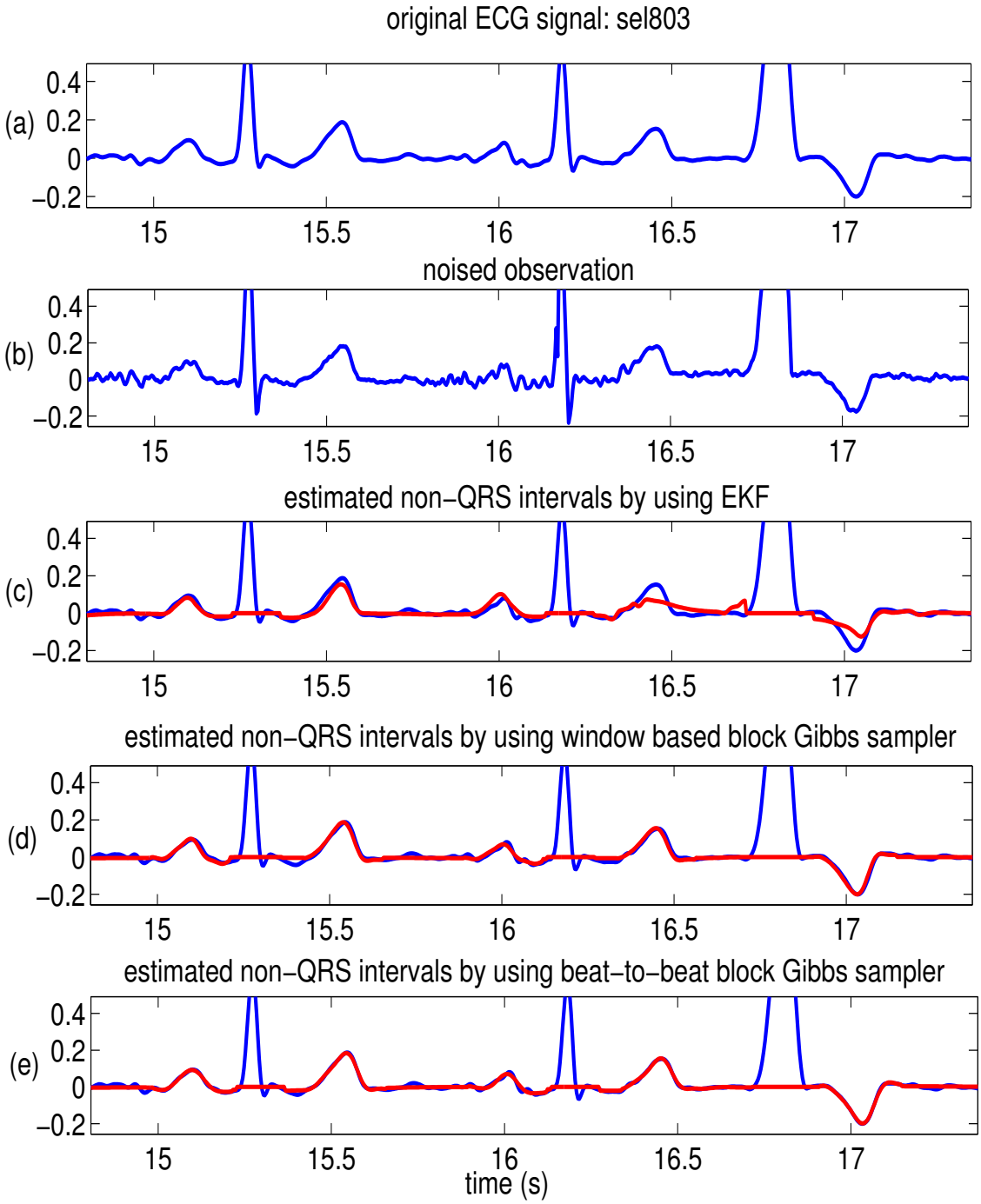


Figure 3.5: (a) Segment from QTDB sel803; (b) noisy version including MA noise with SNR = 10dB; (c) non-QRS signal component estimated by the EKF method of [SS09] (red) and noise-free original signal (blue); (d) non-QRS signal component estimated by the window-based method of Section 2.4 (red) and noise-free original signal (blue); (e) non-QRS signal component estimated by the proposed beat-to-beat method (red) and noise-free original signal (blue).

Quantitative Analysis

In this section, we provide a quantitative performance comparison of the beat-to-beat method with the previously proposed window based Bayesian model with a block Gibbs sampler and several alternative methods [LJC94, MAO+04, VGP+00], based on an exhaustive evaluation performed on the entire QTDB. For a quantitative analysis of the performance of T and P wave detection, we have computed the *sensitivity* (also referred to as *detection rate*) Se and the *positive predictivity* P^+ as in Section 2.3.5 and Section 2.4.5. The performance of wave delineation has been 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 250Hz. 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 have considered only those of the first cardiologist, who provided annotations for at least 30 beats per dataset.

Table 3.1 shows the results for Se, P^+, and $m \pm s$ obtained for the entire QTDB with the beat-to-beat block Gibbs sampler method, the multi-beat block Gibbs sampler method in Section 2.4.5 and three alternative methods of [LJC94, MAO+04, VGP+00]. It can be seen that the proposed method detects the T and P waves annotated by the cardiologist with high sensitivity: the sensitivity Se is 100% for the T waves and 99.93% for the P waves. Similarly good results were obtained for the positive predictivity P^+, which is 99.30% for the T waves and 99.10% for the P waves. Both the Se values and the P^+ values are clearly better than those obtained with the other methods, including the recently proposed window-based Bayesian method of Section 2.4.5.

Regarding the delineation performance, it is seen from Table 3.1 that the proposed method delineates the annotated T and P waves with mean errors m not exceeding 4ms (with one exception) and with smaller standard deviations s than the other methods (again with one exception). We note that delineation error tolerances have been recommended by the CSE Working Party [TCWP85]. In particular, the standard deviation s for $t_{\mathrm{P,on}}$, $t_{\mathrm{P,end}}$, and $t_{\mathrm{T,end}}$ should be at most $2s_{\mathrm{CSE}}$, which is listed in the last row of Table 3.1. However, a stricter recommendation proposed in [MAO+04] is $s \leq s_{\mathrm{CSE}}$. According to Table 3.1, the delineation results for $t_{\mathrm{P,end}}$ achieved by the proposed method comply with the loose recommendation. For the $t_{\mathrm{T,end}}$ results, the proposed method is the only one that complies with the strict recommendation.

Table 3.1: Comparison of the detection and delineation performance of the proposed method, the window based block Gibbs sampler of [LKT+11], the Wavelet transform based method of [MAO+04] (WT), the low-pass differentiation based method of [LJC94] (LPD), the action potential model of [VGP+00] and the delineation error tolerance of [TCWP85]. (N/A: not available)

Method	Parameters	$t_{\mathrm{P,on}}$	$t_{\mathrm{P,peak}}$	$t_{\mathrm{P,end}}$	$t_{\mathrm{T,on}}$	$t_{\mathrm{T,peak}}$	$t_{\mathrm{T,end}}$
Beat-to-beat block Gibbs sampler (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
Window based block Gibbs sampler [LKT+11]	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 [MAO+04]	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 [LJC94]	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 model [VGP+00]	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
Tolerance	$2s_{\mathrm{CSE}}$ (ms)	10.2	N/A	12.7	N/A	N/A	30.6

3.3 Sequential Monte Carlo methods for beat-to-beat P and T wave analysis

Section 3.2 studied a beat-to-beat Bayesian model that introduced 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, a sequential Monte Carlo (SMC) method [DdFG01] is studied to take into account all the information contained in the past of the current beat to be processed. First, we propose a dynamic model to solve simultaneously P and T wave delineation and waveform estimation in a beat-to-beat basis. The proposed dynamic model is similar to the Bayesian models introduced in Section 3.2. However, it adapts to the morphology variations across the ECG beats by assigning a random walk model to the waveform coefficients. Following the SMC principle, particle filters (PFs) are introduced to estimate the unknown parameters of the proposed model. The key idea is to represent the required posterior density by a set of random samples with associated weights and to compute parameter estimates from these samples and weights. Despite the simplicity of the PF principle, its main drawback is its computational complexity especially for large state dimension. This computational complexity can be reduced for nonlinear dynamic models containing a subset of parameters which are linear and Gaussian, conditional upon the other parameters. In this case, the linear parameters can be optimally estimated through standard linear Gaussian filtering. This technique is often referred to as Rao-Blackwellization [DGK01] or marginalization [SGN05]. In our case, the state equations are linear with respect to a subset of the unknown parameters. Thus we propose to use a marginalized particle filter (MPF) that dismisses the states appearing linearly in the dynamics, generates particles in the space of the remaining states and runs one KF for each of these particles to estimate the "linear" parameters. The proposed approach is evaluated on the annotated QT database [LMGM97] and compared with the window-based Bayesian method, the beat-to-beat Bayesian method based on block Gibbs sampler and with other state-of-art methods.

3.3.1 Dynamic model for non-QRS intervals

As in the previous sections, we assume that the locations of the non-QRS intervals are provided by a preliminary QRS detection step using the Pan-Tompkins algorithm [PT85]. We also assume that both the global and local baseline wanderings are removed by the technique suggested in [CM07].

Signal model

As shown in Fig. 3.6, the non-QRS interval $\mathcal{J}_n$ associated with the nth heartbeat consists of two complementary subintervals: a T search interval $\mathcal{J}_{\mathrm{T},n}$, which may contain a T wave, and a P search interval $\mathcal{J}_{\mathrm{P},n}$, which may contain a P wave. The lengths of the intervals $\mathcal{J}_n$, $\mathcal{J}_{\mathrm{T},n}$, and $\mathcal{J}_{\mathrm{P},n}$ will be denoted by N_n, $N_{\mathrm{T},n}$, and $N_{\mathrm{P},n}$, respectively. Note that $N_{\mathrm{T},n} + N_{\mathrm{P},n} = N_n$. The interval lengths $N_{\mathrm{T},n}$ and $N_{\mathrm{P},n}$ can be determined by a cardiologist or simply fixed as given percentages of N_n. In this chapter, we choose $N_{\mathrm{T},n} = N_{\mathrm{P},n} = N_n/2$. Our goal is to

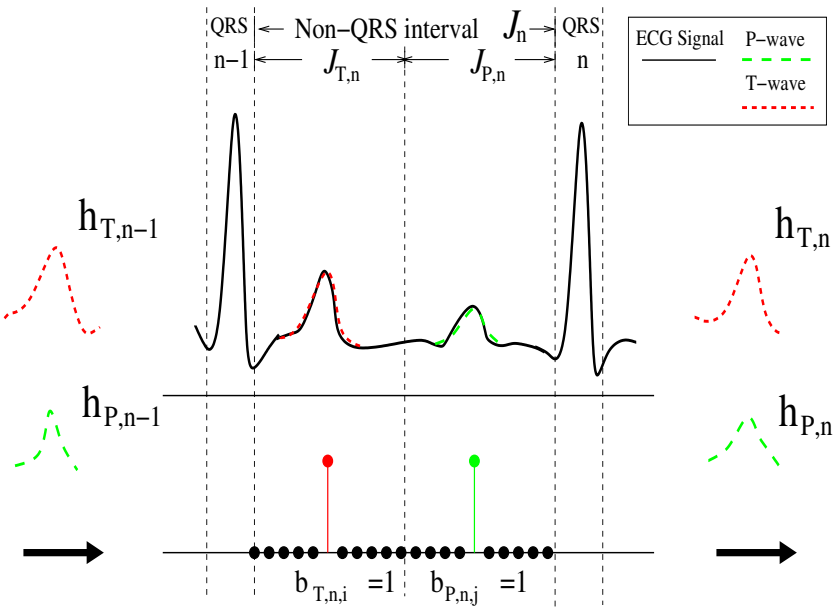


Figure 3.6: Signal model within a non-QRS interval.

estimate the locations and shapes (waveforms) of the T and P waves within their respective search intervals $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$.

Similar to the blind deconvolution problem introduced previously, the signal $x_{n,k}$ in the nth T wave interval $\mathcal{J}_{\mathrm{T},n}$ is modeled by the convolution of an unknown binary "indicator sequence" $\boldsymbol{b}_{\mathrm{T},n} = (b_{\mathrm{T},n,1} \cdots b_{\mathrm{T},n,N_{\mathrm{T},n}})^T$ indicating the wave locations ($b_{\mathrm{T},n,i} = 1$ if there is a wave at location i, $b_{\mathrm{T},n,i} = 0$ otherwise) with an unknown T waveform $\boldsymbol{h}_{\mathrm{T},n} = (h_{\mathrm{T},n,-L} \cdots h_{\mathrm{T},n,L})^T$ such that

$$x_{n,k} = \sum_{j=1}^{N_{\mathrm{T},n}} h_{\mathrm{T},n,k-j}\, b_{\mathrm{T},n,j} + w_{n,k}\,, \quad k \in \mathcal{J}_{\mathrm{T},n} \tag{3.25}$$

where $w_{n,k}$ is a white Gaussian noise with variance $\sigma^2_{w,n}$ and $h_{\mathrm{T},n,k} = 0$ for $k \notin \{-L, \ldots, L\}$. 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 T wave (e.g., $2L+1 = N_n/3$ in this work). The signal model ensures that the position of a nonzero detected indicator (such that $\hat{b}_{\mathrm{T},n,k} = 1$) directly indicates the center $k=0$ of the waveform support interval $\{-L, \ldots, L\}$. Note that since there is at most one T wave in each T wave interval, there will be at most one nonzero entry of $\boldsymbol{b}_{\mathrm{T},n}$ which corresponds to the T wave location.

Similarly, the P wave within the P wave interval is modeled by the convolution of $\boldsymbol{h}_{\mathrm{P},n} = (h_{\mathrm{P},n,-L} \cdots h_{\mathrm{P},n,L})^T$ with $\boldsymbol{b}_{\mathrm{P},n} = (b_{\mathrm{P},n,1} \cdots b_{\mathrm{P},n,N_{\mathrm{P},n}})^T$. The P interval signal component $x_{n,k}$ can then be written as

$$x_{n,k} = \sum_{j=1}^{N_{\mathrm{P},n}} h_{\mathrm{P},n,k-j}\, b_{\mathrm{P},n,j} + w_{n,k}\,, \quad k \in \mathcal{J}_{\mathrm{P},n} \tag{3.26}$$

with $h_{\mathrm{P},n,k} = 0$ for $k \notin \{-L, \ldots, L\}$.

Following the idea developed in Chapter 2 and in Section 3.2, we represent the T and P waveforms by a basis expansion using discrete-time versions of Hermite functions. Thus, the waveform vectors can be written as

$$\boldsymbol{h}_{\mathrm{T},n} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n}\,, \quad \boldsymbol{h}_{\mathrm{P},n} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P},n} \tag{3.27}$$

where $\boldsymbol{H}$ is a $(2L+1) \times G$ matrix whose columns are the first G Hermite functions (with $G \leq 2L+1$), suitably sampled and truncated to length $2L+1$, and $\boldsymbol{\alpha}_{\mathrm{T},n}$ and $\boldsymbol{\alpha}_{\mathrm{P},n}$ are unknown coefficient vectors of length G.

Using (3.27), we obtain the following vector representation of the T wave interval in (3.25)

$$\boldsymbol{x}_{\mathrm{T},n} = \boldsymbol{B}_{\mathrm{T},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n} + \boldsymbol{w}_n \tag{3.28}$$

where $\boldsymbol{x}_{\mathrm{T},n} = (x_{n,1} \cdots x_{n,N_{\mathrm{T},n}})^T$, $\boldsymbol{B}_{\mathrm{T},n}$ is the $(N_{\mathrm{T},n}) \times (2L+1)$ Toeplitz matrix with first row $(b_{n,L+1} \cdots b_{n,1}\ 0 \cdots 0)$ and first column $(b_{n,L+1} \cdots b_{n,N_{\mathrm{T},n}}\ 0 \cdots 0)^T$ and $\boldsymbol{w}_n = (w_{n,1} \cdots w_{n,N_{\mathrm{T},n}})^T$ is a Gaussian vector with covariance matrix $\sigma_w^2 \boldsymbol{I}_{N_{\mathrm{T},n} \times N_{\mathrm{T},n}}$. Note that similar vector representation can be obtained for the signal vector of the P wave interval $\boldsymbol{x}_{\mathrm{P},n} = (x_{n,N_{\mathrm{T},n}+1} \cdots x_{n,N_n})^T$.

Prior distributions and dynamic model

This section describes the prior distributions and the dynamic model assigned to the unknown parameters. The proposed algorithm successively processes the T and P wave intervals within the same non-QRS component. Note that only the T wave dynamic model is presented in the following and the subscript indicating T wave is omitted for notation convenience.

Due to the parametrization (3.28), the state parameter vector for the nth T wave interval (time step n) is given by

$$\boldsymbol{\theta}_n = \left(\boldsymbol{b}_n^T,\ \boldsymbol{\alpha}_n^T\right)^T. \tag{3.29}$$

Concerning the indicator vector $\boldsymbol{b}_n$, since there is no known relation between the wave locations of each beat, the parameter vectors $\boldsymbol{b}_n, \boldsymbol{b}_{n-1}, \boldsymbol{b}_{n-2}, \ldots$ are modeled as a priori mutually independent. The indicators $b_{n,k}$ contained in $\boldsymbol{b}_n$ are subject to a *block constraint*: within $\mathcal{J}_{\mathrm{T},n}$, there is one T wave (such that $\|\boldsymbol{b}_n\| = 1$) or none (with $\|\boldsymbol{b}_n\| = 0$). Therefore, we define the prior of $\boldsymbol{b}_n$ as a discrete distribution on the set $\left\{\boldsymbol{\beta}_0, \ldots, \boldsymbol{\beta}_j, \ldots, \boldsymbol{\beta}_{N_{\mathrm{T},n}}\right\}$ where $\boldsymbol{\beta}_j$ is a $N_{\mathrm{T},n} \times 1$ vector whose j-th entry is 1 and all remaining entries are zero ($\boldsymbol{\beta}_0$ is an all zero vector which represents the case where there is no T wave). The prior of $\boldsymbol{b}_n$ can then be defined as

$$\Pr\left(\boldsymbol{b}_n = \boldsymbol{\beta}_j\right) = \frac{1}{N_{\mathrm{T},n}+1}\,, \quad j \in \{1, \ldots, N_{\mathrm{T},n}\}\,. \tag{3.30}$$

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 $\boldsymbol{\alpha}$

$$\boldsymbol{\alpha}_n = \boldsymbol{\alpha}_{n-1} + \boldsymbol{v}_{n-1} \tag{3.31}$$

where $\boldsymbol{\alpha}_{n-1}$ denotes the T waveform estimates of the $(n-1)$th beat and $\boldsymbol{v}_{n-1}$ is an additive Gaussian white noise vector denoted as $\boldsymbol{v}_{n-1} \sim \mathcal{N}\left(\mathbf{0}, \sigma_\alpha^2 \boldsymbol{I}_G\right)$, where $\boldsymbol{I}_G$ is the $G \times G$ identity matrix. The variance σ_α^2 determines how fast the waveform coefficients are expected to change with time. Since the non-QRS components are normalized by using the corresponding R peak values, we propose a non-informative prior of $\sigma_\alpha^2 = 0.1$ to handle all possible waveform variations in this work. Note that the value of σ_α^2 can be further adjusted either by an expert or by calculating the ECG waveform variance of an example ECG portion in an off-line parameter selection procedure [SSJC07].

The T wave delineation and waveform estimation problem consists of estimating recursively the wave location $\boldsymbol{b}_n$ and the waveform coefficients $\boldsymbol{\alpha}_n$ from the measurements $\boldsymbol{x}_n$ defined in (3.28).

3.3.2 Particle filters for beat-to-beat wave analysis

Our goal is to estimate jointly the discrete-valued indicator vector $\boldsymbol{b}_n$ and the waveform vectors $\boldsymbol{\alpha}_n$, i.e., 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 $p\left(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n}\right)$ with $\boldsymbol{\theta}_{0:n} = (\boldsymbol{\theta}_0, \ldots, \boldsymbol{\theta}_n)$.

Particle filters (PFs) are a class of sequential Monte Carlo (SMC) methods well-suited to perform the estimation of the hybrid state vector $\boldsymbol{\theta}_{0:n}$. The key idea is to represent the required posterior density function by a set of random samples (particles) with associated weights and to compute estimates based on these samples and weights. Let $\left\{\boldsymbol{\theta}_{0:n}^i, i = 0, \ldots, N_s\right\}$ denote a set of support points (referred to as particles) with associated weights $\{\mu_n^i, i = 0, \ldots, N_s\}$ and $\boldsymbol{\theta}_{0:n}$ is the set of all states up to time n (referred to as n-th beat in our case). Then, the target posterior distribution at time n can be approximated as

$$\hat{p}\left(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n}\right) = \sum_{i=1}^{N_s} \mu_n^i \delta\left(\boldsymbol{\theta}_{0:n} - \boldsymbol{\theta}_{0:n}^i\right), \quad \sum_{i=1}^{N_s} \mu_n^i = 1 \tag{3.32}$$

where δ is the Dirac delta function. The weights μ_n^i and the particles $\boldsymbol{\theta}_{0:n}^i$ are classically obtained by applying sequentially the importance sampling (IS) technique. Ideally, the particles should be sampled directly from the target distribution $p\left(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n}\right)$, and assigned equal weights. Since it is usually impossible to sample $p\left(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n}\right)$, they are drawn instead from a proposal distribution $q(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n})$, called importance density [AMG02]. Then, the weights are used to correct the discrepancy between p and q

$$\mu_n^i \propto \frac{p(\boldsymbol{\theta}_{0:n}^i|\boldsymbol{x}_{1:n})}{q(\boldsymbol{\theta}_{0:n}^i|\boldsymbol{x}_{1:n})} \tag{3.33}$$

where "$\propto$" means "proportional to".

In the sequential case, one could leave the previous particles $\boldsymbol{\theta}_{0:n-1}^i$ unchanged by only simulating at time step n

$$\boldsymbol{\theta}_n^i \sim q(\boldsymbol{\theta}_n|\boldsymbol{\theta}_{0:n-1}^i, \boldsymbol{x}_{1:n}). \tag{3.34}$$

Then the importance weights can be updated using

$$\mu_n^i \propto \mu_{n-1}^i \frac{p(\boldsymbol{x}_n|\boldsymbol{\theta}_n^i)p(\boldsymbol{\theta}_n^i|\boldsymbol{\theta}_{0:n-1}^i)}{q(\boldsymbol{\theta}_n^i|\boldsymbol{\theta}_{0:n-1}^i, \boldsymbol{x}_{1:n})}. \tag{3.35}$$

A common problem with the IS technique is the degeneracy phenomenon, where after a few iterations, all but one particle have negligible normalized weights. To overcome this inherent limitation, a selection step is introduced which consists of resampling the set of particles according to the estimated empirical distribution [DdFG01].

A standard particle filter

We propose in a first step a standard particle filter (SPF) which generates jointly the particles for both the indicator vector $\boldsymbol{b}_n$ and the waveform vector $\boldsymbol{\alpha}_n$. The PF recursions are summarized in Algorithm 7. The different steps involved in this algorithm are detailed in the rest of this section and their derivations are provided in the Appendix D. Note that N_s is the number of particles.

Proposal distribution for the indicators. It is well-known that the choice of the importance distribution is a critical issue to design efficient PF algorithms. To generate samples in interesting 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 optimal importance distribution in the sense that it minimizes the variance of the importance weights is $q(\boldsymbol{\theta}_n|\boldsymbol{\theta}_{0:n-1}^i, \boldsymbol{x}_{0:n}) = p(\boldsymbol{\theta}_n|\boldsymbol{\theta}_{0:n-1}^i, \boldsymbol{x}_n)$ [DGA00].

Since the wave indicator vector $\boldsymbol{b}_n$ is independent from $\boldsymbol{b}_{0:n-1}$, the optimal proposal distribution for $\boldsymbol{b}_n$ can be written as

$$p(\boldsymbol{b}_n = \boldsymbol{\beta}_j|\boldsymbol{\alpha}_{n-1}^i, \boldsymbol{x}_n) \propto \exp\left(\boldsymbol{\mu}_{1,j}\boldsymbol{\Sigma}_{1,j}^{-1}\boldsymbol{\mu}_{1,j}\right) \tag{3.36}$$

with

$$\boldsymbol{\mu}_{1,j} = \boldsymbol{\Sigma}_{1,j}\left(\frac{\boldsymbol{H}^T\boldsymbol{B}_{n,j}^T\boldsymbol{x}_n}{\sigma_w^2} + \frac{\boldsymbol{\alpha}_{n-1}^i}{\sigma_\alpha^2}\right),$$
$$\boldsymbol{\Sigma}_{1,j} = \left(\frac{\boldsymbol{H}^T\boldsymbol{B}_{n,j}^T\boldsymbol{B}_{n,j}\boldsymbol{H}}{\sigma_w^2} + \frac{\mathbf{I}_G}{\sigma_\alpha^2}\right)^{-1}.$$

where $\boldsymbol{B}_{n,j}$ is the $N_{\mathrm{T},n} \times (2L+1)$ Toeplitz matrix with first row $(b_{n,L+1} \cdots b_{n,1}\ 0 \cdots 0)$ and first column $(b_{n,L+1} \cdots b_{n,N_{\mathrm{T},n}}\ \ 0 \cdots 0)^T$ which corresponds to $\boldsymbol{b}_n = \boldsymbol{\beta}_j$.

Algorithm 7 A standard particle filter (SPF)

{Initialization}
for particles $i = 1, \ldots, N_s$ **do**
  Set $\boldsymbol{b}_0^i = \mathbf{0}_{N_n \times 1}$, $\boldsymbol{\alpha}_0^i = \boldsymbol{H}^{-1}\hat{\boldsymbol{h}}_0$, $P_0^i = \mathbf{0}_{G\times G}$, $\mu_0^i = 1$.
end for
{Iterations}
for $n = 1, 2, \ldots,$ **do**
  for particles $i = 1, \ldots, N_s$ **do**
    {Particles generation}
    Sample $\boldsymbol{b}_n^i \sim \Pr(\boldsymbol{b}_n = \boldsymbol{\beta}_j | \boldsymbol{\alpha}_{n-1}^i, \boldsymbol{x}_n)$ (see (3.36))
    Sample $\boldsymbol{h}_n^i \sim p(\boldsymbol{\alpha}_n | \boldsymbol{b}_n^i, \boldsymbol{\alpha}_{n-1}^i, \boldsymbol{x}_n)$ (see (3.37))
    Evaluate weights
      $\mu_n^i = \mu_{n-1}^i \sum_{j \in \mathcal{J}_{\mathrm{T},n}} p(\boldsymbol{x}_n | \boldsymbol{b}_n^i = \boldsymbol{\beta}_j, \boldsymbol{x}_{1:n-1}) p(\boldsymbol{b}_n^i = \boldsymbol{\beta}_j)$
  end for
  {Weight normalization}
  for particles $i = 1, \ldots, N_s$ **do**
    $\mu_n^i = \mu_n^i / \sum_{i=1}^{N_s} \mu_n^i$
  end for
  { State estimation}
  Estimation of $\boldsymbol{b}_n$ (see (3.38)) and $\boldsymbol{\alpha}_n$ (see (3.39))
  {Particle resampling}
  Calculate $\hat{N}_{\text{eff}} = 1 / \sum_{i=1}^{N_s} \left(\mu_n^i\right)^2$
  if $\hat{N}_{\text{eff}} \leq 0.7 N_s$ **then**
    Resample using systematic sampling scheme [DdFG01]
  end if
end for

Proposal distribution for the waveform. Concerning the waveform coefficients, the proposal distribution can be written as

$$p(\boldsymbol{\alpha}_n | \boldsymbol{b}_n^i, \boldsymbol{\alpha}_{n-1}^i, \boldsymbol{x}_n) = \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \tag{3.37}$$

with

$$\boldsymbol{\mu}_2 = \boldsymbol{\Sigma}_2 \left(\frac{\boldsymbol{H}^T \left(\boldsymbol{B}_n^i\right)^T \boldsymbol{x}_n}{\sigma_w^2} + \frac{\boldsymbol{\alpha}_{n-1}^i}{\sigma_\alpha^2} \right)$$

$$\boldsymbol{\Sigma}_2 = \left(\frac{\boldsymbol{H}^T \left(\boldsymbol{B}_n^i\right)^T \boldsymbol{B}_n^i \boldsymbol{H}}{\sigma_w^2} + \frac{\boldsymbol{I}_G}{\sigma_\alpha^2} \right)^{-1}$$

State estimation. The sample-based blockwise MAP detector is used for estimating the binary sequence $\boldsymbol{b}_n$, while for the waveform coefficients $\boldsymbol{\alpha}_n$, a smooth state estimation (an

approximation of the MMSE estimator) is applied

$$\hat{\boldsymbol{b}}_n = \underset{\boldsymbol{b}_n^i \in \{0,1\}^{N_{\mathrm{T},n}}}{\operatorname{argmax}} \; \hat{p}(\boldsymbol{b}_n^i | \boldsymbol{x}_{1:n}, \boldsymbol{\alpha}_{0:n-1}) \tag{3.38}$$

$$\hat{\boldsymbol{\alpha}}_n = \sum_{i=1}^{N_s} \boldsymbol{\alpha}_n^i \mu_n^i. \tag{3.39}$$

where $\hat{p}(\boldsymbol{b}_n|\boldsymbol{x}_{1:n}, \boldsymbol{\alpha}_{0:n-1})$ denotes the approximated posterior distribution of the binary sequence $\boldsymbol{b}_n$. The wave delineation consists of determining the peak and boundaries of the detected T and P waves. As mentioned previously, the wave indicator estimated by the PF directly indicates the middle of the allocated waveform time window. Thus, the peak of the respective T or P wave can be obtained by shifting the indicator to the maximum position of the estimated waveform. Concerning the wave boundaries, since the estimated waveforms carry information about the wave morphology, they can be located by using the delineation criterion based on the waveform estimate as in Section 3.2.4.

A marginalized particle filter (MPF)

While the classical PF introduced before is fairly easy to implement, a main drawback is that the required number of particles increases quickly with the state dimension. The MPF can reduce the number of parameters estimated by the particle filtering and therefore the number of particles can be reduced. More specifically, when there is a linear Gaussian sub-structure in the state parameters $\boldsymbol{\theta}_n$, state estimates can be obtained by exploiting this structure. The key idea is to split $\boldsymbol{\theta}_n$ as follows

$$\boldsymbol{\theta}_n = \left[(\boldsymbol{\theta}_n^{\mathrm{L}})^T, \; (\boldsymbol{\theta}_n^{\mathrm{NL}})^T \right]^T \tag{3.40}$$

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. Using Bayes' theorem we can then marginalize out the linear and Gaussian state parameters and estimate them using the Kalman filter, which is the optimal filter for this case. The nonlinear state variables are then estimated using a PF. Note that one KF is then associated with each particle.

It can be observed from (3.28) that both the discrete parameter vector $\boldsymbol{b}_n$ and the continuous parameter vector $\boldsymbol{\alpha}_n$ are linear sub-structures with respect to the observation $\boldsymbol{x}_n$. Since the continuous parameters are more suitable for the Kalman filtering, we choose $\boldsymbol{\theta}_n^{\mathrm{L}} = \boldsymbol{\alpha}_n$ and $\boldsymbol{\theta}_n^{\mathrm{NL}} = \boldsymbol{b}_n$. Thus $\boldsymbol{\alpha}_n$ can be handled by the KF. Analytically marginalizing out the linear state variables from $p(\boldsymbol{\theta}_{0:n}|\boldsymbol{x}_{1:n})$ and using Bayes' theorem yields

$$p(\boldsymbol{b}_{0:n}, \boldsymbol{\alpha}_n | \boldsymbol{x}_{1:n}) = \underbrace{p(\boldsymbol{\alpha}_n | \boldsymbol{b}_{0:n}, \boldsymbol{x}_{1:n})}_{\text{Optimal KF}} \underbrace{p(\boldsymbol{b}_{0:n} | \boldsymbol{x}_{1:n})}_{\text{PF}} \tag{3.41}$$

The proposed marginalized particle filter (MPF) is summarized in Algorithm 8. The different steps involved in this algorithm are detailed in the rest of this section.

Algorithm 8 A marginalized particle filter (MPF)

{Initialization}
for particles $i = 1, \ldots, N_s$ **do**
  Set $\boldsymbol{b}_0^i = \mathbf{0}_{N_n \times 1}$, $\boldsymbol{\alpha}_0^i = \boldsymbol{H}^{-1}\hat{\boldsymbol{h}}_0$, $P_0^i = \mathbf{0}_{G\times G}$, $\mu_0^i = 1$.
end for
{Iterations}
for $n = 1, 2, \ldots,$ **do**
  for particles $i = 1, \ldots, N_s$ **do**
    {KF and PF propagation}
    KF prediction for $\boldsymbol{\alpha}_n^i$ (see (3.42))
    Sample $\boldsymbol{b}_n^i \sim \Pr(\boldsymbol{b}_n = \boldsymbol{\beta}_j | \boldsymbol{b}_{0:n-1}^i, \boldsymbol{x}_{1:n})$ (see (3.43))
    KF correction for $\boldsymbol{\alpha}_n^i$ (see (3.44))
    Evaluate weights
      $\mu_n^i = \mu_{n-1}^i \sum_{j \in \mathcal{J}_{\mathrm{T},n}} p(\boldsymbol{x}_n | \boldsymbol{b}_n^i = \boldsymbol{\beta}_j, \boldsymbol{x}_{1:n-1}) p(\boldsymbol{b}_n^i = \boldsymbol{\beta}_j)$
  end for
  {Weight normalization}
  for particles $i = 1, \ldots, N_s$ **do**
    $\mu_n^i = \mu_n^i / \sum_{i=1}^{N_s} \mu_n^i$
  end for
  { State estimation}
  Estimation of $\boldsymbol{b}_n$ (see (3.38)) and $\boldsymbol{\alpha}_n$ (see (3.39))
  {Particle resampling}
  Calculate $\hat{N}_{\mathrm{eff}} = 1 / \sum_{i=1}^{N_s} \left(\mu_n^i\right)^2$
  if $\hat{N}_{\mathrm{eff}} \leq 0.7 N_s$ **then**
    Resample using systematic sampling scheme
  end if
end for

KF prediction. In KF recursions, the mean and the covariance matrix of the state are propagated [AM79]. By using (3.31), the prediction step in the KF can be written as follows

$$\boldsymbol{\alpha}_{n|n-1}^i = \boldsymbol{\alpha}_{n-1}^i, \quad P_{n|n-1}^i = P_{n-1}^i + Q \tag{3.42}$$

where $\boldsymbol{\alpha}_{n|n-1}^i = \mathrm{E}\left[\boldsymbol{\alpha}_n | \boldsymbol{x}_{1:n-1}, \boldsymbol{b}_{0:n-1}^i\right]$, $Q = \sigma_\alpha^2 \boldsymbol{I}_{2L+1}$ and $P_{n|n-1}^i = \mathrm{Cov}\left[\boldsymbol{\alpha}_n | \boldsymbol{x}_{1:n-1}, \boldsymbol{b}_{0:n-1}^i\right]$. Note that the predicted state vector and its covariance computed by the KF are directly used to propagate the particles and compute their importance weights.

Proposal distribution for the indicators. Similar to the proposal distribution of the SPF, the optimal importance distribution in the sense that it minimizes the variance of the importance weights is $q(\boldsymbol{\theta}_n | \boldsymbol{\theta}_{0:n-1}^i, \boldsymbol{x}_{0:n}) = p(\boldsymbol{\theta}_n | \boldsymbol{\theta}_{0:n-1}^i, \boldsymbol{x}_n)$. By inserting the KF prediction of (3.42), the proposal distribution for $\boldsymbol{b}_{\mathrm{T},n}$ can be written as

$$\Pr(\boldsymbol{b}_n = \boldsymbol{\beta}_j | \boldsymbol{b}_{0:n-1}^i, \boldsymbol{x}_{1:n}) \propto \mathcal{N}\left(\tilde{\boldsymbol{x}}_{n,j}^i, \tilde{S}_{n,j}^i\right) \tag{3.43}$$

with

$$\tilde{\boldsymbol{x}}^i_{n,j} = \boldsymbol{B}_{n,j}\boldsymbol{H}\boldsymbol{\alpha}^i_{n|n-1}$$
$$\tilde{S}^i_{n,j} = \boldsymbol{B}_{n,j}P^i_{n|n-1}\boldsymbol{B}^T_{n,j} + R$$

where $\boldsymbol{B}_{n,j}$ is the $N_{\mathrm{T},n} \times (2L+1)$ Toeplitz matrix with first row $(b_{n,L+1} \cdots b_{n,1}\ 0 \cdots 0)$ and first column $(b_{n,L+1} \cdots b_{n,N_{\mathrm{T},n}}\ \ 0 \cdots 0)^T$ which corresponds to $\boldsymbol{b}_n = \boldsymbol{\beta}_j$, and $R = \sigma_w^2 \boldsymbol{I}_{N_{\mathrm{T},n}}$. Note that with the marginalization, the proposal distribution for the indicators no longer depends on the impulse response $\boldsymbol{b}_{n-1}$ which has been marginalized out while it depends on the past sequence $\boldsymbol{b}_{1:n-1}$ whereas this is not the case with the standard particle filter.

KF correction. After receiving the observation at time step n, the waveform coefficients can be updated for each generated wave indicator particle $\boldsymbol{b}^i_n$. The KF correction procedure can be written as

$$\begin{aligned} S^i_n &= \boldsymbol{B}^i_n P^i_{n|n-1} \left(\boldsymbol{B}^i_n\right)^T + R \\ K &= P^i_{n|n-1} \left(\boldsymbol{B}^i_n\right)^T \left(S^i_n\right)^{-1} \\ \boldsymbol{\alpha}^i_n &= \boldsymbol{\alpha}^i_{n|n-1} + \boldsymbol{H}^{-1} K(\boldsymbol{x}_n - \boldsymbol{B}^i_n \boldsymbol{H}\boldsymbol{\alpha}^i_{n|n-1}) \\ P^i_n &= (\boldsymbol{I} - K\boldsymbol{B}^i_n) P^i_{n|n-1} \end{aligned} \tag{3.44}$$

where $\boldsymbol{B}^i_n$ is the Toeplitz matrix $(N_{\mathrm{T},n} + L) \times (L + 1)$ corresponding to the wave indicator particle vector $\boldsymbol{b}^i_n$.

3.3.3 Boundary issue between P and T wave intervals

Different from the Gibbs sampler based method in Section 3.2 which considers the whole non-QRS signal component in the processing, the PF based method processes the T and P wave intervals within the same non-QRS component sequentially one after the other. This is because for the Gibbs sampler, one can generate samples for the T wave parameters conditioned on the rest of the parameters including those of the P wave interval in the same iteration, while this is not feasible for the PF. Since the two intervals can not be considered simultaneously, the border situation (where a T or P wave has part of its waveform across the border between $\mathcal{J}_{\mathrm{T},n}$ and $\mathcal{J}_{\mathrm{P},n}$) becomes an issue.

In the signal models (3.25) and (3.26), the non zero entry of the wave indicator vector can be anywhere within the wave search interval. Meanwhile, we adopted a "zero boundary" assumption, which assumes that for a T wave search interval, the observation sequence x_k is defined for all $\{\mathcal{J}_{\mathrm{T},n} - L, \cdots, \mathcal{J}_{\mathrm{T},n} + L\}$ (where L is the delay introduced by the convolution model), while it is not defined for $k \notin \mathcal{J}_{\mathrm{T},n}$. However, this assumption is not suitable for signals that have part of the waveform crossing the border of the P and T wave intervals. Fig. 3.7 shows a typical example with the premature ventricular contraction (PVC) signal, where the T wave is positioned almost at the middle of the non-QRS interval and the P wave is missing. In this case, if we only consider the observation within $k \in \mathcal{J}_{\mathrm{T},n}$, the T waveform may not be

estimated correctly.

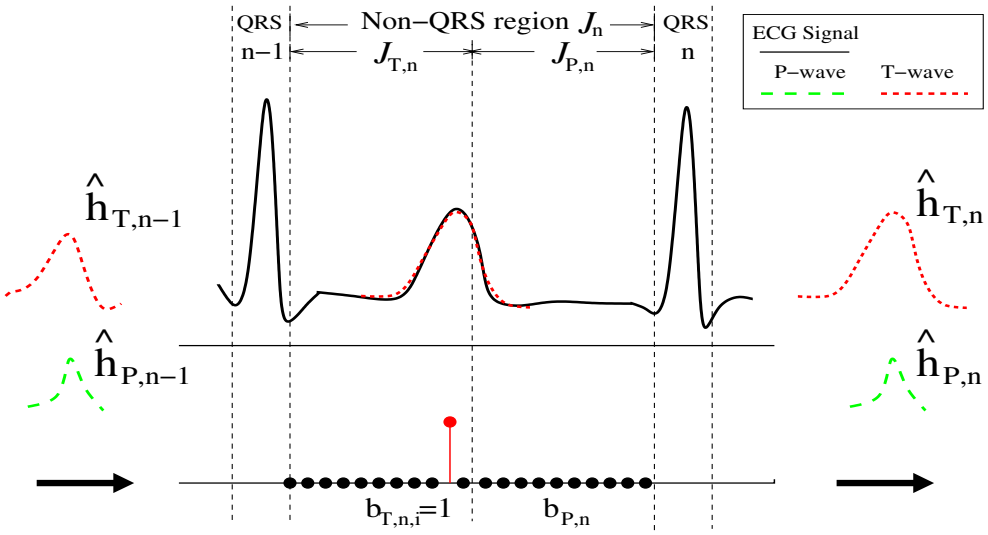


Figure 3.7: An example of the boundary problem with PVC signal.

In order to overcome this problem, we propose to include in the T wave interval observations $\boldsymbol{x}_{\mathrm{T}}$ the first L samples of the following P wave interval. As shown in Fig. 3.8, instead of adding L zeros at the beginning and the end of the T wave search interval $\mathcal{J}_{\mathrm{T},n}$, the first samples of $\mathcal{J}_{\mathrm{P},n}$ are considered to complete the end of $\boldsymbol{x}_{\mathrm{T}}$. Since one non-QRS interval is processed sequentially from the T wave interval to the P wave interval, the estimated signal within the overlapped part can be extracted. More precisely, the first L samples of the P wave interval observations $\boldsymbol{x}_{\mathrm{P}}$, which are overlapped by $\boldsymbol{x}_{\mathrm{T}}$, are obtained by subtracting $\hat{\boldsymbol{x}}_{\mathrm{T}}$ from the observation of the nth non-QRS component $\boldsymbol{x}$.

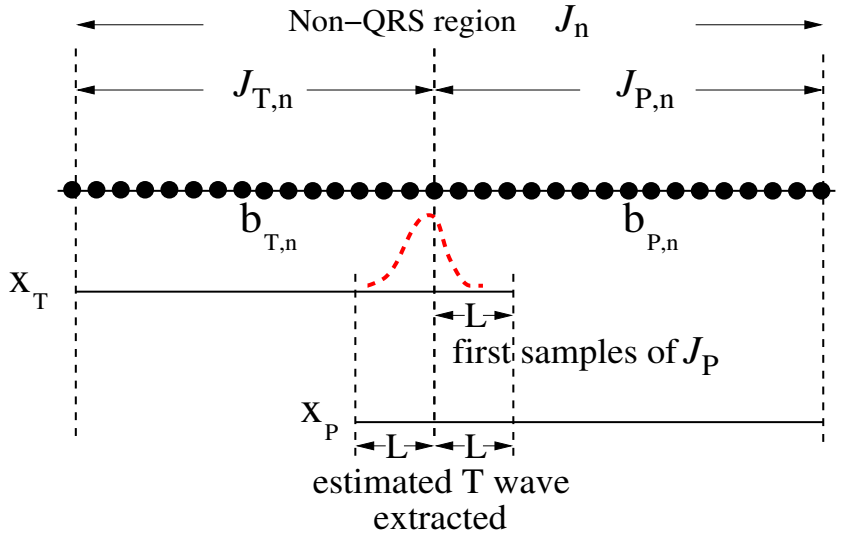


Figure 3.8: A new boundary definition.

3.3.4 Simulation results

Simulations have been conducted to validate the proposed PF algorithms. To assess the algorithm performance of the two PFs, Fig. 3.9 shows the empirical normalized mean square error (NMSE) of the estimated non-QRS components $\hat{\boldsymbol{x}}$ versus the number of particles N_s for SPF and MPF. The empirical NMSE is defined as the average (over the 100 realizations) of $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2$ normalized by the average of $\|\boldsymbol{x}\|^2$. The fixed hyperparameters involved in the prior distributions were chosen as $\sigma_\alpha^2 = 0.1$ and $\sigma_w^2 = 0.1$. These values allow for an appropriate waveform variability from one beat to another. Note that the non-QRS components are normalized by using the corresponding R peak values to handle different amplitude resolutions. The waveform vector $\hat{\boldsymbol{h}}_0$ was initialized with a $2L+1$ Hanning window whose amplitude was half the R peak. As can be seen, benefiting from the optimal importance distribution, good estimation performance is obtained with a moderate number of particles for both SPF and MPF. Furthermore, it is shown that the MPF requires less particles than SPF to achieve similar performance. We have chosen the MPF method with $N_\mathrm{s} = 200$ particles for all the following simulations.

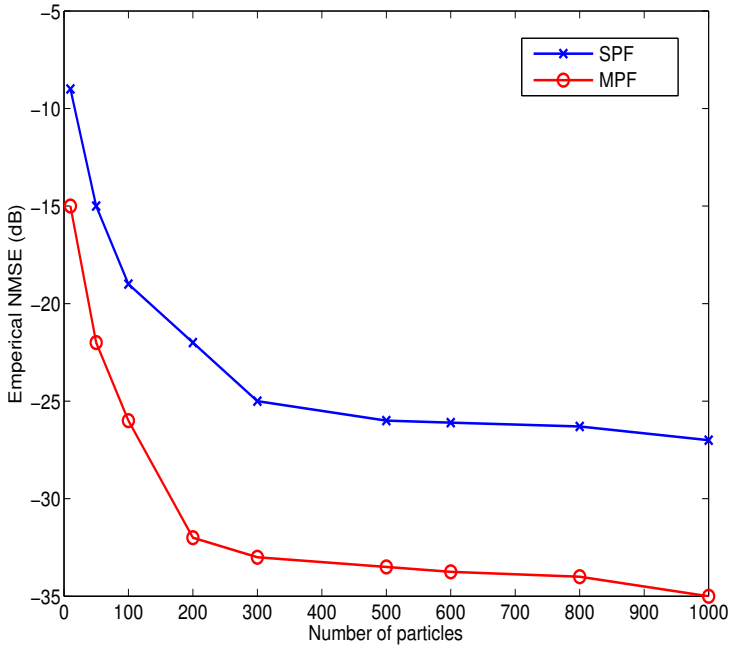


Figure 3.9: Estimation performance versus the number of particles.

Fig. 3.10 shows qualitative comparisons of the MPF with the previously introduced beat-to-beat block Gibbs sampler method. Here, muscular activity (MA) noise has been added to the clean ECG signal as in Figs. 3.4 and 3.5. Fig. 3.10(a) shows a segment of QTDB sele0136. Fig. 3.10(b) shows the same segment corrupted by MA noise with a signal-to-noise ratio (SNR) of 10dB. The estimated non-QRS signal components obtained from the noisy signal by the two methods are depicted in Fig. 3.10(c) and Fig. 3.10(d). The original (noise-free) ECG signal is

also shown for comparison. It can be seen that the MPF method and the beat-to-beat block Gibbs sampler method provide similar results. This confirms that the block Gibbs sampler method performs similarly to the sequential MC method which is the optimal solution for the ECG beat-to-beat dynamic model. Fig. 3.11 shows analogous results for a segment of sel803 that contains PVCs. As explained in Section 3.3.3, the boundary issue caused by the sequential processing scheme can be handled by including in the T wave interval observations $\boldsymbol{x}_{\mathrm{T}}$ the first L samples of the following P wave interval, where L is the half of the allocated waveform length. This allows us to estimate the T waveform even if the wave indicator is located at the last sample of $\mathcal{J}_{\mathrm{T}}$. Since one non-QRS interval is processed sequentially from a T wave interval to a P wave interval, the estimated T wave portion within the overlapped part can be extracted. Again, it can be seen that MPF method and the beat-to-beat block Gibbs sampler method provide similar estimates.

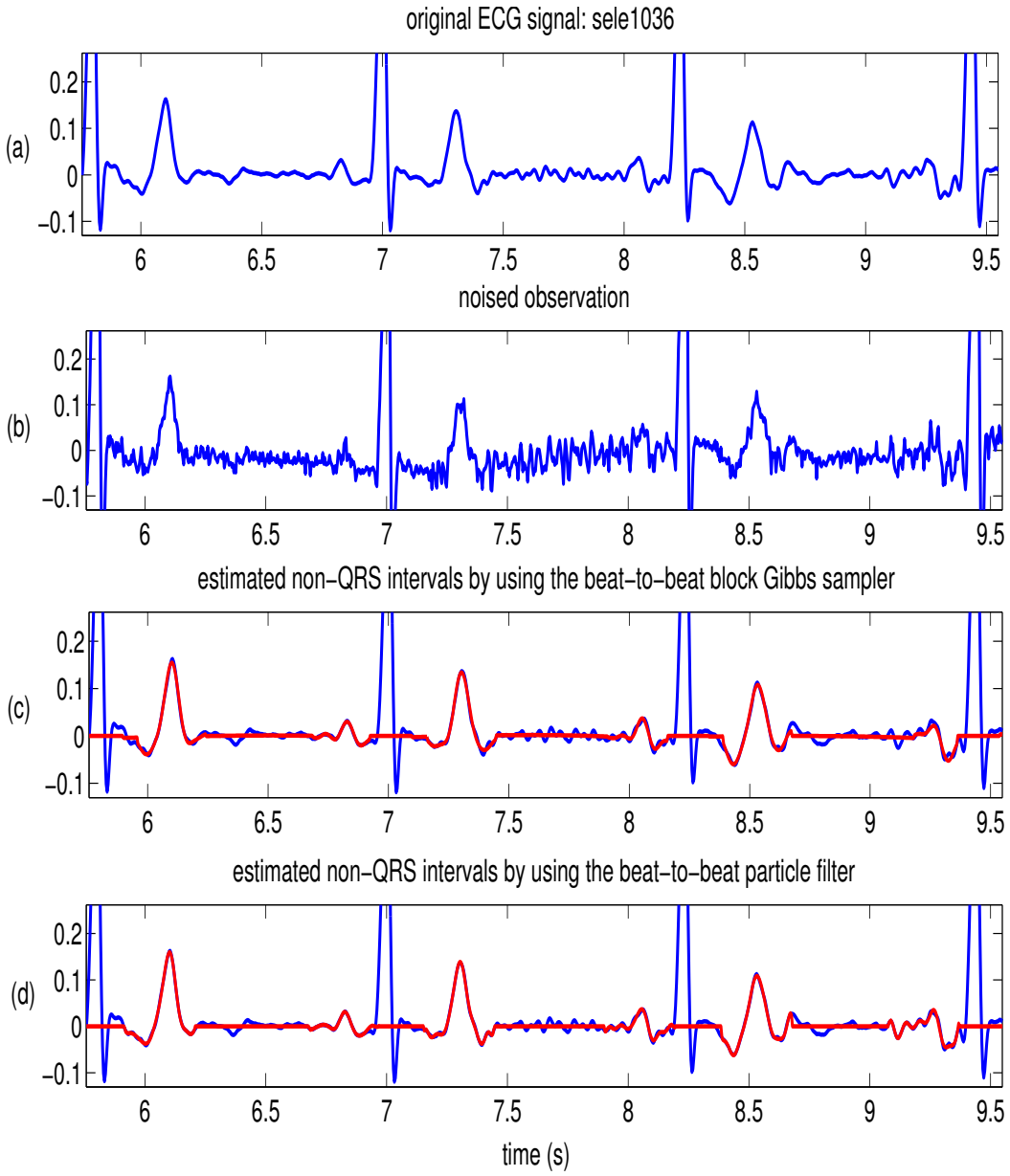


Figure 3.10: (a) Segment from QTDB sele0136; (b) noisy version including MA noise with SNR = 10dB; (c) non-QRS signal component estimated by the beat-to-beat block Gibbs sampler method (red) and noise-free original signal (blue); (d) non-QRS signal component estimated by the proposed beat-to-beat MPF method (red) and noise-free original signal (blue).

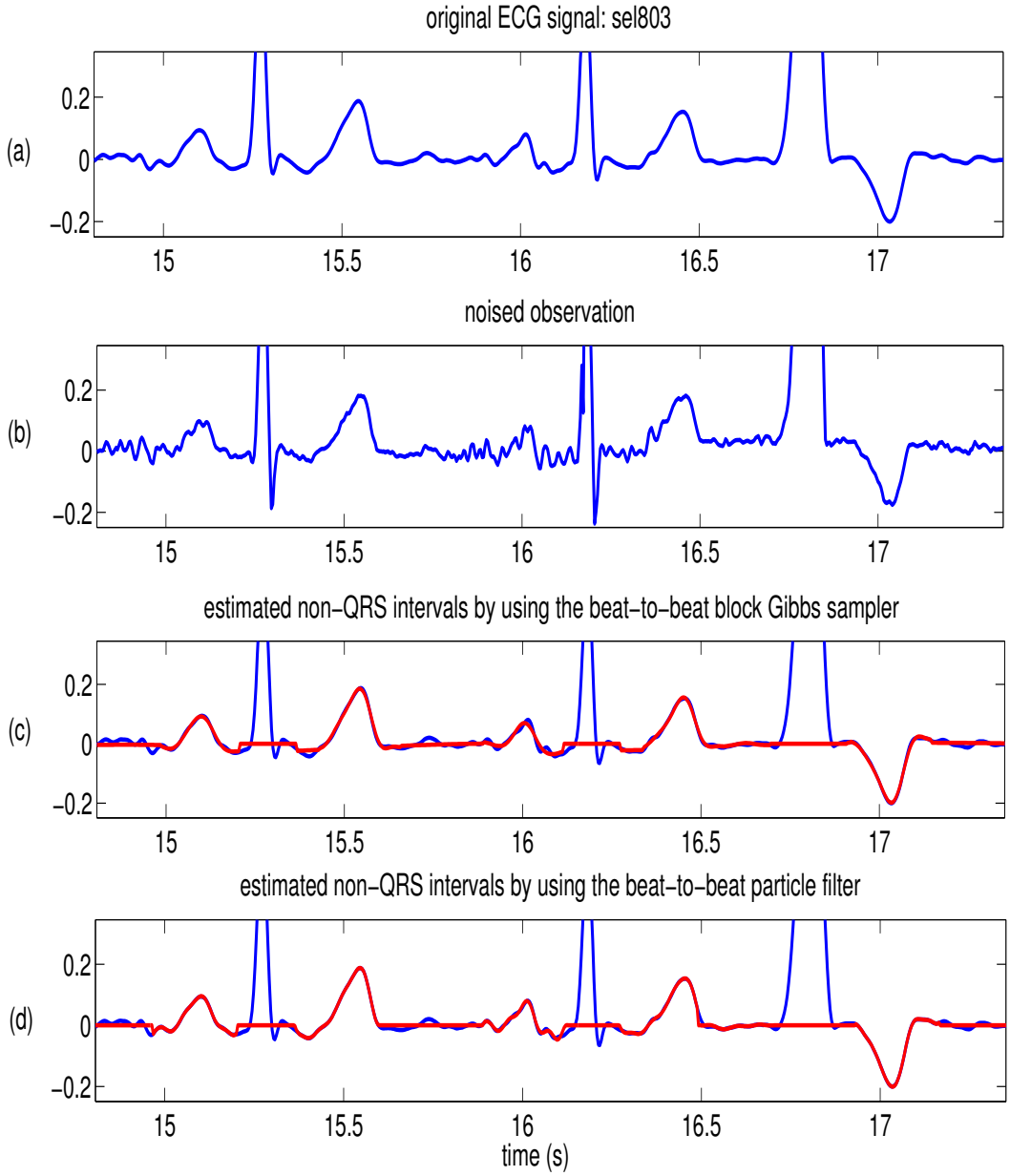


Figure 3.11: (a) Segment from QTDB sel803; (b) noisy version including MA noise with SNR = 10dB; (c) non-QRS signal component estimated by the beat-to-beat block Gibbs sampler method (red) and noise-free original signal (blue); (d) non-QRS signal component estimated by the proposed beat-to-beat MPF method (red) and noise-free original signal (blue).

For a quantitative comparison between the block Gibbs sampler and the MPF beat-to-beat methods, Table 3.2 presents the means (m) and standard deviations (s) of the differences between the automated delineation results and the manual annotations. It is seen that the detection and delineation results obtained with the two methods are comparable. However, the MPF method slightly outperforms the block Gibbs sampler method in terms of delineation accuracy. This can be explained by the fact that the MPF considers all the available beats while the beat-to-beat block Gibbs sampler considers only the estimates of the last beat. These two methods have comparable computational complexity, with a slightly higher cost for the MPF method. Note also that despite the moderate number of particles needed by the MPF, the fact that one KF is associated with each particle introduces matrix inversion operations which make the MPF more complex than the block Gibbs sampler. For the MPF using 200 particles, the processing time per beat is approximately 0.5s for a non-optimized MATLAB implementation running on a 3.0-GHz Pentium IV computer, compared to about 0.3s for the method of beat-to-beat block Gibbs sampler in Section 3.2.

Table 3.2: Beat-to-beat block Gibbs sampler and the MPF method delineation performance comparison.

Method	Parameters	P_{on}	P_{peak}	P_{end}	T_{on}	T_{peak}	T_{end}
Beat-to-beat	annotations	3176	3176	3176	1345	3403	3403
Block Gibbs sampler	$Se\,(\%)$	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±19.3	-0.8±14.0	-3.1±14.0
Beat-to-beat	annotations	3176	3176	3176	1345	3403	3403
MPF	$Se\,(\%)$	99.45	99.45	99.45	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

3.4 Conclusion

This chapter studied Bayesian methods for beat-to-beat delineation and waveform estimation of P and T waves. Based on the multi-beat processing window model, we proposed a new Bayesian model for the non-QRS components of the ECG signal [LKT+12]. Instead of using a processing window containing several successive beats characterized by the same T and P waveforms, the proposed method accounted the T and P waveforms by processing individual beats sequentially. A sample-based Bayesian method was first studied to estimate the unknown parameters of the beat-to-beat Bayesian model. Then, in order to take advantage of all the available information contained in the past of the beat to be processed, a dynamic model is proposed [LBMT11, LGMT12]. This dynamic model exploited the sequential nature of the ECG by introducing a random walk model for the waveforms. Both the classical particle filter and a marginalized particle filter were studied to estimate the unknown parameters of the dynamic model. The main features and advantages of this chapter can be summarized as follows:

1. Beat-to-beat block Gibbs sampler [LKT+12]
    - The proposed Bayesian model uses the T and P waveform estimates of the previous beat as prior information for detecting/estimating the current T and P waves.
    - By properly accounting for the strong local dependencies in ECG signals and the sequential nature of ECG signals, the proposed block Gibbs sampler exhibits a significantly faster convergence compared to the related samplers used in Chapter 2.
    - The high accuracy of the proposed technique for T and P 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 window-based Bayesian methods. Moreover, it is ideally suited for on-line processing and, thus, for real-time ECG monitoring.
2. Marginalized particle filter [LBMT11, LGMT12]
    - Following the sequential Monte Carlo analysis principle, the sequential nature of the ECG signal is exploited by using a dynamic model under the Bayesian framework.
    - A marginalized particle filter is proposed to efficiently estimate the unknown parameters of the dynamic model. Benefiting from the marginalized particle filtering scheme, a smaller number of particles is needed compared to the classical particle filter to achieve a good estimation performance.
    - Compared to the beat-to-beat block Gibbs sampler method, the MPF method is optimal because all the available beats are considered in the waveform estimation.

As in Chapter 2, the resulting beat-to-beat Bayesian algorithms were validated using the QT database. A comparison with the multiple-processing window based Bayesian methods and with other benchmark methods demonstrated that both the beat-to-beat block Gibbs sampler and the MPF methods can provide significant improvements regarding T and P wave detection

rate, positive predictivity, and delineation accuracy. Meanwhile, the beat-to-beat processing mode is ideally suited for on-line processing and, thus, for real-time ECG monitoring. Further advantage includes the possibility of analyzing the beat-to-beat variation and evolution of the T and P waveforms. The next chapter introduces how the beat-to-beat Bayesian model and the associated Gibbs sampling method can be applied to a specific clinical research problem: T wave alternans (TWA) detection in intra-cardiac electrograms.

Chapter 4

Application in clinical research: T-wave alternans (TWA) detection

Contents

4.1 Introduction

This chapter presents an application of the Bayesian models introduced in this thesis to a specific clinical research topic referred to as T wave alternans (TWA) detection. Invisible (microvolt-level) TWA detection is one of the most challenging problems in cardiovascular signal processing. This issue has been the subject of the annual PhysioNet/Computers in Cardiology Challenge in 2008 [Phy08]. It refers to subtle beat-to-beat alternations in amplitude, shape or durations of the repolarization waves and has been closely related to the occurrence of spontaneous malignant ventricular arrhythmias [Nar06, RSS+94, RAC96].

Based on the window based Bayesian model presented in Chapter 2, a Bayesian approach is first proposed to deal with the surface ECG TWA detection problem. More precisely, the

window based Bayesian model introduced in Section 2.4 is slightly modified in order to perform T-wave delineation by taking into account a distinction between odd and even beats. The odd and even T-wave amplitudes generated by the block Gibbs sampler can be used to build statistical tests for TWA detection. This work concentrates on two tests: the two-sample Kolmogorov-Smirnov test (which is a non-parametric and robust method for comparing two samples) and the two-sample Student's t-test which is based on the assumption of normality for comparing the mean of two samples [Leh97]. Note that contrary to the statistical test proposed in [SLKG02], the proposed method computes multiple test statistics for each observation window (one per iteration of the Gibbs sampler) that can be used advantageously to derive detection performance (detection probability, probability of false alarm, receiver operational characteristics, etc.). The proposed algorithm is evaluated on real surface ECG signals subjected to synthetic TWA and compared with other classical algorithms. Compared with a test based on a single estimation, the proposed method provides information about the reliability of the detection which is important for medical diagnosis. It can also better handle small processing windows with few beats compared to the spectral analysis [SCV+94].

In Section 4.4, TWA detection on intracardiac electrogram (EGM) signals stored in implantable cardioverter defibrillator (ICD) memories is addressed by using the beat-to-beat based Bayesian model introduced in Chapter 3. Contrary to the surface ECG, the analysis of ICD stored signals suffers some important limitations due to the short periods of recordings (usually 10 to 15 sec, i.e., 10 to 20 beats) which hinder usual spectral and temporal analysis, and due to the fixed filtering, amplification and sampling rate preventing upstream tailored signal processing. Furthermore, if surface ECG TWA appears as a consistent fluctuation in the repolarization morphology on an every-other-beat basis (A-B-A-B...), the EGM TWA could happen with other patterns (A-B-C-A-B-C..., etc.). The proposed beat-to-beat Bayesian approach is appropriate for EGM TWA detection because it allows a beatwise waveform estimation which enables an analysis of the beat-to-beat variation and evolution of T waveforms. Besides the wave amplitude, additional parameters exploring the whole repolarisation (see Table 4.2) defined by a cardiologist from the Toulouse Rangueil Hospital can also be measured by using the waveform estimates of the proposed method. Dimensionality reduction techniques and statistical tests are then employed to exploit these parameters to analyze beat-to-beat T wave variations. The proposed approach is applied on real clinical data provided by St. Jude Medical, Inc and is shown to be promising for endocardial TWA detection.

4.2 TWA detection problem

T wave alternans (TWA) refers to subtle beat-to-beat alternations in amplitude, shape or duration of the repolarisation waves (T waves) and has been closely related to the occurrence of spontaneous malignant ventricular arrhythmias and sudden cardiac death [RSS+94]. In surface ECG, TWA is defined as a consistent fluctuation in the repolarization morphology which repeats on an every-other-beat basis (see Fig. 4.1 A-B-A-B repetition pattern).

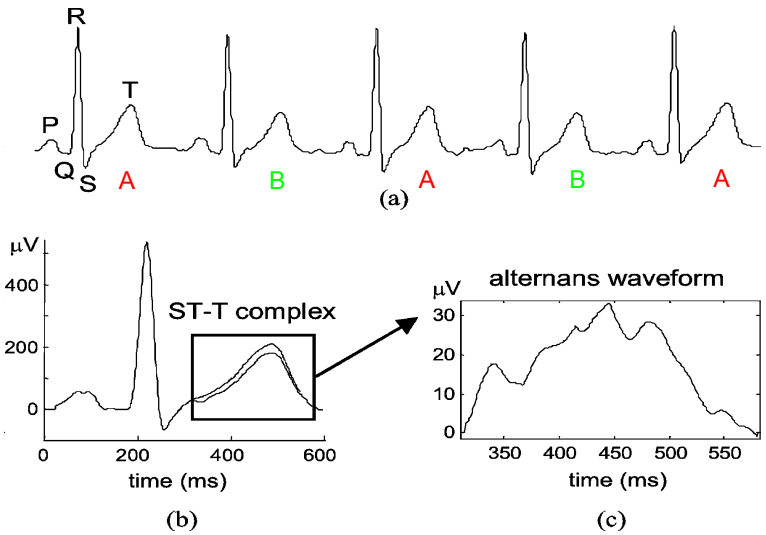


Figure 4.1: (a) ECG signal with visible TWA. (b) Superposition of two consecutive beats. (c) Alternans waveform: difference between odd and even beats. Image adapted from [MLM09].

Since the first report of non-visible (microvolt-level) TWA by Adam *et al.* in the 1980s, intensive research has been conducted on developing TWA detection and estimation algorithms [AAC81]. A complete and comprehensive review of signal processing methods to detect and estimate TWA proposed before 2005 can be found in [MO04]. The most widely used techniques are the spectral method (SM) [RSS+94, SCV+94] and the modified moving average method [NV02]. Alternative techniques are the statistical test method [SLKG02] and the Laplacian likelihood ratio method [MOWL06]. More recently new techniques have appeared including the multilead TWA detection by using principal component analysis [MLM09] and an empirical-mode decomposition based method [BVCRGLB10]. The fact that TWA amplitude is in the range of microvolts, together with the presence of the baseline and the physiological noise in the ECG make the TWA detection a difficult task. The main drawback of existing TWA analysis approaches is either their sensitivity to the presence of nonalternant components with high amplitude or their poor sensitivity to low-level TWA. Another problem with existing methods is that they generally require preprocessing steps for baseline suppression, rough segmentation or alignment of ST-T complexes. Thus, their performance is strongly influenced by the quality of these preprocessing procedures. Moreover, few existing methods allow an accurate TWA waveform estimation, while it has been proved that the TWA waveform characterization is important, e.g., to detect arrhythmic risk [NS99a]. In this section, the preprocessing procedure and several widely used methods are briefly reviewed.

4.2.1 Preprocessing

The aim of the preprocessing step is to condition the acquired ECG for posterior analysis. As TWA is a beat-to-beat phenomenon and is associated with the cardiac repolarization (ST-T complex), an alignment and segmentation procedure must be defined. QRS detection and T wave delineation are, therefore, necessary tasks. Also in this stage, the signal quality can be enhanced by filtering the signal and by removing the baseline wander. A general preprocessing scheme for TWA detection is shown in Fig. 4.2 and each step of this stage is listed below.

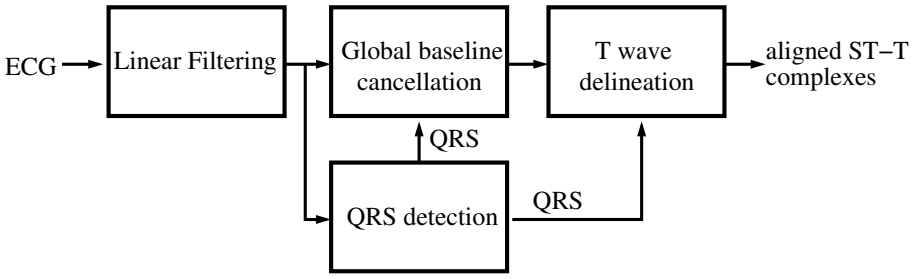


Figure 4.2: General TWA preprocessing stage.

Linear filtering. First, low-pass linear filtering is used to reject out-of-band noise. Similar techniques can be used as in the general ECG signal preprocessing procedure introduced in Section 1.4.1. In the references, the cutoff frequencies range from 50 to 360 Hz (which depend on the ECG sampling frequency). However, QRS complexes would lose their high-frequency components, which might degrade QRS detection and alignment. Moreover, the widened QRS could also invade the adjacent ST-T complex. Note that the requirements of this stage may also be increased or relaxed depending on the robustness of subsequent stages.

Baseline wander cancellation. Baseline fluctuations can manifest as high amplitude noise when analyzing the beat-to-beat ECG signal, degrading the analysis performance (masking subtle TWA or generating false detections). Cancellation of baseline changes in the preprocessing stage is therefore beneficial. As pointed out in [BZB06], the traditional interpolation methods (such as the widely used third-order spline interpolation [BZM99, MOL00]) allow removal of only baseline fluctuations at frequencies lower than heart rate, which makes the residual *local* baseline problematic for TWA detection.

ST-T complex segmentation. The criteria to define the segments of interest are diverse. In most methods, QRS locations are provided by a QRS detection. The ST-T complex is then selected with a fixed or RR-adjusted time window relative to the QRS position. Since the TWA analysis is affected by the performance of the alignment, T-wave delineator used for TWA analysis must show inter-beat stability in the fiducial point determination [NS99b].

Our proposed Bayesian approach as a preprocessing step for TWA detection. The Bayesian delineation algorithms proposed in chapters 2 and 3 can be good candidates for this preprocessing step. As introduced previously, the proposed Bayesian approaches all rely on the QRS detection method of [PT85]. The linear filtering that is done prior to this method is found to be satisfactory. Thus, no additional filtering is required. Moreover, the median filtering methods proposed in [CM07] are employed to remove the global baseline drift, and the

local baseline is included in the signal model and jointly estimated with the T wave parameters. Finally, ST-T complex segmentation and alignment is replaced by the proposed Bayesian algorithms. The Bayesian approaches proposed in this work allow simultaneous estimations of the T-wave location, amplitude and waveform and have shown good delineation performance. Thus no additional step is required for TWA analysis.

4.2.2 TWA detection methods

Based on the extracted and aligned ST-T complex, different estimation / detection techniques have been proposed in the literature to decide about the TWA presence or absence and possibly estimate the TWA amplitude and its waveform. Due to the transient, non-stationary nature of TWA, detection and estimation must involve a limited set of neighbor beats. In most cases, a $2D$-beat analysis window is shifted in order to cover the whole signal. Then, detectors and estimators are computed for the D even and D odd beats within each window. The analysis can be performed on non-overlapping or overlapping windows, the extreme case being the beat-by-beat sliding window analysis. Three state-of-the-art TWA detection methods are briefly introduced in this section. Note that only the basic principles of these methods are presented here, while many improvements or alternatives to these methods can be found in the literature. Interested readers are invited to consult [MO04] for more details.

Spectral method (SM)

The SM was first proposed in 1988 by Smith *et al.* [SCV+94] as a more elaborated version of the energy spectral method [AAC81]. After ECG beat alignment, the SM computes a periodogram-based power spectrum estimator for each sample of the segment to be processed. Since the TWA in surface ECG is defined as a consistent fluctuation in the repolarization morphology which repeats on an every-other-beat basis, the value of an aggregate spectrum at 0.5 cycles-per-beat is compared with the spectral noise level to decide if TWA is present. A slightly modified version has been presented by the same group in 1994 [RSS+94]. Since then, it has been extensively used for clinical research. It is included in commercial equipments, such as CH2000 and Heartwave (Cambridge Heart Inc, Bedford, MA).

Consider a $2D \times N$ aligned ST-T complexes matrix associated with a $2D$-beat window an of a $2D$-beat processing window

$$\boldsymbol{T} = \begin{pmatrix} T_1(1) & T_1(2) & \dots & T_1(N) \\ \vdots & \vdots & \ddots & \vdots \\ T_{2D}(1) & T_{2D}(2) & \dots & T_{2D}(N) \end{pmatrix}. \tag{4.1}$$

The periodogram of the n-th sample is obtained along the columns of $\boldsymbol{T}$ as

$$\hat{S}_n(f) = \frac{1}{2D} \left| \mathrm{TF} \left\{ T_k(n),\, k = 1, \dots, 2D \right\} \right|^2 \tag{4.2}$$

with $n \in \{1, \ldots, N\}$. And the mean periodogram can be obtained as follows

$$\overline{S}(f) = \frac{1}{N} \sum_{n=1}^{N} \widehat{S}_n(f) \tag{4.3}$$

The decision rule is defined in terms of a significance measure called TWA ratio (TWAR)

$$\frac{\overline{S}(0.5) - \hat{\mu}_{\text{noise}}}{\hat{\sigma}_{\text{noise}}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \tag{4.4}$$

where $\overline{S}(0.5)$ denotes the estimated mean spectrum at 0.5 cycle-per-beat and where $\hat{\mu}_{\text{noise}}$ and $\hat{\sigma}_{\text{noise}}$ are the mean and standard deviation of the equivalent noise measured at a properly chosen spectral window such as

$$\hat{\mu}_{\text{noise}} = \frac{1}{N_f} \sum_{f=f_{\min}}^{f_{\max}} \overline{S}(f) \tag{4.5}$$

$$\hat{\sigma}^2_{\text{noise}} = \frac{1}{N_f} \sum_{f=f_{\min}}^{f_{\max}} \left(\overline{S}(f) - \hat{\mu}_{\text{noise}}\right)^2 \tag{4.6}$$

where N_f is the number of the frequency points within the spectral window $[f_{\min}, f_{\max}]$. In the literature, a typical range for the spectral window is defined by the normalized frequencies $f_{\min} = 0.33$ and $f_{\max} = 0.48$. Note that applying a fixed threshold (typically $\gamma = 3$) to the TWAR is equivalent to applying a variable noise-dependent threshold.

The advantage of the SM is its simplicity and possibility to adapt to changing noise conditions. The main drawback is that a large window size (typically $2D = 128$) is required to have a good resolution in the spectral domain. This demands a large amount of ST-T complexes to be selected (the premature or ectopic beats should be rejected), extracted and aligned by using various techniques. The detection performance can be affected by these preliminary steps. It also should be noted that having a good detection performance for a small window length is required for medical diagnostics.

Modified moving average (MMA)

The MMA method was proposed by Nearing and Verrier in 2002 [NV02]. It is a time-domain procedure computing "continuously" a recursive running average of odd and even beats, where a limiting nonlinearity is applied to the innovation of every new beat to avoid the effect of impulsive artifacts. It is included in commercial equipments such as CASE-8000 (GE Medical Systems, Milwaukee, WI).

In the MMA method, a moving average is computed for both even and odd T waves. Let us denote $\bar{T}_k(n)$ the moving average of the nth sample of the T wave at the kth beat. This

moving average is computed recursively as follows

$$\begin{aligned}\bar{T}_0(n) &= T_0(n)\\ \bar{T}_1(n) &= T_1(n)\\ \bar{T}_k(n) &= \bar{T}_{k-2}(n) + g\left(\frac{T_{k-2}(n) - \bar{T}_{k-2}(n)}{2}\right),\ k = 2, 3, \ldots\end{aligned} \tag{4.7}$$

where $T_k(n)$ denotes the nth sample of the T wave at the kth beat and $g(x)$ is a function which is used to describe the average T wave evolution. If the T wave is larger (respectively lower) than the present moving average, the next moving average has to be increased (respectively decreased). Therefore, the proposed function g is a nonlinear limiting function taking into account a fraction of the difference between the T wave and the computed moving average

$$g(x) = \begin{cases} -K & \text{if } x < -K \\ x & \text{if } |x| \leq K \\ K & \text{if } x > K \end{cases} \tag{4.8}$$

An arbitrary value of $K = 32$ was proposed in [NV02].

The TWA at the kth beat is computed as the absolute difference between even and odd estimates

$$z_k(n) = \left|\bar{T}_k(n) - \bar{T}_{k-1}(n)\right| \tag{4.9}$$

and the global detection statistic of the kth beat is obtained as

$$Z_k = \max_n \left[z_k(n)\right] = \max_n \left|\bar{T}_k(n) - \bar{T}_{k-1}(n)\right| \tag{4.10}$$

with $n \in \{1, \ldots, N\}$. The detection statistic is then compared to a threshold which is determined from the receiver-operator characteristic (ROC) curve based on synthetic or real data. The main characteristic of the MMA method is its intrinsic robustness to outliers in the beat-to-beat series compared to linear methods (e.g., SM). However, it still requires the ST-T complexes to be well aligned and the TP segments that were not relatively isoelectric (corruption by electronic noise) to be eliminated.

Statistical test (ST)

A student t test has been proposed in [SLKG02] to decide whether the maxima of the odd and even beats have the same mean or not. More precisely, consider the ST-T complex matrix defined in (4.1). Let $a_{\mathrm{o},j}$ ($j \in \{1, \ldots, D\}$) denote the maximum values of the D odd ST-T complexes within the $2D$-beat window and $a_{\mathrm{e},j}$ denote the maxima of the even ST-T complexes

$$\begin{aligned}a_{\mathrm{o},j} &= \max\left[T_{2j-1}(1), T_{2j-1}(2), \ldots, T_{2j-1}(N)\right],\\ a_{\mathrm{e},j} &= \max\left[T_{2j}(1), T_{2j}(2), \ldots, T_{2j}(N)\right].\end{aligned} \tag{4.11}$$

The TWA detection problem can be formulated as

$$\mathcal{H}_0 : \mu_{\mathrm{o}} = \mu_{\mathrm{e}}, \quad \mathcal{H}_1 : \mu_{\mathrm{o}} \neq \mu_{\mathrm{e}} \tag{4.12}$$

where μ_{o} and μ_{e} are the means of $a_{\mathrm{o},j}$ and $a_{\mathrm{e},j}$. The t-test statistic can then be computed as follows

$$T = \frac{\hat{\mu}_{\mathrm{o}} - \hat{\mu}_{\mathrm{e}}}{\hat{S}_{\mathrm{eo}}\sqrt{\frac{2}{D}}} \tag{4.13}$$

with

$$\hat{\mu}_{\mathrm{o}} = \frac{1}{D}\sum_{j=1}^{D} a_{\mathrm{o,j}}, \quad \hat{\mu}_{\mathrm{e}} = \frac{1}{D}\sum_{j=1}^{D} a_{\mathrm{e,j}}$$

$$\hat{S}_{\mathrm{eo}} = \sqrt{\frac{1}{2D-2}\left(\sum_{j=1}^{D}\left(a_{\mathrm{o,j}} - \hat{\mu}_{\mathrm{o}}\right)^2 + \sum_{j=1}^{D}\left(a_{\mathrm{e,j}} - \hat{\mu}_{\mathrm{e}}\right)^2\right)}.$$

The p-value [1] of the test is compared to a significance level (the α-level) to make the decision. Due to the monotonic relation between the p-value and $|T|$, this decision rule is equivalent to decide TWA if $|T| > \gamma$ where γ is the critical value of the test statistic for the considered α-level. The resulting TWA detection strategy is

$$|T| \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{4.14}$$

As the SM, the decision rule in the t test is robust to changing noise conditions. The required significance level can be set by adjusting the parameter γ. The main drawback is that the test is based on the assumption of normality for the random variables $a_{\mathrm{o},j}$ and $a_{\mathrm{e},j}$. If this normality assumption does not hold, the test performance can decrease significantly. Furthermore, in [SLKG02], the beat features $a_{\mathrm{o},j}$ and $a_{\mathrm{e},j}$ are obtained by a rough estimation of the ST-T maximum. Thus it is sensible to noise and to the length of the estimation window.

[1] In statistical significance testing, the p-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true.

4.3 TWA detection in surface ECG using the window based Bayesian approach

4.3.1 Signal model for T waves in a $2D$-beat processing window

As presented in Section 4.2, most TWA detection methods are based on consecutive extracted and aligned T waves. The ST-T complex is usually selected with a fixed or RR-adjusted time window and then aligned by one of the various existing techniques. However, as it was verified in [NS99b], TWA analysis can be highly affected by the choice of alignment techniques since the T-wave delineator used for TWA analysis must show inter-beat stability in the fiducial point determination.

In this section we propose a modified Bayesian model, which is adapted from the window based Bayesian model introduced in Chapter 2. This model has the advantage of compensating the alignment errors and thus to carry out TWA detection. The main differences with the previous presented method is that the signal model focuses only on T wave estimation and delineation with a splitting into odd and even T wave sequences. However, to understand how this modified model is used for TWA detection, the signal model and the associated Bayesian inference are detailed in what follows.

The proposed method first detects QRS-complexes, which are the most prominent parts of the ECG signal. A $2D$-beat processing window is then shifted in a nonoverlapping way to cover the whole signal. In the processing window, the right hand neighborhood of each successive pair of QRS-offset constitutes a T-wave search interval (shown in Fig. 4.3(a)). The length of the nth T-wave search interval $N_{\mathrm{T},n}$ can be fixed either according to the cardiologists or simply as a fixed percentage of N_n, which is the length of a non-QRS interval ($n \in \{1, \ldots, 2D\}$). The T-wave search intervals are divided into odd T-wave blocks denoted as $\mathcal{J}_{\mathrm{o}} = \{\mathcal{J}_{\mathrm{o},1}, \cdots \mathcal{J}_{\mathrm{o},D}\}$ (containing the D odd T-wave search intervals) and even T-wave blocks denoted as $\mathcal{J}_{\mathrm{e}} = \{\mathcal{J}_{\mathrm{e},1}, \cdots \mathcal{J}_{\mathrm{e},D}\}$ (containing the even intervals).

As shown in Fig. 4.3(b), signals within each T-wave search interval can be approximated by one main pulse representing the T wave plus a local baseline. The odd (resp. even) waveforms are assumed to be constant within each processing window, contrary to the amplitudes and locations that vary with n. Therefore, T waves within D odd search intervals can be modeled by the convolution of an unknown waveform $\boldsymbol{h}_{\mathrm{o}} = (h_{\mathrm{o},-L} \cdots h_{\mathrm{o},L})^T$ (of length $2L+1$) with an unknown "impulse" sequence $\boldsymbol{u}_{\mathrm{o}} = (u_{\mathrm{o},1} \cdots u_{\mathrm{o},M})^T$ indicating the odd T-wave locations and amplitudes (see Fig. 4.3(c)). The impulse sequences can be defined as the products $u_{\mathrm{o},k} = b_{\mathrm{o},k} a_{\mathrm{o},k}$ of binary indicator sequences $b_{\mathrm{o},k} \in \{0,1\}$ and amplitude factors $a_{\mathrm{o},k} \in \mathbb{R}$. Each $b_{\mathrm{o},k} = 1$ indicates the location of an odd T wave, and the corresponding $a_{\mathrm{o},k}$ is the respective amplitude. Note that the $a_{\mathrm{o},k}$ are not defined for the k satisfying $b_{\mathrm{o},k} = 0$. Similarly, the even T waves within a window are modeled by the convolution of $\boldsymbol{h}_{\mathrm{e}} = (h_{\mathrm{e},-L} \cdots h_{\mathrm{e},L})^T$ with $\boldsymbol{u}_{\mathrm{e}} = (u_{\mathrm{e},1} \cdots u_{\mathrm{e},M})^T$. Let K denote the corresponding signal length. The ECG signal within

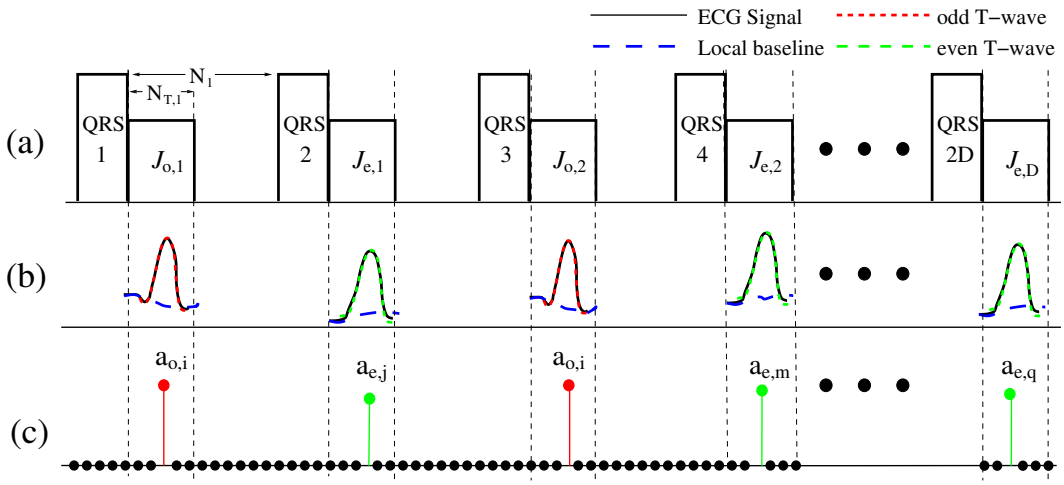


Figure 4.3: Signal model for TWA Bayesian analysis: (a) T-wave search intervals within the $2D$-beat processing window. (b) T waves within each non-QRS region. (c) T-wave amplitudes. Here we set $N_{\mathrm{T},n} = N_n/2$.

the processing window can be written as

$$x_k = \sum_{l=-L}^{L} h_{\mathrm{o},l}\left(a_{\mathrm{o},k-l} b_{\mathrm{o},k-l}\right) + \sum_{l=-L}^{L} h_{\mathrm{e},l}\left(a_{\mathrm{e},k-l} b_{\mathrm{e},k-l}\right) + c_k + w_k \tag{4.15}$$

where c_k denotes the baseline and w_k is the additive white Gaussian noise with unknown variance σ_w^2. Note that only indexes k belonging to a T-wave search interval are considered.

Baseline removal is generally recognized as an important processing step that is beneficial to TWA detection, while some of the traditional techniques (such as the widely used third-order spline interpolation) only allow the baseline fluctuations at frequencies lower than heart rate to be removed [BZB06]. This makes the estimation of local baseline essential for TWA detection. As in Section 2.4, the local baseline within the nth T search interval $\mathcal{J}_{\mathrm{T},n}$ is modeled by using a 4th-degree polynomial, i.e.,

$$c_{n,k} = \sum_{i=1}^{5} \gamma_{n,i} k^{i-1}, k = 1, \ldots, N_{\mathrm{T},n} \tag{4.16}$$

for each $n \in \{1, \ldots, 2D\}$. In vector-matrix form, (4.16) can be written as $\boldsymbol{c}_n = \boldsymbol{M}_n \boldsymbol{\gamma}_n$, where $\boldsymbol{M}_n$ is a known $N_{\mathrm{T},n} \times 5$ Vandermonde matrix and $\boldsymbol{\gamma}_n = (\gamma_{n,1} \cdots \gamma_{n,5})^T$ contains the unknown baseline coefficients. The baseline sequence for the entire $2D$-beat window can then be written as $\boldsymbol{c} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{2D}) = \boldsymbol{M}\boldsymbol{\gamma}$ where $\boldsymbol{M}$ is a $K \times 10D$ matrix and $\boldsymbol{\gamma}$ is a $10D \times 1$ vector.

Following the idea of Section 2.4.1, the odd and even T waveforms are represented by a basis expansion using discrete-time versions of Hermite functions to reduce the unknown parameter dimension

$$\boldsymbol{h}_{\mathrm{o}} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{o}}\,, \quad \boldsymbol{h}_{\mathrm{e}} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{e}} \tag{4.17}$$

where $\boldsymbol{H}$ is a $(2L+1) \times G$ matrix whose columns are the first G Hermite functions (with $G \leq 2L+1$), suitably sampled and truncated to length $2L+1$, and $\boldsymbol{\alpha}_{\mathrm{T},n}$ and $\boldsymbol{\alpha}_{\mathrm{P},n}$ are unknown coefficient vectors of length G.

Let $\boldsymbol{b}_{\mathrm{o}}$, $\boldsymbol{b}_{\mathrm{e}}$, $\boldsymbol{a}_{\mathrm{o}}$, and $\boldsymbol{a}_{\mathrm{e}}$ denote the $M \times 1$ vectors corresponding to $b_{\mathrm{o},k}$, $b_{\mathrm{e},k}$, $a_{\mathrm{o},k}$, and $a_{\mathrm{e},k}$, and $\boldsymbol{B}_{\mathrm{o}} \triangleq \mathrm{diag}(\boldsymbol{b}_{\mathrm{o}})$, $\boldsymbol{B}_{\mathrm{e}} \triangleq \mathrm{diag}(\boldsymbol{b}_{\mathrm{e}})$ denote the diagonal $M \times M$ matrices whose diagonal elements are formed by the components of $\boldsymbol{b}_{\mathrm{o}}$ and $\boldsymbol{b}_{\mathrm{e}}$. By concatenating (4.15) for $k = 1, \ldots, K$, where K is the number of ECG signal samples, the following matrix equation can be obtained

$$\boldsymbol{x} = \boldsymbol{F}_{\mathrm{o}}\boldsymbol{B}_{\mathrm{o}}\boldsymbol{a}_{\mathrm{o}} + \boldsymbol{F}_{\mathrm{e}}\boldsymbol{B}_{\mathrm{e}}\boldsymbol{a}_{\mathrm{e}} + \boldsymbol{M}\boldsymbol{\gamma} + \boldsymbol{w} \tag{4.18}$$

where $\boldsymbol{F}_{\mathrm{o}}$ is the $K \times M$ Toeplitz matrix with first row $[(\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{o}})_0 \ \cdots \ (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{o}})_0 \ \ \boldsymbol{0}]$, and first column $\left[(\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{o}})_0^T \ \cdots \ (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{o}})_L^T \ \ \boldsymbol{0}^T\right]^T$, while $\boldsymbol{F}_{\mathrm{e}}$ is the Toeplitz matrix of size $K \times M$ with first row $[(\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{e}})_0 \ \cdots \ (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{e}})_0 \ \ \boldsymbol{0}]$, and first column $\left[(\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{e}})_0^T \ \cdots \ (\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{e}})_L^T \ \ \boldsymbol{0}^T\right]^T$.

4.3.2 Bayesian inference

The unknown parameter vector resulting from the above parametrization is $\boldsymbol{\theta} = (\boldsymbol{\theta}_{\mathrm{o}}^T \ \boldsymbol{\theta}_{\mathrm{e}}^T \ \boldsymbol{\theta}_{\mathrm{cw}}^T)^T$, where $\boldsymbol{\theta}_{\mathrm{o}} \triangleq (\boldsymbol{b}_{\mathrm{o}}^T \ \boldsymbol{a}_{\mathrm{o}}^T \ \boldsymbol{\alpha}_{\mathrm{o}}^T)^T$ and $\boldsymbol{\theta}_{\mathrm{e}} \triangleq (\boldsymbol{b}_{\mathrm{e}}^T \ \boldsymbol{a}_{\mathrm{e}}^T \ \boldsymbol{\alpha}_{\mathrm{e}}^T)^T$ are related to the odd and even T waves, and $\boldsymbol{\theta}_{\mathrm{cw}} \triangleq (\boldsymbol{\gamma}^T \ \sigma_w^2)^T$ is related to the baseline and noise. Bayesian detection/estimation relies on the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x}) \propto p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ where $p(\boldsymbol{x}|\boldsymbol{\theta})$ is the likelihood function and $p(\boldsymbol{\theta})$ is the prior distribution of $\boldsymbol{\theta}$.

Likelihood function. Using our model (4.18) and the fact that $\boldsymbol{\omega}$ is white Gaussian, the likelihood function is obtained as

$$p(\boldsymbol{x}|\boldsymbol{\theta}) \propto \frac{1}{\sigma_w^K} \exp\left(-\frac{1}{2\sigma_w^2} \|\boldsymbol{x} - \boldsymbol{F}_{\mathrm{o}}\boldsymbol{B}_{\mathrm{o}}\boldsymbol{a}_{\mathrm{o}} - \boldsymbol{F}_{\mathrm{e}}\boldsymbol{B}_{\mathrm{e}}\boldsymbol{a}_{\mathrm{e}} - \boldsymbol{M}\boldsymbol{\gamma}\|^2\right),$$

where $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$.

Prior distributions. Since there is no known relation between $(\boldsymbol{b}_{\mathrm{o}}, \boldsymbol{a}_{\mathrm{o}})$, $(\boldsymbol{b}_{\mathrm{e}}, \boldsymbol{a}_{\mathrm{e}})$, $\boldsymbol{\alpha}_{\mathrm{o}}$, $\boldsymbol{\alpha}_{\mathrm{e}}$, $\boldsymbol{\gamma}$, and σ_n^2, all these sets of parameters are assumed to be *a priori* statistically independent. We will now discuss the prior distributions chosen for these parameters. Let $\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}$, $n \in \{1, \ldots, D\}$ contain all entries of the odd T-wave indicator vector $\boldsymbol{b}_{\mathrm{o}}$ that are indexed by the odd T-wave interval $\mathcal{J}_{\mathrm{o},n}$. The indicators are subject to a *block constraint*: within $\mathcal{J}_{\mathrm{o},n}$, there is one T wave (i.e., $\|\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}\| = 1$) or none (i.e., $\|\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}\| = 0$), the latter case being very unlikely. Therefore, we define the prior of $\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}$ as

$$p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}) = \begin{cases} p_0 & \text{if } \|\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}\| = 0 \\ p_1 & \text{if } \|\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}\| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.19}$$

where $p_1 = (1-p_0)/N_{\mathrm{o},n}$ and p_0 is chosen very small. The indicators $\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}}$ are supposed independent, and all remaining entries of the total vector $\boldsymbol{b}_{\mathrm{o}}$ (i.e., entries outside the search

intervals $\mathcal{J}_{\mathrm{o}}$) are zero. Thus, the prior of $\boldsymbol{b}_{\mathrm{o}}$ is the product of the priors $p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}})$

$$p(\boldsymbol{b}_{\mathrm{o}}) = \prod_{n=1}^{D} p(\boldsymbol{b}_{\mathcal{J}_{\mathrm{o},n}})\,.$$

For the T-wave amplitudes $a_{\mathrm{o},k}$ corresponding to $b_{\mathrm{o},k}=1$ (recall that the $a_{\mathrm{o},k}$ are undefined otherwise), we choose a zero-mean Gaussian prior, i.e., $p(a_{\mathrm{o},k}|b_{\mathrm{o},k}=1)=\mathcal{N}(0,\sigma_a^2)$. This allows for both positive and negative amplitudes. Amplitudes at different k are modeled as statistically independent. It follows that $u_{\mathrm{o},k}=b_{\mathrm{o},k}a_{\mathrm{o},k}$ is the kth element of a Bernoulli-Gaussian sequence with block constraints. The priors of the even T-wave indicators $b_{\mathrm{e},k}$ and amplitudes $a_{\mathrm{e},k}$ are defined in a fully analogous way, with the same fixed hyperparameters p_0, p_1, and σ_a^2. Moreover, the even T-wave variables are supposed to be independent of the odd T-wave variables. The odd T-waveform vector is assigned a zero-mean Gaussian prior, i.e., $p(\boldsymbol{\alpha}_{\mathrm{o}}) = \mathcal{N}(\mathbf{0},\sigma_{\boldsymbol{h}}^2\mathbf{I}_{L+1})$, where $\mathbf{I}_{L+1}$ denotes the identity matrix of size $(L+1)\times(L+1)$. The same prior is chosen for the even T-wave coefficients, i.e., $p(\boldsymbol{\alpha}_{\mathrm{e}})=\mathcal{N}(\mathbf{0},\sigma_{\boldsymbol{\alpha}}^2\mathbf{I}_{L+1})$. The baseline coefficients $\gamma_{n,i}$ are also modeled as independent identically distributed zero-mean Gaussian, i.e., $p(\boldsymbol{\gamma})=\mathcal{N}(\mathbf{0},\sigma_{\gamma}^2\mathbf{I}_{5D})$. Note that these prior distributions are similar to those of the window based Bayesian model in Section 2.4.2.

Posterior distribution. The posterior of the parameter vector $\boldsymbol{\theta}$ is given by

$$p(\boldsymbol{\theta}|\boldsymbol{x}) \propto p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}) = p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}_{\mathrm{o}})p(\boldsymbol{\theta}_{\mathrm{e}})p(\boldsymbol{\theta}_{\mathrm{cw}}) \tag{4.20}$$

with $p(\boldsymbol{\theta}_{\mathrm{cw}}) = p(\boldsymbol{\gamma})p(\sigma_w^2)$, $p(\boldsymbol{\theta}_{\mathrm{o}}) = p(\boldsymbol{a}_{\mathrm{o}}|\boldsymbol{b}_{\mathrm{o}})p(\boldsymbol{b}_{\mathrm{o}})p(\boldsymbol{\alpha}_{\mathrm{o}})$ and $p(\boldsymbol{\theta}_{\mathrm{e}}) = p(\boldsymbol{a}_{\mathrm{e}}|\boldsymbol{b}_{\mathrm{e}})p(\boldsymbol{b}_{\mathrm{e}})p(\boldsymbol{\alpha}_{\mathrm{e}})$. The block Gibbs sampler introduced in Section 2.4.3 is used to generate samples asymptotically distributed according to $p(\boldsymbol{\theta}|\boldsymbol{x})$. From these samples, the discrete parameters $\boldsymbol{b}_{\mathrm{o}}$ and $\boldsymbol{b}_{\mathrm{e}}$ can be detected by means of the sample-based maximum a posteriori (MAP) detector whereas the continuous parameters $\boldsymbol{a}_{\mathrm{o}}$, $\boldsymbol{a}_{\mathrm{e}}$, $\boldsymbol{\alpha}_{\mathrm{o}}$, $\boldsymbol{\alpha}_{\mathrm{e}}$, $\boldsymbol{\gamma}$, and σ_w^2 can be estimated by means of the sample-based minimum mean square error estimator.

4.3.3 Bayesian TWA detection

As explained in the last section, the T-wave locations and amplitudes can be sampled according to their joint posterior within each T-wave search interval, i.e., according to (4.20). Taking advantage of the Gibbs sampling method, different statistical tests can be carried out on the T-wave amplitudes generated by the proposed sampler to detect TWA[2]. First, we consider the two-sample Kolmogorov-Smirnov (KS) test which is a classical nonparametric method for comparing two samples. Let $\boldsymbol{a}_{\mathrm{o}}^{(i)} = \left(a_{\mathrm{o},1}^{(i)},\ldots,a_{\mathrm{o},D}^{(i)}\right)^T$ and $\boldsymbol{a}_{\mathrm{e}}^{(i)} = \left(a_{\mathrm{e},1}^{(i)},\ldots,a_{\mathrm{e},D}^{(i)}\right)^T$ denote the odd and even T-wave amplitudes within the $2D$-beat window generated at the i-th iteration of the Gibbs sampler. The TWA detection problem can be formulated as the following binary hypothesis test

$$\mathcal{H}_0 : F_{\mathrm{o}} = F_{\mathrm{e}}, \quad \mathcal{H}_1 : F_{\mathrm{o}} \neq F_{\mathrm{e}} \tag{4.21}$$

[2]Note that the first iterations belonging to the so-called burn-in period are not considered for parameter estimation or for TWA detection.

where F_{o} and F_{e} are the cumulative distribution functions of the odd and even T-wave amplitude samples. The KS test statistic is defined as

$$s^{(i)} = \sup_{a} \left| \hat{F}_{\mathrm{o}}^{(i)}(a) - \hat{F}_{\mathrm{e}}^{(i)}(a) \right| \tag{4.22}$$

where $\hat{F}_{\mathrm{o}}^{(i)}$ and $\hat{F}_{\mathrm{e}}^{(i)}$ are the empirical distribution functions of $\boldsymbol{a}_{\mathrm{o}}^{(i)}$ and $\boldsymbol{a}_{\mathrm{e}}^{(i)}$, respectively.

The two-sample Student t-test can also be applied to compare the means of the two samples $\boldsymbol{a}_{\mathrm{o}}^{(i)}$ and $\boldsymbol{a}_{\mathrm{e}}^{(i)}$. Indeed, the TWA detection problem can be formulated as

$$\mathcal{H}_0 : \mu_{\mathrm{o}} = \mu_{\mathrm{e}}, \quad \mathcal{H}_1 : \mu_{\mathrm{o}} \neq \mu_{\mathrm{e}} \tag{4.23}$$

where μ_{o} and μ_{e} are the means of the odd and even T-wave amplitude samples. The t-test statistic is classically defined as

$$t^{(i)} = \frac{\overline{\boldsymbol{a}}_{\mathrm{o}}^{(i)} - \overline{\boldsymbol{a}}_{\mathrm{e}}^{(i)}}{S_{\mathrm{eo}}^{(i)} \sqrt{\frac{2}{D}}} \tag{4.24}$$

with

$$\overline{\boldsymbol{a}}_{\mathrm{o}}^{(i)} = \frac{1}{D} \sum_{j=1}^{D} a_{\mathrm{o},j}^{(i)}, \quad \overline{\boldsymbol{a}}_{\mathrm{e}}^{(i)} = \frac{1}{D} \sum_{j=1}^{D} a_{\mathrm{e},j}^{(i)}$$

$$S_{\mathrm{eo}}^{(i)} = \sqrt{\frac{1}{2D-2} \left(\sum_{j=1}^{D} \left(a_{\mathrm{o},j}^{(i)} - \overline{\boldsymbol{a}}_{\mathrm{o}}^{(i)} \right)^2 + \sum_{j=1}^{D} \left(a_{\mathrm{e},j}^{(i)} - \overline{\boldsymbol{a}}_{\mathrm{e}}^{(i)} \right)^2 \right)}$$

By computing the test statistics (4.22) or (4.24) at each iteration of the Gibbs sampler, we obtain $N_{\mathrm{eff}} = N_{\mathrm{r}} - N_{\mathrm{bi}}$ (N_{r} is the number of iterations after convergence and N_{bi} is the number of burn-in iterations) samples of the test statistics corresponding to the same $2D$-beat block. Then several solutions can be proposed to deal with the N_{eff} test results.

Since the sampling distribution of the test statistic $s^{(i)}$ or $t^{(i)}$ is known under the null hypothesis, a first solution could be to compare this distribution to the empirical one of the N_{eff} observed values of $s^{(i)}$ or $t^{(i)}$. Hence, thanks to a second statistical hypothesis test such as a Kolmogorov test, it could be possible to conclude on the binary hypothesis test (4.23).

In our case, we have proposed a simpler solution. A test level α is chosen (typically $\alpha = 5\%$), corresponding to the rejection of the null hypothesis while $\mathcal{H}_0$ is true (false alarm). Then the results of the N_{eff} tests are taken into account with this α level. Based on the percentage of acceptance or rejection of the null hypothesis, among all these N_{eff} tests, a decision is made. The interest of such a solution is to be able to give a decision on the binary hypothesis test (4.23) with reliability index which is defined as

$$\tau = \frac{\text{Number of test statistic samples rejecting hypothesis } \mathcal{H}_0}{\text{Total number of test statistic samples}}.$$

This reliability information about the decision can be useful for medical diagnostics.

Remark: It is reported that the TWA waveform (difference between two adjacent T waveforms) contains interesting information for clinical diagnostics. The TWA waveform estimator can be obtained by using the MMSE estimators of the T-waveform difference between the odd and even beats

$$\hat{\boldsymbol{h}}_{\mathrm{TWA}} = \frac{1}{N_{\mathrm{eff}}} \sum_{t=1}^{N_{\mathrm{eff}}} \left[\boldsymbol{h}_{\mathrm{o}}^{(N_{\mathrm{bi}}+t)} - \boldsymbol{h}_{\mathrm{e}}^{(N_{\mathrm{bi}}+t)} \right] \tag{4.25}$$

4.3.4 Simulation results

Biomedical signal processing techniques are usually evaluated on standard databases, where the output of the technique is compared to manual expert annotations. However, because TWA is often non visible due to its low amplitude (sometimes below the noise level), the lack of validation database has been a major problem for TWA analysis. Simulated alternans with real nonalternant ECG recordings are widely used in the community [MO04]. In the following simulations, 20 healthy ECG segments (128 beats each) with no positive TWA have been selected from different databases. TWA episodes are then simulated by adding and subtracting alternatively (on a every-other-beat basis) a Hanning window to the delineated T waves as in [BVCRGLB10]. Fig. 4.4 illustrates the TWA synthesizer scheme. A small TWA amplitude value of V_{alt}=35μV has been chosen for the evaluations. Two different physiological noise sources have been considered to evaluate the proposed TWA detectors under real noise conditions: electrode motion ("em") and muscular activity ("ma"). Note that the "em" and "ma" noises have been extracted from the MIT-BIH noise stress test database.

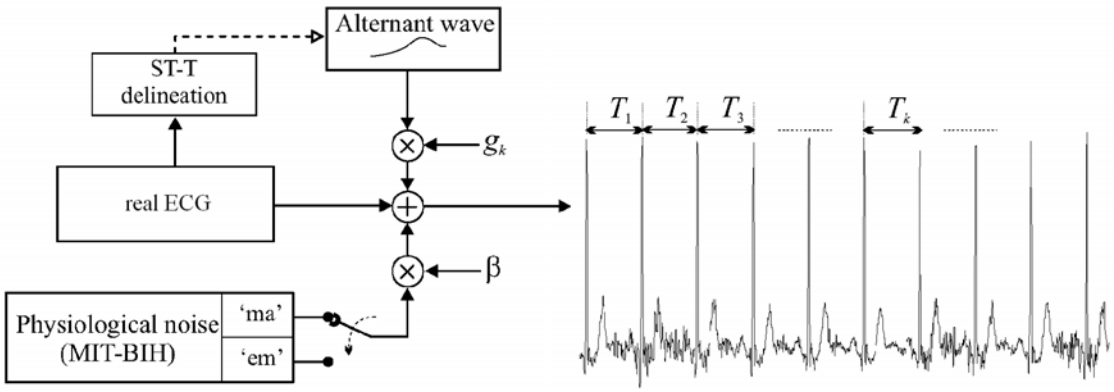


Figure 4.4: TWA synthesizer scheme with real ECG signal.

As a preprocessing step, the QRS complexes have been detected using the algorithm proposed in [PT85]. Based on the detected QRS complex locations, T-wave search intervals have been defined. The processing window length has been set to $2D$=16 beats, which is the smallest window length among the methods mentioned in [MO04]. Note that having a good detection performance with a small window length is beneficial for medical diagnostics. The Gibbs sampler studied in Section 2.4.3 has been run for each processing window with $N_{\mathrm{bi}} = 40$ burn-in iterations and $N_{\mathrm{eff}} = 100$ iterations to compute the estimates.

Figure 4.5 shows the estimation results for an ECG signal segment from the European ST-T dataset "e0303" with synthetic TWA. The "ma" noise has been added to the signal with SNR=10 dB. Typical estimates for the baseline and odd/even T waves are depicted in Fig. 4.5(a). Fig. 4.5(b) and (c) show the averages of T-wave estimates resulting from 20 Monte Carlo runs and the corresponding confidence intervals (error bars) for the odd and even beats within the 16-beat window.

As explained in Section 4.3.3, the KS test and t-test statistics can be determined for each iteration of the Gibbs sampler according to (4.22) and (4.24) providing $N_{\text{eff}} = 100$ decisions for each processing window. Fig. 4.6 shows representative situations for three different processing windows. Fig. 4.6 (a) and (d) show the KS test and t-test decisions for one window of dataset "e0303" with no synthetic TWA and "ma" noise (SNR=10dB). As can be seen, both tests have accepted the null hypothesis 100 times, therefore the null hypothesis can be accepted with full certainty ($\tau = 0$) for this window. Fig. 4.6 (c) and (f) show the decisions for one window of dataset "e0303" corrupted with synthetic 35 μv TWA and "ma" noise (SNR=10dB). The null hypothesis can be rejected with full certainty ($\tau = 1$) since both tests have rejected the null hypothesis 100 times. Fig. 4.6 (b) and (e) show results for a more complicated case where one window of dataset "e0303" is corrupted with synthetic 35 μv TWA and with a higher "ma" noise level than in the previous case (SNR=5dB). For this window, the null hypothesis can be rejected with detection probabilities $\tau = 0.93$ and $\tau = 0.82$ for the KS test and the t-test, respectively. It can be seen that, depending on signal characteristics (presence of physiological noise, baseline behavior, etc.), the null hypothesis rejection rate can be exactly unity (Fig. 4.6 (c) and (f)) or just below 1 (Fig. 4.6 (b) and (e)). Note again that using several samples from the Gibbs sampling iterations provides multiple test statistics allowing decision with an interesting reliability information.

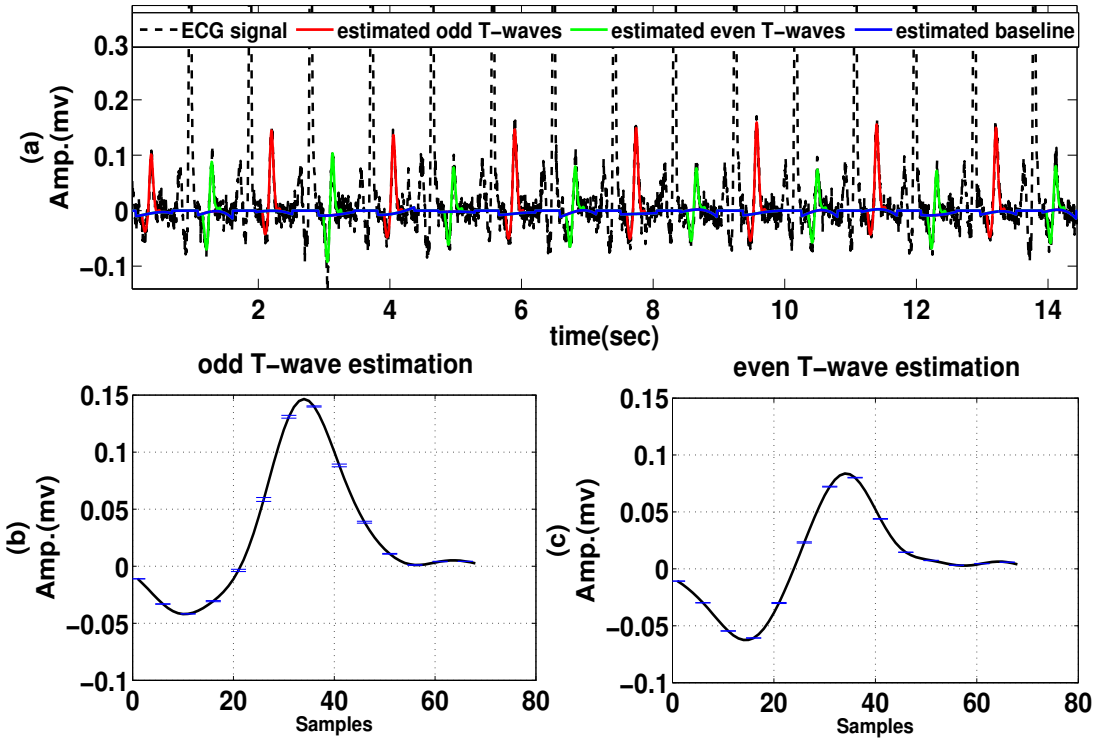


Figure 4.5: (a) Segment of dataset "e0303" with synthetic TWA and "ma" noise SNR=10 dB (black), estimated local baseline (blue), and estimated odd (red) and even (green) T waves. (b) Odd T-wave estimation averages (black) and the corresponding confidence intervals (blue) for the 16-beat window. (c) Even T-wave estimation averages (black) and the corresponding confidence intervals (blue) for the 16-beat window.

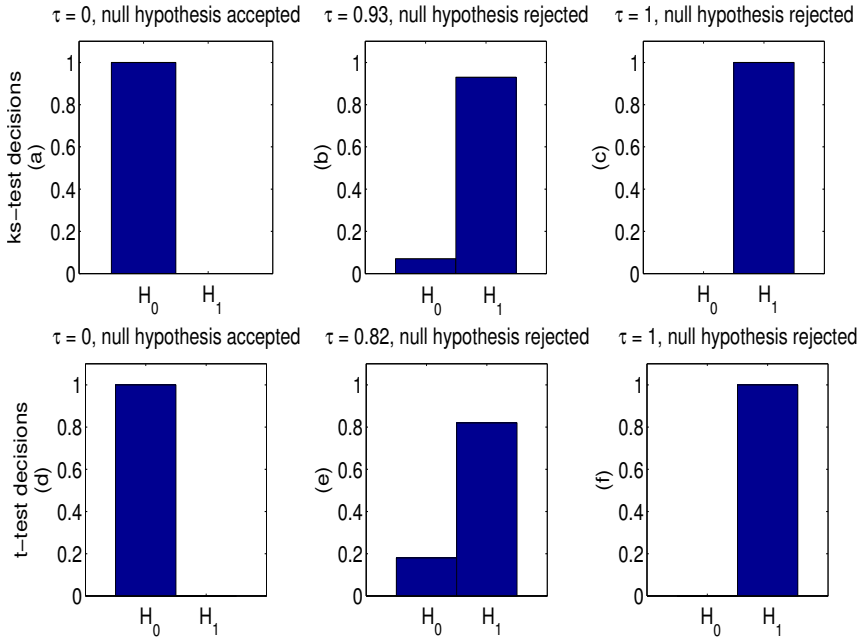


Figure 4.6: The KS-test (top) and the t-test (bottom) decisions made for three different 16-beat windows: (a) and (d) show the test decisions for one window of dataset "e0303" with no synthetic TWA and "ma" noise SNR=10dB; (b) and (e) show the test decisions for one window of dataset "e0303" with synthetic 35 μv TWA and "ma" noise SNR=5dB; (c) and (f) show the test decisions for one window of dataset "e0303" with synthetic 35 μv TWA and "ma" noise SNR=10dB.

For a quantitative comparison, we have implemented two classical methods, the spectral method (SM) [SCV+94] and the statistical test based on the maximum amplitude of the ST-T complex (ST) [SLKG02]. Fig. 4.7 shows the detection results achieved with real signals with synthetic $35\mu V$ TWA corrupted by "ma" and "em" noises for the Bayesian Gibbs sampler with KS test (BGS-KS) and t-test (BGS-T), the SM and the ST methods. Note that 20 Monte Carlo runs have been carried out for each SNR value, where the noise realizations have been changed from one simulation to another. The processing window length has been set to 16 beats for all the methods. As can be seen, the proposed Bayesian tests (referred to as BGS-KS and BGS-T) yield better results for both "ma" and "em" noise compared to ST and SM (e.g., an improvement of 10dB is achieved for having $P_{\mathrm{D}} = 1$). The KS test gives slightly better results than the t-test. Note however that the proposed methods have higher computational costs especially when using large processing windows (e.g., 128 beats). Thus their use is generally recommended for small processing windows.

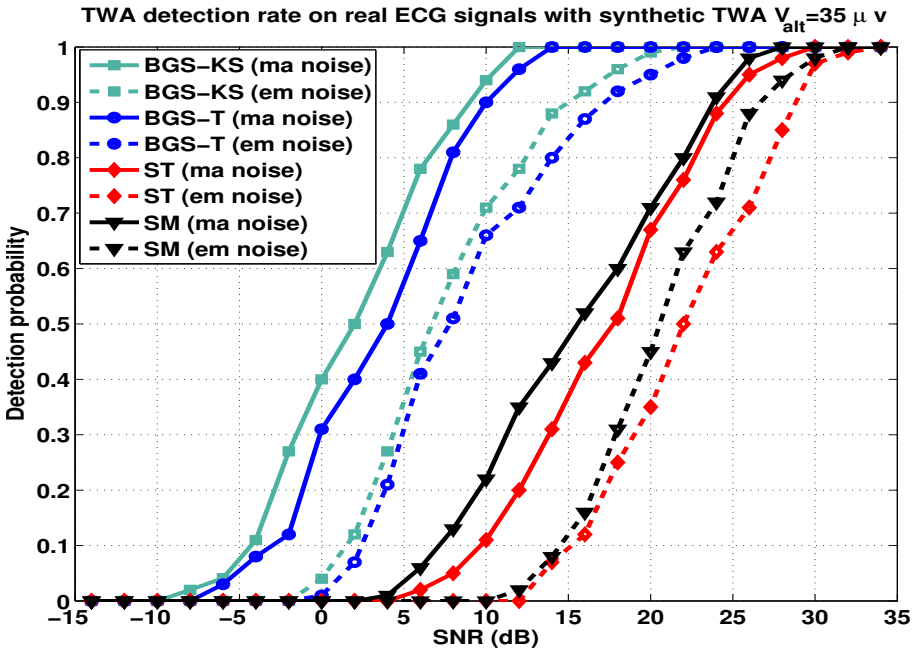


Figure 4.7: Detection performance for real ECGs with synthetic $35\mu V$ TWA. The proposed Bayesian Gibbs sampler with KS test (green square markers), Bayesian Gibbs sampler with t-test (blue round markers), the ST method (red diamond markers) and the SM (black triangle markers) are tested in both "ma" (continuous lines) and "em" (dotted lines) noise conditions.

4.4 Endocardial TWA detection using the beat-to-beat Bayesian approach

Section 4.3 presented a Bayesian TWA detection approach for surface ECGs by using the window based Bayesian model. In this section, a novel detection strategy based on the T wave **beat-to-beat** variations is proposed to address the TWA detection issue in the endocardial context, i.e., for electrocardiograms directly stored from implantable cardioverter defibrillators (ICDs). This strategy relies on the beat-to-beat Bayesian model introduced in Section 3.2 to estimate the T waveforms in each beat. Then, ten parameters defined by the cardiologist are subsequently extracted from the estimated waveforms to explore the whole repolarization. The absolute values of beat-to-beat differences between consecutive beats are then calculated for each extracted parameter. Discriminant analysis, uni-variate and multivariate statistic tests are then proposed to compare the absolute values of beat-to-beat parameter differences of the control signals (references) and the pre-onset [3] signals (episodes). The methodology developed in this section serves as a post-hoc analysis tool to the Endocardial T Wave Alternans Study (ETWAS) project in collaboration with St. Jude Medical Inc. and Toulouse Rangueil Hospital, which aims to prospectively assess the feasibility of TWA detection in inracardiac electrograms (EGMs). The contribution of our work is that our beat-to-beat Bayesian model allows a good beat-wise T waveform estimation and a dimensionality reduction step which enhances the detection performance. Moreover, the results obtained with our Bayesian algorithm have been useful to confirm results of other experiments conducted by the engineers from St. Jude Medical Inc.

4.4.1 TWA detection in ICD-stored intracardiac electrograms

EGMs stored in ICD memories represent a unique opportunity for detecting TWA immediately before ventricular tachycardia (VT) or ventricular fibrillation (VF) onset in patients prone to malignant ventricular arrhythmia. Compared to the surface ECG TWA detection problem, the analysis of ICD-stored signals suffers some important limitations due to

- the short periods of recordings available (usually 10 to 20 beats),
- the nature of observed TWA: contrary to the surface ECG case, where the TWA is only reported as a 0.5 cycle-per-beat (cpb) phenomenon (A-B-A-B...) [SCV+94], the EGM TWA could happen with other patterns (A-B-C-A-B-C..., etc.).

These limitations hinder usual spectral [SCV+94] and temporal [NV02] analysis which interest in detecting the alternant in 0.5 cpb. To the best of our knowledge, few papers in the biomedical engineering literature are related to automated endocardial TWA detection methods. The surface ECG TWA detection methods are often employed to deal with EGM TWA analysis problem. The possibility of using spectral methods for long-term EGMs (recorded during more than 5 min) has been discussed in [CSH+03]. The MMA method has been used to compute

[3] The term "pre-onset" defines the part of the signal before a cardiac problem such as a ventricular tachycardia or ventricular fibrillation (see Fig. 4.8).

repolarization alternans magnitude in porcine EGM signals in [MHI+05]. Moreover, from a clinical point of view, TWA has been reported up to now to be manually detectable in ICD-stored EGM in an isolated case [AAC+04] and in a recent study [SCD+11] as an irregular fluctuation on T wave amplitude.

Control and pre-onset data collection

A study on Endocardial T wave alternans (ETWAS) has been launched by St. Jude Medical to collect real clinical data and to assert the feasibility of detecting TWA in recorded EGMs. 57 patients from French university hospitals of Toulouse, Montpellier, Nîmes and Aix-en-Provence were enrolled in this study between November 2006 and December 2007. They were all implemented with single-chamber, dual-chamber or biventricular St. Jude Medical ICDs. First, the ICDs were programmed to record the EGM at a sampling rate of 128 Hz with 8-bit resolution and to filter the signal with a band-pass filter of 1.5 to 300 Hz. These values are considered to be acceptable for detecting beat-to-beat changes in intracardiac T wave, as reported in [PZG+06, SPN+07, KPP+09]. Second, in order to collect the largest amount of signal before the onset of each arrhythmia, the pre-trigger duration for stored EGM was programmed to its maximal value, i.e., 32 seconds. Third, the recorded intracardiac lead was programmed to be "ring-to-CAN", which means that the sensor has been positioned such that it yields high levels of TWA and is intended to explore wider myocardial areas (far-field bipolar recording) [PZG+06]. Fourth, the setting of amplification was tailored for each patient to record T waves as large as possible while avoiding signal clipping: this was usually programmed to a value of ± 7.4 mV, but was sometimes programmed to ± 3.7 mV (when T wave amplitude was smaller than 3 mV) or to ± 10.6 mV (for T wave amplitude larger than 7 mV).

Episode electrograms for which TWA are suspected were automatically stored in the device memory upon detection of a spontaneous tachyarrhythmia. Following the implantation, all patients were followed for 12 months with interim visits at 3, 6 and 9 months. All episode EGMs were retrieved from device memory at each follow-up visit.

Control baseline electrograms (references) were real-time electrograms streamed from the device memory to an external computer via an analog-to-digital acquisition card DAQcard 6062E (National Instruments, Austin, TX, USA) and a custom software LabWindows CVI. Real-time electrograms were sampled at 1kHz and 12-bit resolution in order to ensure signal quality. These EGMs which can be considered as TWA-free were recorded at rest during sinus rhythm at implantation (prior to anesthesia for ICD-testing) and/or during follow-up visits.

Thus, the real clinical data available consists of two sets:

- a set of "reference" EGMs for which it can be assumed that no TWA is present,
- a set of "episode" EGMs for which TWA is suspected but maybe not present in all signals from this set.

Data selection

The recorded EGMs were carefully reviewed by a cardiologist from the Hospital of Toulouse Rangueil. Patients with paroxysmal atrio-ventricular block or with permanent atrial fibrillation with spontaneous atro-ventricular conduction were not included in the study. EGM featuring noise, artefacts, T wave clipping or alternation between intrinsic and paced ventricular beats were eliminated from analysis. EGMs with atrial fibrillation before arrhythmia onset, EGM beats suffering from premature ventricular or supraventricular beats, arrhythmia episodes with a supraventricular origin or whose ventricular origin could not be formally proved were also eliminated from analysis. The same methodology was applied to select the control reference EGMs. Fig. 4.8 shows an example of a selected pre-onset episode followed by a ventricular fibrillation, and Fig. 4.9 shows a sinus rhythm reference signal recorded for the same patient during rest. Among the total 57 patients enrolled in the study, 16 patients (28.7%) experienced at least one episode of ventricular tachycardia or fibrillation identified by the cardiologist. Table 4.1 shows the number of selected episodes and references. On average, an episode contains approximately 30-40 beats (a 32 second window) and a reference has about 100-120 beats.

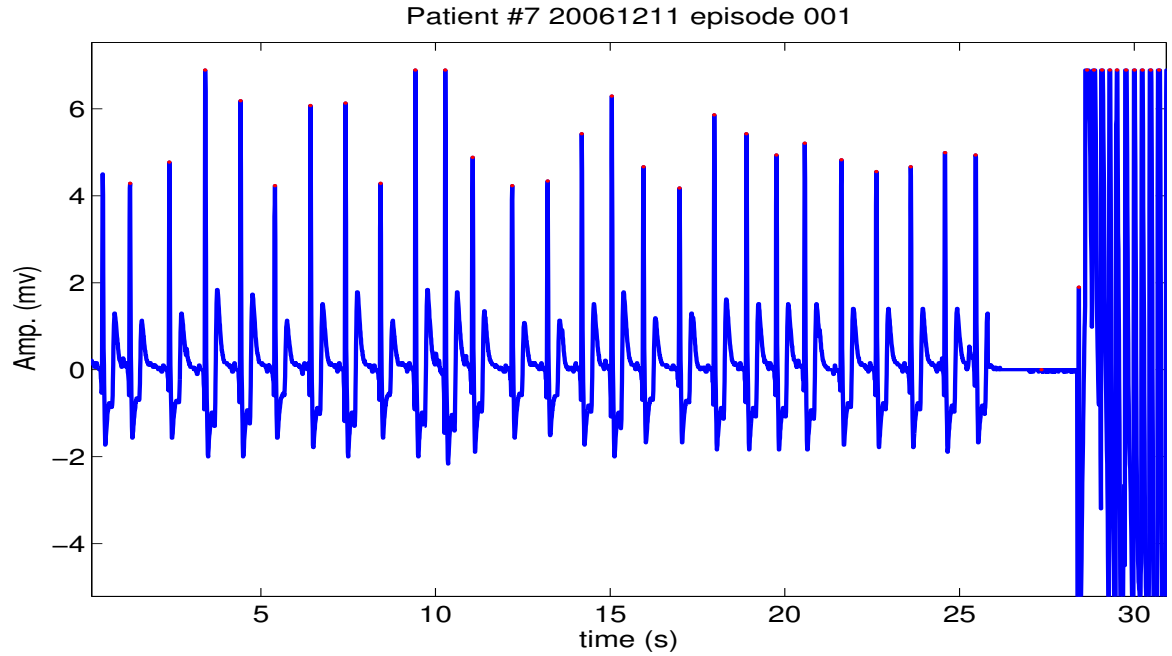


Figure 4.8: An example of selected EGM episode signal followed by a ventricular fibrillation.

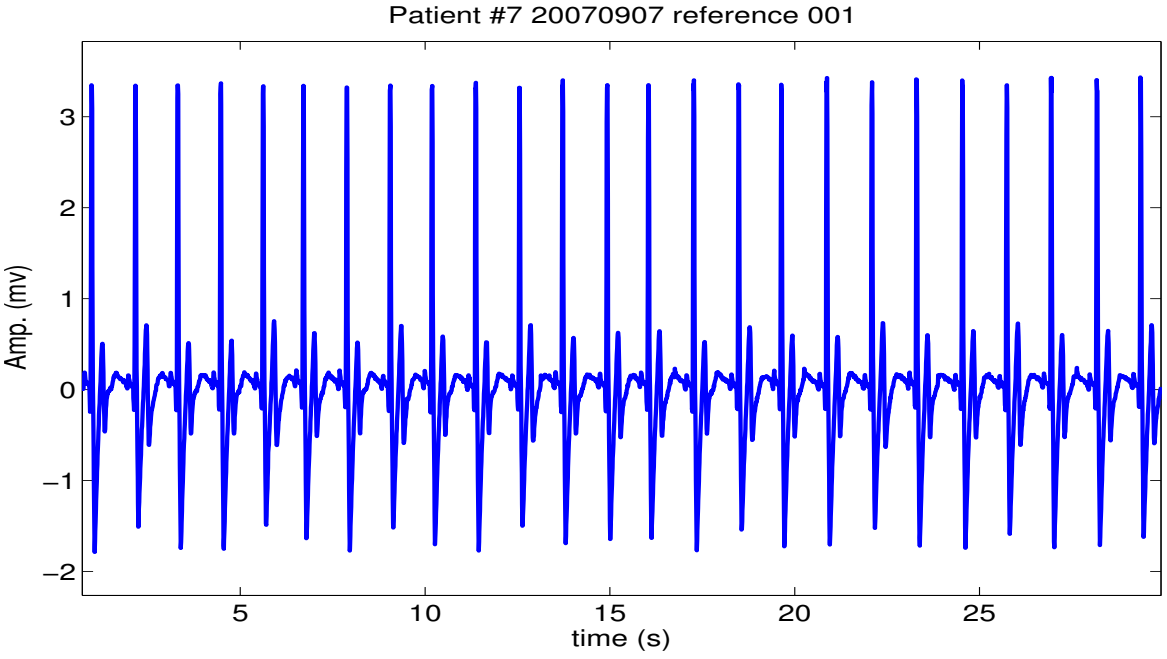


Figure 4.9: An example of EGM reference signal recorded for the same patient during sinus rhythm.

Table 4.1: Selected pre-onset episodes and control references for each patient.

Patients	Number of episodes	Number of references
Patient ♯1	1	6
Patient ♯2	1	3
Patient ♯3	1	6
Patient ♯4	1	5
Patient ♯5	3	12
Patient ♯6	1	5
Patient ♯7	4	15
Patient ♯8	1	6
Patient ♯9	1	4
Patient ♯10	1	3
Patient ♯11	3	16
Patient ♯12	3	7
Patient ♯13	1	2
Patient ♯14	1	6
Patient ♯15	13	16
Patient ♯16	1	4
Total datasets	**37**	**116**

4.4.2 T wave variation analysis and TWA detection

Beat-to-beat T waveform estimation

In the ETWAS project, selected EGMs (pre-onset and control recordings) were originally segmented and aligned on their QRS using a fixed time window relative to the QRS position. Despite the simplicity of this method, the segmentation and alignment quality could be affected due to the complexity of the signal nature. To address this problem, a modified version of the beat-to-beat Bayesian model introduced in Section 3.2 is proposed to estimate simultaneously the T-wave location and waveform in each heartbeat. The modified statistical model and its associated Bayesian model are detailed in the Appendix E. A block Gibbs sampler is employed to resolve the unknown parameters of the resulting posterior distribution. This sampler has a very similar structure to the block Gibbs sampler proposed in Section 3.2.3, except that the sampling distributions for some parameters are different due to the omission of P wave processing. From these samples, the discrete indicator parameter $\boldsymbol{b}_n$ is then estimated by means of the sample-based maximum a posteriori (MAP) estimator, and the continuous parameters $\boldsymbol{\alpha}_{\mathrm{T},n}$, $\boldsymbol{\gamma}_n$, and $\sigma^2_{w,n}$ are estimated by means of the sample-based minimum mean square error (MMSE) estimator, as described in Section 3.2.4. The next section discusses the T wave variation analysis based on the T waveform estimates $\hat{\boldsymbol{h}}_{\mathrm{T},n} = \boldsymbol{H}\hat{\boldsymbol{\alpha}}_{\mathrm{T},n}$ obtained by the MMSE estimator for the nth T wave search interval.

T wave parameters

Since the TWA could be more complicated than a simple amplitude or duration variation, the cardiologist defined six landmarks and $P = 10$ associated parameters to characterize the whole T wave. A post-hoc waveform analysis is then carried out based on the values of those parameters. The landmarks are defined as follows.

- L_1: The onset of the T wave.
- L_2: The point related to the maximum slope of the initial ascending part of the T wave.
- L_3: The peak of the T wave.
- L_4: The point related to the maximum slope of the terminal descending part of the T wave.
- L_5: The end of the T wave
- L_{QRS}: The R peak of the associated QRS complex.

As shown in Fig. 4.10, the six landmarks are calculated on the estimated T waveform $\hat{\boldsymbol{h}}_{\mathrm{T}}$. Note that the T wave onset P_1, peak P_3 and end P_5 can be obtained by using the delineation strategy proposed in Section 3.2.4. The slope can be obtained by computing the discrete-time approximation of the first derivative of the estimated T waveform, and the points P_2 and P_4 can be obtained as the local maxima of the slope before and after the wave peak, respectively.

Table 4.2 describes the $P = 10$ parameters related to the landmarks while Fig. 4.11 illustrates the schema of them on a EGM beat. The absolute values of parameter differences between consecutive beats are then calculated. For an $D+1$ beat EGM signal portion, these values are collected into a $P \times D$ difference matrix

$$\mathbf{\Delta} = \{\delta_{p,n}\}_{p=1,\cdots,P,\, n=1,\cdots,D} \tag{4.26}$$

with $\delta_{p,n}$ represents the absolute difference of the pth parameter between beats n and $n+1$.

Given one such beat-to-beat difference matrix $\mathbf{\Delta}$, the TWA detection can be formulated as a multivariate two-class problem

$\mathcal{H}_0$: No significant beat-to-beat wave parameter variation.
$\mathcal{H}_1$: Significant beat-to-beat wave parameter variation.

However, the $P = 10$ parameters may contribute unequally to the characterization of the beat-to-beat variation. For example, if the TWA is expressed as an amplitude alternation, the amplitude related parameters ("T_amplitude", "T_area", $\cdots$) will be more discriminant than wave shape related parameters ("T_max_asc_slope", $\cdots$) or rhythm related parameters ("QRS_T_apex_dur", $\cdots$). Most of the classifiers are strongly affected by the *curse of dimensionality*. In other words, when signals are represented in too high dimensional subspaces, the classifier performance can be impaired by the over-fitting problem. This problem is reduced by compressing the signal down to a lower-dimensional subspace by using dimensionality reduction techniques [DHS00]. In our case, 10 parameters may not introduce a curse of dimensionality. However, since these TWA detection algorithms are studied with the aim of implementing in the defibrillators, performing real-time processing, it is of great importance to study any implementation cost reduction.

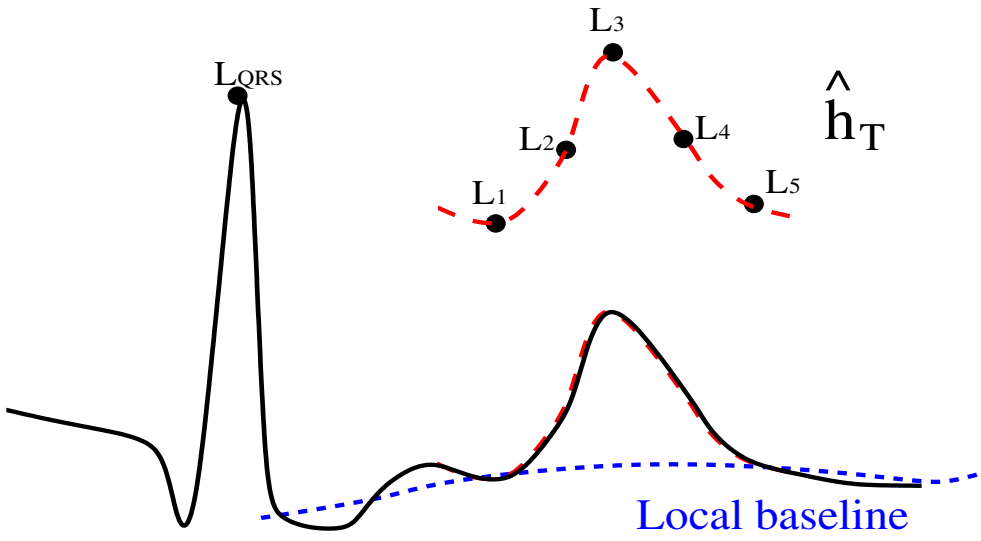


Figure 4.10: T wave landmarks calculated for each heartbeat.

Table 4.2: Description of the T wave parameters.

ID	Parameter	Description
1	T wave amplitude (μV) (T_amplitude)	T wave amplitude from peak (L_3) to local baseline
2	T wave area (μV·ms) (T_area)	T area between onset (L_1) and end (L_5)
3	Max T wave ascending slope (μV/ms) (T_max_asc_slope)	Absolute value of T wave slope maximal before L_3
4	Max T wave descending slope (μV/ms) (T_max_desc_slope)	Absolute value of T wave slope maximal after L_3
5	T wave duration (ms) (T_duration)	Timing between points of onset (L_1) and end (L_5)
6	T wave peak-end duration (ms) (T_apex_end_dur)	Timing between T wave peak (L_3) and T wave end (L_5)
7	R peak-T wave peak duration (ms) (QRS_T_apex_dur)	Timing of T wave end (L_5) relative to R peak (L_{QRS})
8	R peak-T wave Max ascending slope duration (ms) (QRS_T_max_asc_dur)	Timing of point L_2 relative to R peak (L_{QRS})
9	R peak-T wave Max descending slope duration (ms) (QRS_T_max_desc_dur)	Timing of point P_4 relative to R peak (L_{QRS})
10	R peak-T wave end duration (ms) (QRS_T_end_dur)	Timing of T wave end (L_5) relative to R peak (L_{QRS})

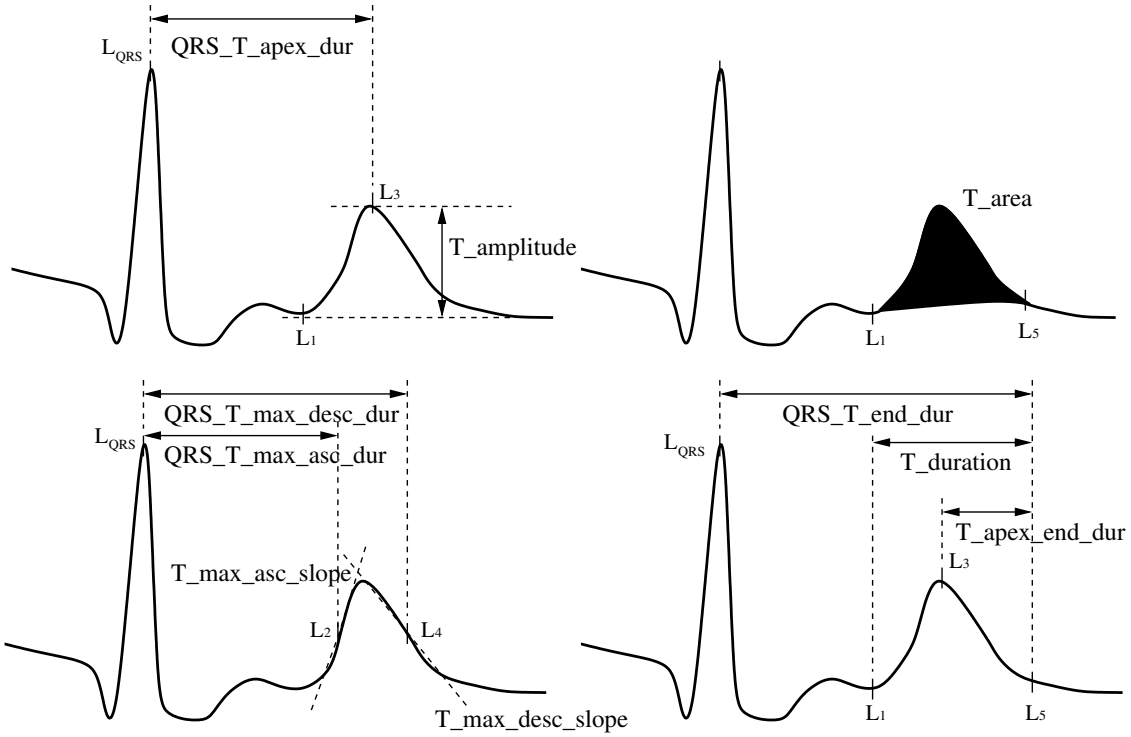


Figure 4.11: Schematic representation of the T wave parameters.

Dimensionality reduction: Fisher score

The Fisher criterion [Fuk90, DHS00] plays an important role for dimensionality reduction. It aims at finding a feature (parameter) representation minimizing an appropriate within-class distance and maximizing a between-class distance. Based on the Fisher criterion, two classes of dimensionality reduction methods have been proposed in the literature. The first class of methods is classically referred to as Feature Extraction with the *Linear Discriminant Analysis* (LDA) as an example. The other class contains feature selection methods, such as the *Fisher score*. One can find a thorough introduction of these methods in [DHS00]. In this chapter, the Fisher score has been used to reduce the dimensionality of the matrix $\boldsymbol{\Delta}$ and to select the most discriminant parameters to separate different classes. Note that one might think of using the feature extraction technique LDA. However, since there are only two classes in our case, the LDA suffers from the weakness of being able to provide only one discriminant axis. Furthermore, from a clinical point of view, it is also interesting to know which parameters are more discriminant when the alternation occurs.

The Fisher score can be computed for each parameter independently according to the criterion introduced in [DHS00]. Let $\boldsymbol{\Delta}^{(0)}$ and $\boldsymbol{\Delta}^{(1)}$ denote the difference matrices for a (D_0+1) beat control signal portion and for a (D_1+1) beat pre-onset episode portion, respectively. The mean and variance of the pth parameter difference for the control signal matrix $\boldsymbol{\Delta}^{(0)}$ can be obtained as

$$\hat{\mu}_{p,0} = \frac{1}{D_0}\sum_{n=1}^{D_0}\delta_{p,n}^{(0)}, \quad \hat{\sigma}_{p,0}^2 = \frac{1}{D_0}\sum_{n=1}^{D_0}\left(\delta_{p,n}^{(0)} - \hat{\mu}_{p,0}\right)^2 \tag{4.27}$$

where $\delta_{p,n}^{(0)}$ is the pth row and nth column of $\boldsymbol{\Delta}^{(0)}$. Similarly, the mean and standard deviation of the pth parameter difference of the pre-onset episode matrix $\boldsymbol{\Delta}^{(1)}$ can be obtained as

$$\hat{\mu}_{p,1} = \frac{1}{D_1}\sum_{n=1}^{D_1}d_{p,n}^{(1)}, \quad \hat{\sigma}_{p,1}^2 = \frac{1}{D_1}\sum_{n=1}^{D_1}\left(d_{p,n}^{(1)} - \hat{\mu}_{p,1}\right)^2. \tag{4.28}$$

Let $\boldsymbol{\Delta} = \left[\boldsymbol{\Delta}^{(0)}\ \boldsymbol{\Delta}^{(1)}\right]$ denote the joint difference matrix for the two classes. The mean and variance of the joint matrix corresponding to the pth parameter can be obtained as

$$\hat{\mu}_p = \frac{1}{D}\sum_{n=1}^{D}\delta_{p,n}, \quad \hat{\sigma}_p^2 = \frac{1}{D}\sum_{n=1}^{D}\left(\delta_{p,n} - \hat{\mu}_p\right)^2 \tag{4.29}$$

where $D = D_0 + D_1$ and $\delta_{p,n}$ is the pth row and nth column of $\boldsymbol{\Delta}$.

By using (4.27), (4.28) and (4.29), the Fisher score of the pth parameter can be computed as follows

$$F(\boldsymbol{\delta}_p) = \frac{D_0\left(\hat{\mu}_{p,0} - \hat{\mu}_p\right)^2 + D_1\left(\hat{\mu}_{p,1} - \hat{\mu}_p\right)^2}{D_0\hat{\sigma}_{p,0}^2 + D_1\hat{\sigma}_{p,1}^2}. \tag{4.30}$$

After computing the Fisher score for each parameter $\boldsymbol{\delta}_p$, we have to select the top-M ranked parameters which yield a normalized cumulative Fisher score superior to a given predefined value

(95% in our study). In a second step, different statistical tests (univariate t-test, Wilcoxon test, Kolmogorov Smirnov test and multivariate t-test [Leh97]) can be applied on the selected parameters. The general flowchart for the proposed EGM T wave beat-to-beat analysis algorithm (including beat-to-beat T waveform estimation), wave parameter extraction (including discriminant analysis) and statistical test used for TWA detection is summarized in Fig. 4.12.

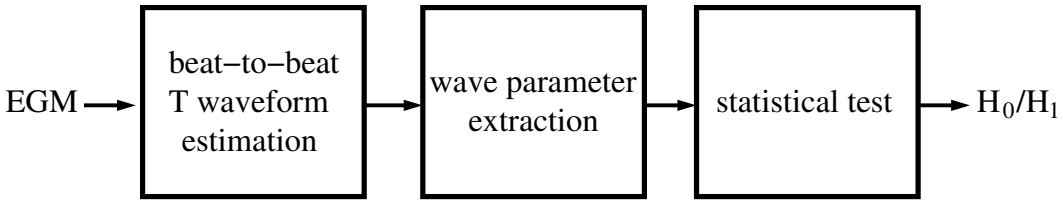


Figure 4.12: General block diagram for the beat-to-beat T wave variation analysis of EGM signals.

4.4.3 Clinical results and discussion

This section presents the simulation results of the proposed T wave analysis approach for both synthetic and real clinical data.

Synthetic data

Synthetic data is first used to validate the algorithm. Among the set of selected reference EGMs, 10 reference signals (approximately 500 beats) from 6 patients are chosen by the cardiologist as representative reference (control) signals. By using the TWA synthesis scheme presented in Fig. 4.4, TWA episodes are simulated by adding or subtracting alternatively (still on a every-other-beat basis) a Hanning window to the delineated T waves. Different TWA amplitude values of 50μv and 25μv have been chosen for the evaluations. Fig. 4.13 shows an example of EGM portion with synthetic 50μv and 25μv TWA.

Let $\mathbf{\Delta}_{\text{ref}}$ and $\mathbf{\Delta}_{\text{syn50}}$ denote the difference matrices computed for the control signals, and for the signals with synthesized 50μv TWA, respectively. The Fisher score of the two classes $\mathbf{\Delta}_{\text{ref}}$ and $\mathbf{\Delta}_{\text{syn50}}$ is obtained by using (4.30). As shown in Fig. 4.14, the wave amplitude and wave shape related parameters "T_amplitude", "T_area", "T_max_asc_slope" and "T_max_desc_slope" are more discriminant than the other parameters to separate $\mathbf{\Delta}_{\text{syn}}$ from $\mathbf{\Delta}_{\text{ref}}$. This result is coherent with the fact that the TWA has been synthetized by simulating alternans on the T wave amplitude and shape. Table 4.3 shows the examples of univariate test decisions for each parameter. Note that the different parameters are classified in descending order according to their respective Fisher scores and the significance level of the tests is fixed at 5% ($\alpha = 0.05$). As can be seen from the table, the first four parameters have a cumulative Fisher score of 98% and all the three univariate tests returned positive results. This means that the beat-to-beat variations of the two samples on the four most discriminant parameters are significantly different, which indicates an amplitude or shape alternation. Finally, in order to have a general decision on the two examples, a multivariate two sample t test is carried

out on the first four discriminant parameters which have a cumulative Fisher score superior to 0.95, and a positive decision is obtained.

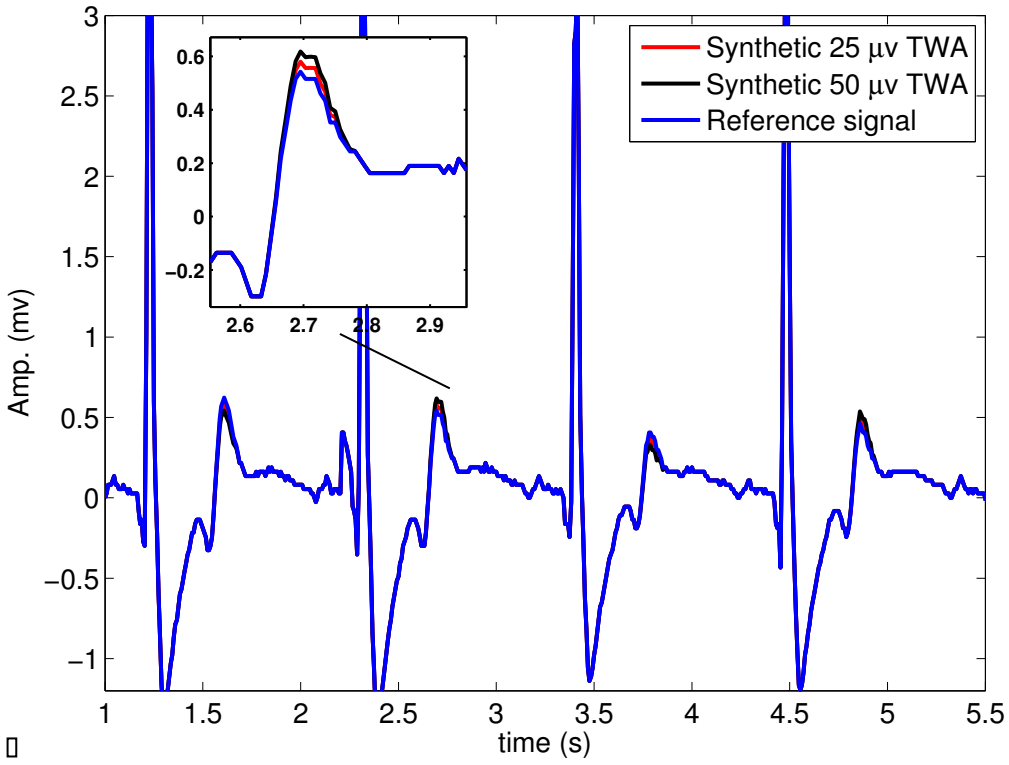


Figure 4.13: One EGM portion with synthetic TWA by adding alternatively a Hanning window to the delineated T waves. (Original reference signal in blue, reference signal with 25 μv TWA in red and reference signal with 50 μv TWA in black).

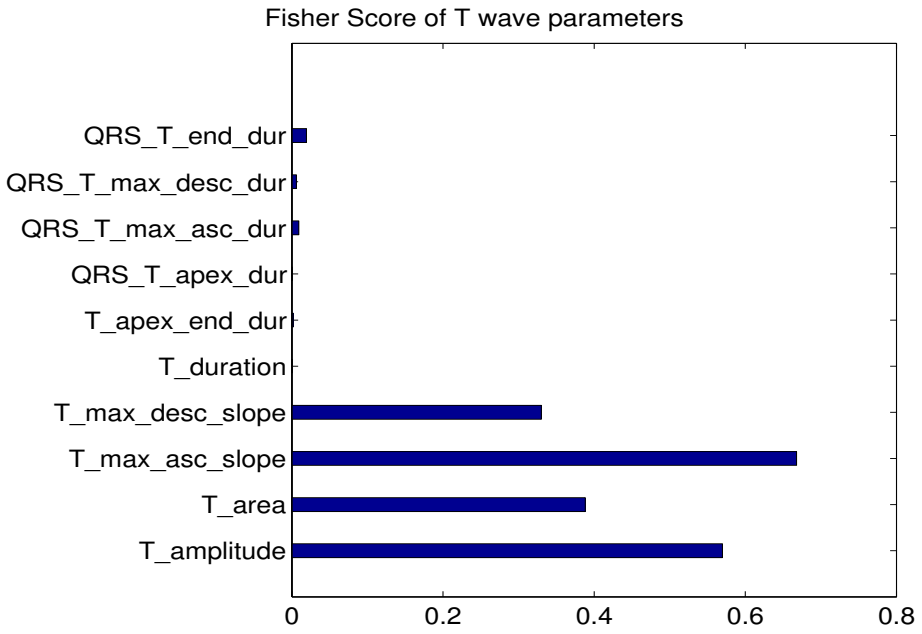


Figure 4.14: Fisher score of each parameter calculated on Beat-to-beat variations of reference and synthesized 50μv TWA data.

Table 4.3: Univariate test results on two samples $\mathbf{\Delta}_{\text{syn50}}$ and $\mathbf{\Delta}_{\text{ref}}$. Note that the parameters are classified in descending order according to their Fisher scores.

Parameter	**normalized Fisher score**	**cumulative Fisher score**	t**-test**	**KS-test**	**Wilcoxon-test**
T_max_asc_slope	0.3352	0.3352	H_1	H_1	H_1
T_amplitude	0.2859	0.6211	H_1	H_1	H_1
T_area	0.1949	0.8160	H_1	H_1	H_1
T_max_desc_slope	0.1656	0.9816	H_1	H_1	H_1
QRS_T_max_desc_dur	0.0097	0.9913	H_1	H_0	H_1
QRS_T_end_dur	0.0046	0.9959	H_0	H_0	H_0
QRS_T_apex_dur	0.0031	0.9990	H_0	H_0	H_0
T_apex_end_dur	0.0009	0.9959	H_0	H_0	H_0
QRS_T_max_asc_dur	0.0000	1.0000	H_0	H_0	H_0
T_duration	0.0000	1.0000	H_0	H_0	H_0

Another example is based on the reference signal with synthesized 25μv TWA. Let $\mathbf{\Delta}_{\text{syn25}}$ denote the difference matrix of the control signals with synthesized 25μv TWA. As can be seen in Fig. 4.15, the parameters that have high Fisher scores are still the amplitude and wave shape related parameters as in Fig. 4.14. However, the difference between the maximum and the minimum Fisher score is much smaller than those obtained in the previous example. This can be justified by the fact that the synthesized 25μv TWA data has smaller simulated beat-to-beat variations than those of the synthesized 50μv TWA. Table 4.4 presents the univariate test results on two samples $\mathbf{\Delta}_{\text{syn25}}$ and $\mathbf{\Delta}_{\text{ref}}$. Despite the small simulated alternans, the test results for the four most discriminant parameters return correct decisions, i.e., indicate an amplitude or shape alternans. Finally, the multivariate two sample t test carried on the four most discriminant parameters also provided positive decision.

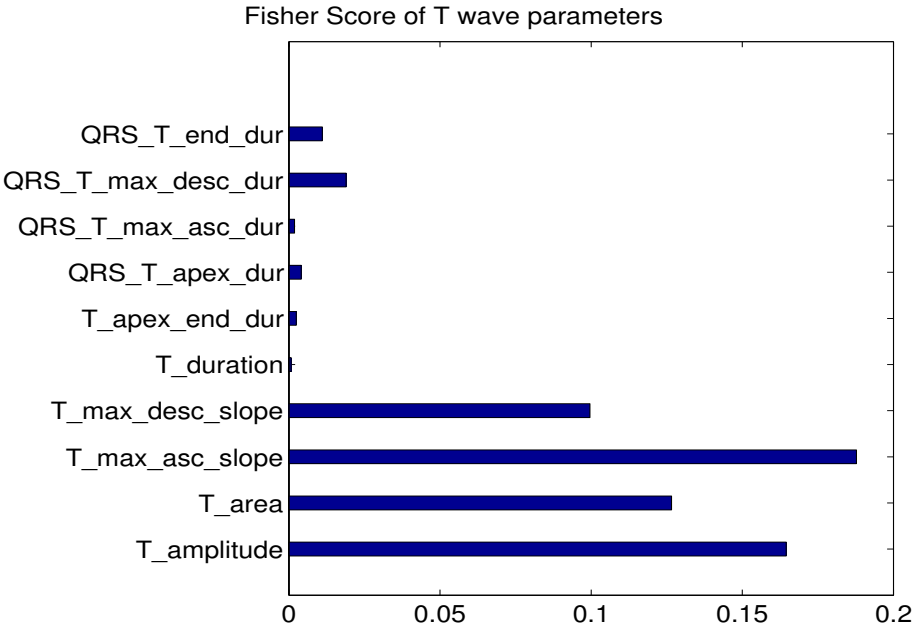


Figure 4.15: Fisher score of each parameter calculated on beat-to-beat variations of reference and synthesized 25μv TWA data.

Table 4.4: Univariate test results on two samples $\mathbf{\Delta}_{\text{syn25}}$ and $\mathbf{\Delta}_{\text{ref}}$. Note that the parameters are classified in descending order according to their Fisher scores.

Parameter	**normalized Fisher score**	**cumulative Fisher score**	***t*-test**	**KS-test**	**Wilcoxon-test**
T_max_asc_slope	0.2511	0.2511	H_1	H_1	H_1
T_amplitude	0.2509	0.5020	H_1	H_1	H_1
T_area	0.1764	0.6784	H_1	H_1	H_1
T_max_desc_slope	0.1083	0.7867	H_1	H_1	H_1
QRS_T_max_desc_dur	0.1021	0.8888	H_0	H_0	H_0
QRS_T_end_dur	0.0511	0.9399	H_0	H_0	H_0
QRS_T_apex_dur	0.0244	0.9643	H_0	H_0	H_0
T_apex_end_dur	0.0197	0.9840	H_0	H_0	H_0
QRS_T_max_asc_dur	0.0090	0.9930	H_0	H_0	H_0
T_duration	0.0070	1.0000	H_0	H_0	H_0

Results on real clinical data

This section presents the results obtained with real clinical data. As introduced in Section 4.4.1, a total of 37 pre-onset episodes of 16 different patients are selected under the guidance of an expert. For a patient who experiences at least one pre-onset episode, more than 3 control signals recorded during rest are selected as references. From a clinical point of view, it is more reasonable to compare a pre-onset episode of a given patient with his own reference data. Thus, the discriminant analysis and the statistical tests are carried out to compare each episode of a given patient with his total reference.

Example 1. The first example refers to a patient who has significant alternans in his ICD recorded episode compared to his reference signals (from the cardiologist point of view). Fig. 4.16(a) shows four consecutive heartbeats from a reference signal of patient ♯8. The corresponding sample-based estimates of the marginal posterior probabilities of having a T wave at a given location are depicted in Fig. 4.16(b). Fig. 4.16(c) shows the T waveforms estimated by the beat-to-beat Bayesian approach for each search interval. Fig. 4.17 shows the corresponding simulation results for one episode signal portion of the same patient. The two difference matrices of the reference and episode signals are then calculated from the estimated waveforms, respectively. Fig. 4.18 shows the Fisher score obtained for each parameter to separate the two samples. As can be seen, the wave amplitude related parameters are the most discriminant. Fig. 4.19 shows the box-and-whisker diagram (boxplot) of the beat-to-beat variations of the three most discriminant parameters. Note that there is a significant difference between the reference and the episode signals for each parameter. Table 4.5 shows the univariate test results on each of the parameter of the two samples and most of the test decisions are positive. The general test decision, obtained by using the multivariate two sample t test on the most discriminant parameters, also provides a positive decision.

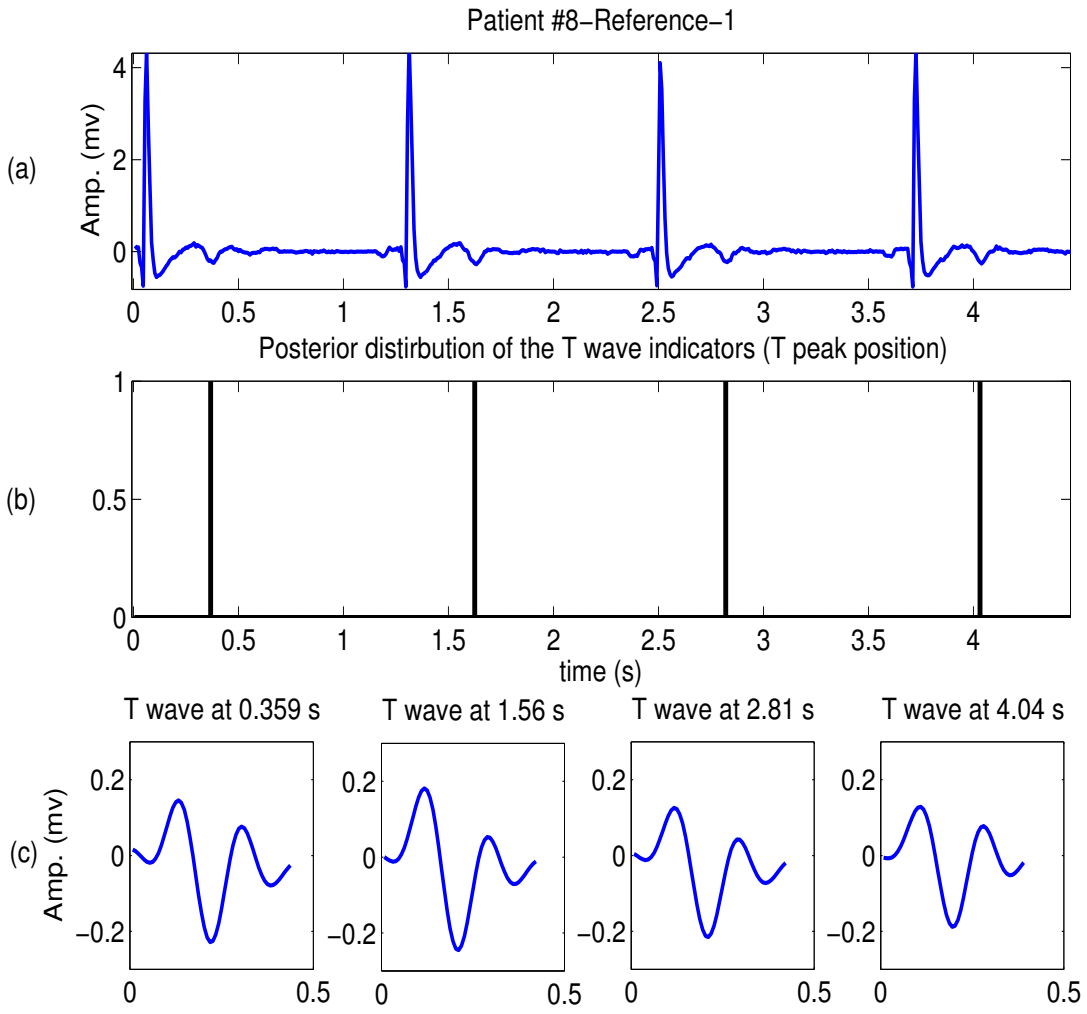


Figure 4.16: (a) Four consecutive heartbeats from one reference signal of patient ♯8; (b) Posterior distribution of the T wave indicators (wave peaks); (c) Estimated T waveforms.

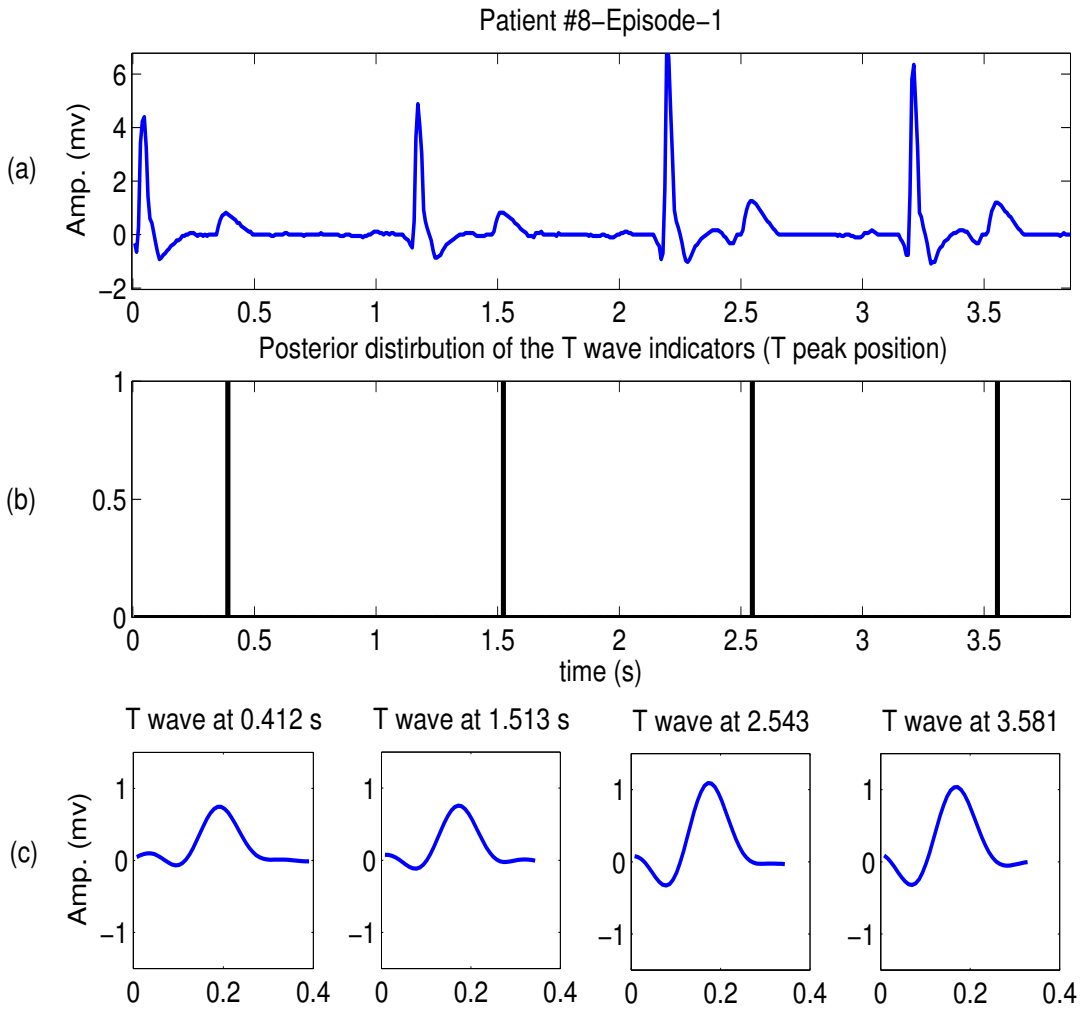


Figure 4.17: (a) Four consecutive heartbeats from one episode signal of patient ♯8; (b) Posterior distribution of the T wave indicators (wave peaks); (c) Estimated T waveforms.

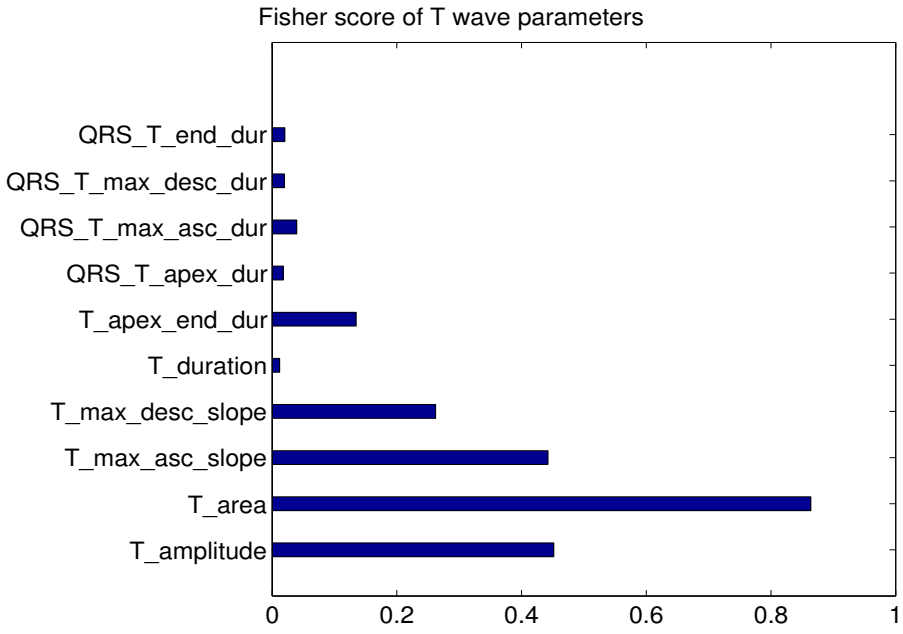


Figure 4.18: Fisher score of each parameter calculated on Beat-to-beat variations of reference and one episode signals of patient ♯8.

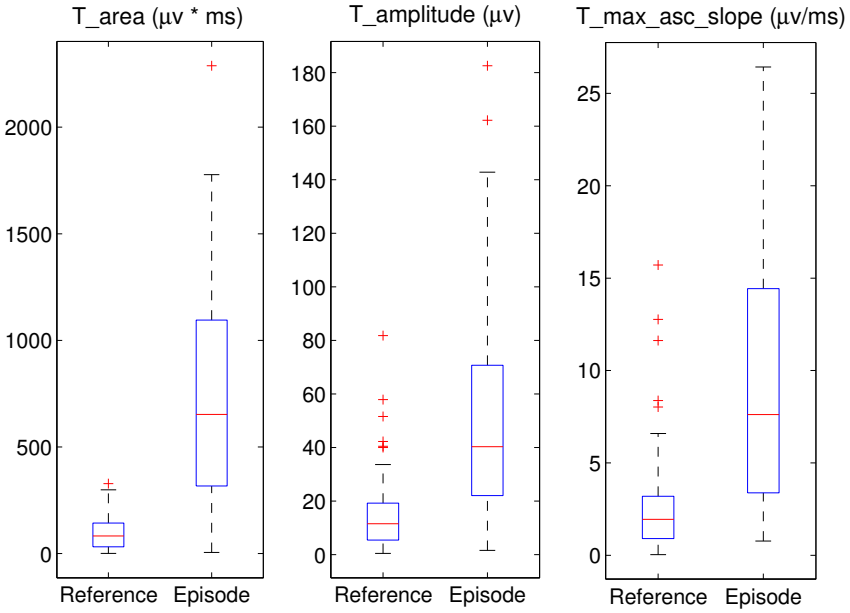


Figure 4.19: Beat-to-beat variation box-and-whisker diagram of the three most discriminant parameters of patient ♯8.

Table 4.5: Univariate test results on reference and episode signals of patient ♯8. Note that the parameters are classified in descending order according to their Fisher scores.

Parameter	**normalized Fisher score**	**cumulative Fisher score**	***t*-test**	**KS-test**	**Wilcoxon-test**
T_area	0.3814	0.3814	H_1	H_1	H_1
T_amplitude	0.1995	0.5809	H_1	H_1	H_1
T_max_asc_slope	0.1953	0.7763	H_1	H_1	H_1
T_max_desc_slope	0.1158	0.8920	H_1	H_1	H_1
T_apex_end_dur	0.0596	0.9516	H_1	H_1	H_1
QRS_T_max_desc_dur	0.0174	0.9690	H_1	H_1	H_1
QRS_T_end_dur	0.0091	0.9781	H_1	H_1	H_1
QRS_T_max_asc_dur	0.0087	0.9868	H_0	H_1	H_1
QRS_T_apex_dur	0.0079	0.9947	H_0	H_1	H_1
T_duration	0.0053	1.0000	H_0	H_1	H_1

Example 2. The second example is from a patient whose episode signal has no significant alternans compared to his reference signals. The Fisher scores of the 10 parameters to separate the episode from the reference signals are shown in Fig. 4.20. The rhythm related parameters have relatively higher scores yet the difference between the maximum and the minimum is much smaller than in Fig. 4.18. This is also confirmed by the boxplot of the most discriminant parameters as shown in Fig. 4.21. Table 4.6 presents the test results on each of the parameter of the two samples and most of the parameters show no significant difference between the episode signal and the reference. The parameters which have a cumulative Fisher score superior to 0.95 are included in the multivariate t test, and the general test decision for this episode is negative.

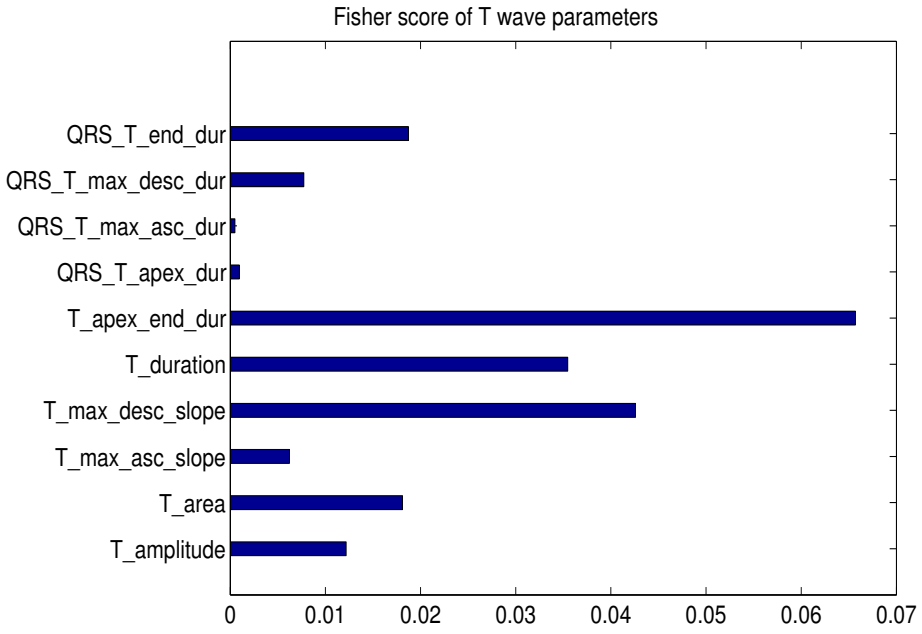


Figure 4.20: Fisher score of each parameter calculated on Beat-to-beat variations of reference and one episode signals of patient ♯15.

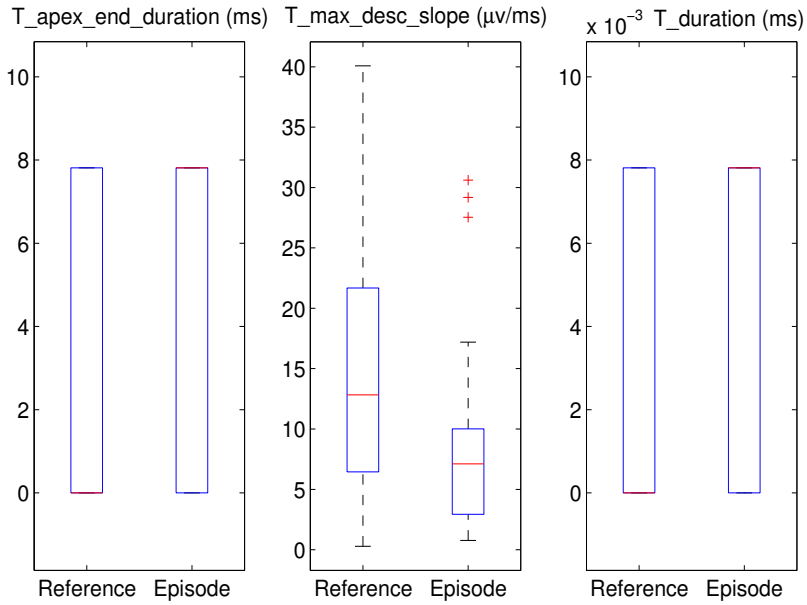


Figure 4.21: Beat-to-beat variation box-and-whisker diagram of the three most discriminant parameters of patient ♯15.

Table 4.6: Univariate test results on reference and episode signals of patient ♯15. Note that the parameters are classified in descending order according to their Fisher scores.

Parameter	**normalized Fisher score**	**cumulative Fisher score**	***t*-test**	**KS-test**	**Wilcoxon-test**
T_apex_end_dur	0.3156	0.3521	H_1	H_1	H_1
T_max_desc_slope	0.5202	0.5809	H_1	H_0	H_1
T_duration	0.6906	0.7763	H_1	H_0	H_1
QRS_T_end_dur	0.7805	0.8920	H_0	H_0	H_0
T_area	0.0596	0.8675	H_0	H_0	H_0
T_amplitude	0.0174	0.9260	H_0	H_0	H_0
QRS_T_max_desc_dur	0.9632	0.9781	H_0	H_0	H_0
T_max_asc_slope	0.9931	0.9868	H_0	H_0	H_0
QRS_T_apex_dur	0.0046	0.9977	H_0	H_0	H_0
QRS_T_max_asc_dur	0.0023	1.0000	H_0	H_0	H_0

Results on the full database. Table 4.7 shows the statistical test result on the 37 selected pre-onset signals. Each pre-onset episode has been tested against all the references of the same patient by using the multivariate t-test. Note that only the parameters which have cumulative Fisher score superior to 0.95 are included in the test. Among the 37 selected pre-onset signals, 13 (i.e., 35.14%) episodes have been found to have significant beat-to-beat variations compared to their respective reference signals.

Table 4.7: Statistic test result on the 37 selected pre-onset signals. Each pre-onset episode has been tested against all the references of the same patient by using the multivariate t-test. Only the parameters which have cumulative Fisher score superior to 0.95 are included in the test. Note that the number of pre-onset episodes labeled as positive TWA for each patient is indicated in the right column.

Patient pseudo	**Number of total episodes**	**Number of TWA positive episodes**
Patient ♯1	1	0
Patient ♯2	1	0
Patient ♯3	1	0
Patient ♯4	1	1
Patient ♯5	3	0
Patient ♯6	1	0
Patient ♯7	4	2
Patient ♯8	1	1
Patient ♯9	1	1
Patient ♯10	1	1
Patient ♯11	3	1
Patient ♯12	3	1
Patient ♯13	1	1
Patient ♯14	1	0
Patient ♯15	13	4
Patient ♯16	1	0
Total	**37**	**13**

Clinical discussion

In this section, we proposed a beat-to-beat Bayesian approach together with discriminant analysis and statistical tests to analyze the ICD stored EGM recordings provided by St. Jude Medical, Inc. First, the beat-to-beat Bayesian approach was used to estimate the T waveform from ICD stored EGM recordings. Then, beat-to-beat variations of various T wave parameters extracted from the estimated waveforms were subjected to discriminant analysis to reduce the dimensionality. Finally, univariate and multivariate statistical tests were carried out on the most discriminant parameters.

From a clinical point of view, the major finding of this prospective and multicenter study is that the beat-to-beat variations (amplitude, shape or duration) are significantly greater immediately before spontaneous VT/VF in ICD stored EGM recordings than during different types of control signals. Our results also reflect the highly variable beat-to-beat T wave variations from one patient to another, since no reproducible T wave behavior was observed in this population. Moreover, this study shows the interest of performing TWA detection in endocardial context using several parameters characterizing the T waves, while most of the related studies [SCD+11] focus only on TWA amplitude.

4.5 Conclusion

This chapter presented applications of the different Bayesian models introduced in this thesis to a specific clinical research topic referred to as T wave alternans (TWA) detection. First, an adapted version of the window based Bayesian model presented in Chapter 2 was proposed to address the surface ECG TWA detection problem. The odd and even T-wave amplitudes generated by the block Gibbs sampler were used to build statistical tests for TWA detection. The proposed algorithm was evaluated on real surface ECG signals subjected to synthetic TWA and compared with other classical algorithms. In the second part of the chapter, the TWA detection problem on intracardiac electrogram (EGM) signals stored in implantable cardioverter defibrillator (ICD) memories was addressed by using the beat-to-beat based Bayesian model introduced in Chapter 3. Based on the good waveform estimation provided by the beat-to-beat Bayesian approach, ten parameters describing the whole repolarisation defined by a cardiologist were extracted. Dimensionality reduction techniques and statistical tests were then employed to exploit these parameters to analyze beat-to-beat T wave variations. The proposed approach was applied on real clinical data provided by St. Jude Medical, Inc.

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

1. Window based Bayesian model for surface ECG TWA detection [LMT11]
    - The window based Bayesian model introduced in Section 2.4 is modified to perform T-wave delineation by taking into account a distinction between odd and even beats.
    - The odd and even T-wave amplitudes generated by the block Gibbs sampler are used to build statistical tests (Student's t-test and Kolmogorov-Smirnov test) for TWA detection.
    - The proposed method computes multiple test statistics for each observation window (one per iteration of the Gibbs sampler) that can be used advantageously to derive detection performance and to provide information about the reliability of the detection.
    - The proposed algorithm is evaluated on real surface ECG signals subjected to synthetic TWA and compared with other classical algorithms.
2. Beat-to-beat Bayesian approach for intracardiac TWA detection [LKT+12]
    - A modified version of the beat-to-beat Bayesian approach introduced in Chapter 3 is proposed to estimate the T waveform in each heartbeat.
    - A waveform characterization procedure which allows the extraction of ten parameters exporting the whole repolarisation from the estimated waveform is proposed.
    - Dimensionality reduction techniques are studied to find the most discriminant parameters.
    - The proposed approach is applied on real clinical data and is shown to be promising to endocardial TWA detection.

Conclusions and Perspectives

This thesis studied Bayesian estimation/detection algorithms for P and T wave analysis in ECG signals. In this work, different statistical models and associated Bayesian methods were proposed to solve simultaneously the P and T wave delineation task (determination of the positions of the peaks and boundaries of the individual waves) and the waveform-estimation problem. These models take into account appropriate prior distributions for the unknown parameters (wave locations and amplitudes, and waveform coefficients). These prior distributions are combined with the likelihood of the observed data to provide the posterior distribution of the unknown parameters. Due to the complexity of the resulting posterior distributions, Markov chain Monte Carlo algorithms were proposed for (sample-based) detection/estimation. On the other hand, to take full advantage of the sequential nature of the ECG, a dynamic model was proposed under a similar Bayesian framework and sequential Monte Carlo methods (SMC) were also investigated.

The first chapter recalled the physiological basis of the ECG. From the point of view of ECG computer-aided detection and computer-aided diagnosis, a brief review of the ECG signal processing literature was given, with the emphasis on P and T wave detection and delineation.

In the second chapter, we introduced a new Bayesian model based on a multiple-beat processing window which simultaneously solves the P and T wave delineation and the waveform estimation problems. This model is based on a modified Bernoulli-Gaussian sequence with minimum distance constraint for the wave locations and appropriate priors for the amplitudes, wave impulse responses and noise variance. A recently proposed partially collapsed Gibbs sampler which exploits this minimum distance constraint was adapted to the proposed model to estimate the unknown parameters. To the best of our knowledge, this is the first application of MCMC methods to P and T wave analysis of ECG signals. Then, a modified version of this Bayesian model was proposed to consider the baseline within each non-QRS component and to represent P and T waves by their respective dimensionality reducing expansion according to Hermite basis functions. The local dependency of the ECG signal was expressed by a block constraint. To alleviate numerical problems related to the modified Bayesian model, a block Gibbs sampler was studied. The proposed PCGS and block Gibbs sampler overcame the slow convergence problem encountered with the classical Gibbs sampler. The resulting algorithms were validated using the entire annotated QT database. A comparison with other benchmark methods showed that the proposed Bayesian methods provide a reliable detection and an accurate delineation for a wide variety of wave morphologies. In addition, the proposed Bayesian methods can provide accurate waveform estimation and allow for the determination

of confidence intervals which indicate reliability information about the estimates.

The third chapter introduced a modified Bayesian model that enables P and T wave delineation and waveform estimation on a beat-to-beat basis. The beat-to-beat Bayesian model used the P and T waveform estimates of the previous beat as prior information for detecting/estimating the current P and T waves. Compared to the window based model which relies on a non-overlapped multiple-beat processing window to estimate the waveforms, the beat-to-beat model is more suitable to the pseudo-stationary nature of the ECG signal and to real time applications. The block Gibbs sampler was used to estimate the parameters of the resulting Bayesian model. Then, in order to consider all the available former beats instead of using the last beat only, a SMC method was studied. Following the SMC analysis principle, the sequential nature of the ECG was exploited by using a dynamic model under a similar Bayesian framework. Particle filters (PFs) were then proposed to resolve the unknown parameters of the dynamic model. Qualitative and quantitative comparisons with the window based Bayesian models showed that the beat-to-beat Bayesian approaches provide better detection and delineation results and they are more suitable to on-line applications. Furthermore, compared to the recently proposed Gaussian mixture model and extended Kalman filter method (refer to Section 1.4.3 and [SS09] for details), the proposed beat-to-beat Bayesian models provide better results when dealing with pathology signals and they allow a good simultaneous waveform estimation.

The last chapter presented applications of the different Bayesian models introduced in this thesis to a specific clinical research topic referred to as T wave alternans (TWA) detection. First, the window based Bayesian model introduced in Chapter 2 is modified to account for a possible distinction between odd and even beats since a difference between successive T waves is a sign of a potential cardiac risk. The odd and even T-wave amplitudes generated by the block Gibbs sampler are used to build statistical tests for TWA detection. Different from the state-of-art statistical tests for TWA detection, the proposed method computes multiple test statistics for each observation window that can be used advantageously to derive detection performance and to provide information about the reliability of the detection. The comparisons with other classical algorithms on real surface ECGs subjected to synthetic TWA showed good detection performance of the proposed method. In the second part of the chapter, the TWA detection problem on intracardiac electrogram (EGM) signals stored in implantable cardioverter defibrillator (ICD) memories is addressed by using the beat-to-beat based Bayesian model introduced in Chapter 3. Based on the good waveform estimation provided by the beat-to-beat Bayesian approach, ten parameters indicating the whole repolarisation defined by a cardiologist are extracted. Dimensionality reduction techniques and statistical tests are then employed to exploit these parameters to analyze beat-to-beat T wave variations. The proposed approach served as a post-hoc analysis tool in the Endocardial T wave Alternans Study (ETWAS) project in collaboration with ST. JUDE MEDICAL, INC and Toulouse Rangueil Hospital, which aims to prospectively assess the feasibility of TWA detection in EGM stored in ICD memories.

Some possible perspectives are illustrated as follows. In this work, the proposed statistical models and associated Bayesian methods were applied on a specific pathology analysis problem: TWA detection. As a perspective, the proposed Bayesian methods can serve as preprocessing tools for numerous other P and T wave pathology analysis problems such as arrhythmia detection [EG96] and P wave morphology classification [CJO01, DG02]. Furthermore, we focused

ourselves on single-lead ECG processing in this thesis. Despite that most of the real-time Holter monitoring systems and the ICD devices provide only single-lead recordings, multi-lead surface ECG signal (typically the 12-lead ECG) recordings are sometimes available in clinical or ambulance monitoring (see Chapter 1). Logically, another interesting perspective of this work is to include multi-lead ECG analysis in a similar Bayesian framework. This can be realized by including the correlation between parameters estimated for each individual lead (i.e., wave indicators) into the Bayesian model. A similar technique has already been proposed in [DTS07] to deal with astronomical time series. Since multi-lead recordings contain spatial characteristics of the different available leads, a more robust delineation and waveform analysis result can be expected.

Appendix

Appendix A

Sampling distributions for the PCGS

This appendix derives the expressions of the sampling distributions provided in Section 2.3.2.

Wave Indicators. The sampling distribution for $\boldsymbol{b}_{J_d(k)}$ can be obtained as follows

$$\begin{aligned}
&p\left(\boldsymbol{b}_{J_d(k)}|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_n^2, \boldsymbol{x}\right) \\
&\quad \propto p\left(\boldsymbol{b}_{J_d(k)}, \boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_n^2, \boldsymbol{x}\right) \\
&\quad \propto p\left(\boldsymbol{x}|\boldsymbol{b}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_n^2\right) p\left(\boldsymbol{a}|\boldsymbol{b}\right) p\left(\boldsymbol{b}\right) \\
&\quad \propto \left[\int p\left(\boldsymbol{x}|\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{h}, \sigma_n^2\right) p\left(\boldsymbol{a}_{J_d(k)}|\boldsymbol{b}_{J_d(k)}\right) \mathrm{d}\boldsymbol{a}_{J_d(k)}\right] p\left(\boldsymbol{b}\right)
\end{aligned}$$

Using the minimum-distance constraint, there can only be one non-zero wave indicator within the neighborhood $J_{d(k)}$. Let $k^{'} \in J_{d(k)}$ denote this only non-zero indicator location, $\boldsymbol{b}_{J_d(k)}$ can be seen as two parts

$$\begin{cases} b_{k'} = 1 \\ b_m = 0, m \in J_d\left(k\right) \setminus \left\{k^{'}\right\} \end{cases}$$

where $J_d\left(k\right) \setminus \left\{k^{'}\right\}$ denotes the set of locations within the neighborhood $J_{d(k)}$ excluding $k^{'}$. The conditional distribution can be further developed by inserting all the prior distributions

$$\begin{aligned}
&p\left(\boldsymbol{b}_{J_d(k)}|\boldsymbol{b}_{\sim J_d(k)}, \boldsymbol{a}_{\sim J_d(k)}, \boldsymbol{h}, \sigma_n^2, \boldsymbol{x}\right) \\
&\quad \propto \int \exp\left[-\frac{1}{2\sigma_n^2}\Big\| \underbrace{\boldsymbol{x} - \boldsymbol{F}_{\sim J_d(k)}\boldsymbol{B}_{\sim J_d(k)}\boldsymbol{a}_{\sim J_d(k)}}_{\tilde{\boldsymbol{x}}} \right. \\
&\quad \left. -\boldsymbol{F}_{J_d(k)}\boldsymbol{B}_{J_d(k)}\boldsymbol{a}_{J_d(k)}\Big\|^2 - \frac{a_{k'}}{2\sigma_a^2}\right] \prod_{m \in J_d(k) \setminus k'} \delta\left(a_m\right) \mathrm{d}\boldsymbol{a}_{J_d(k)} p\left(\boldsymbol{b}\right)
\end{aligned}$$

$$
\begin{aligned}
&\propto \left[\int \exp\left[-\frac{1}{2\sigma_n^2}\left\|\tilde{\boldsymbol{x}}-f_{k'}a_{k'}\right\|^2-\frac{a_{k'}}{2\sigma_a^2}\right]\mathrm{d}a_{k'}\right]p\left(\boldsymbol{b}\right)\\
&\propto \left[\int \exp\left[-\frac{1}{2}\Bigg(a_{k'}\underbrace{\left(\frac{f_{k'}^Tf_{k'}}{\sigma_n^2}+\frac{1}{\sigma_a^2}\right)}_{\frac{1}{\sigma_1^2}}a_{k'}-a_{k'}\underbrace{\frac{f_{k'}^T\tilde{\boldsymbol{x}}}{\sigma_n^2}}_{\frac{\mu_1}{\sigma_1^2}}-\frac{\tilde{\boldsymbol{x}}^Tf_{k'}}{\sigma_n^2}a_{k'}\Bigg)\right]\mathrm{d}a_{k'}\right]p\left(\boldsymbol{b}\right)\\
&\propto \left[\int \exp\left[-\frac{\left(a_{k'}-\mu_1\right)^2}{2\sigma_1^2}+\frac{\mu_1^2}{2\sigma_1^2}\right]\mathrm{d}a_{k'}\right]p\left(\boldsymbol{b}\right)\\
&\propto \sigma_1^2\exp\left[\frac{\mu_1^2}{2\sigma_1^2}\right]p\left(\boldsymbol{b}\right)
\end{aligned}
$$

where σ_1^2 and μ_1 contain information about $\boldsymbol{b}_{J_d(k)}$ and are defined as

$$
\begin{aligned}
\sigma_1^2 &= \left(\frac{\left\|f_{k'}\right\|^2}{\sigma_n^2}+\frac{1}{\sigma_a^2}\right)^{-1}=\left(\frac{\left\|\boldsymbol{F}_{J_d(k)}\boldsymbol{b}_{J_d(k)}\right\|^2}{\sigma_n^2}+\frac{1}{\sigma_a^2}\right)^{-1}\\
\mu_1 &= \frac{\sigma_1^2f_{k'}^T\tilde{\boldsymbol{x}}}{\sigma_n^2}=\frac{\sigma_1^2\boldsymbol{b}_{J_d(k)}^T\boldsymbol{F}_{J_d(k)}^T\tilde{\boldsymbol{x}}}{\sigma_n^2}.
\end{aligned}
$$

Wave Amplitudes. The sampling distribution for a_k can be obtained as follows

$$
\begin{aligned}
&p\left(a_k|b_k=1,\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto\int p\left(\boldsymbol{a}_{J_d(k)}|\boldsymbol{b},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_n^2,\boldsymbol{x}\right)\mathrm{d}\boldsymbol{a}_{J_d(k)\backslash k}\\
&\quad\propto\int p\left(\boldsymbol{x}|\boldsymbol{b},\boldsymbol{a},\boldsymbol{h},\sigma_n^2\right)p\left(\boldsymbol{a}_{J_d(k)}|\boldsymbol{b}_{J_d(k)}\right)\mathrm{d}\boldsymbol{a}_{J_d(k)\backslash k}
\end{aligned}
$$

Consequently, similar to the conditional distribution of $\boldsymbol{b}_{J_d(k)}$, the following results can be obtained

$$
\begin{aligned}
&p\left(a_k|b_k=1,\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto\int\exp\left[-\frac{1}{2\sigma_n^2}\left\|\tilde{\boldsymbol{x}}-\boldsymbol{F}_{J_d(k)}\boldsymbol{B}_{J_d(k)}\boldsymbol{a}_{J_d(k)}\right\|^2+\frac{a_k}{2\sigma_a^2}\right]\prod_{m\in J_d(k)\backslash k}\delta\left(a_m\right)\mathrm{d}\boldsymbol{a}_{J_d(k)\backslash k}\\
&\quad\propto\exp\left[-\frac{1}{2\sigma_n^2}\left\|\tilde{\boldsymbol{x}}-f_ka_k\right\|^2+\frac{a_k}{2\sigma_a^2}\right]\\
&\quad\propto\exp\left[-\frac{1}{2\sigma_1^2}\left(a_k-\mu_1\right)^2\right]
\end{aligned}
$$

that can be summarized as

$$
p\left(a_k|b_k=1,\boldsymbol{b}_{\sim J_d(k)},\boldsymbol{a}_{\sim J_d(k)},\boldsymbol{h},\sigma_n^2,\boldsymbol{x}\right)=\mathcal{N}\left(\mu_1,\sigma_1^2\right).
$$

Waveform coefficients. The sampling distribution for $\boldsymbol{h}$ can be obtained as follows

$$\begin{aligned} p\left(\boldsymbol{h}|\boldsymbol{b},\boldsymbol{a},\sigma_n^2,\boldsymbol{x}\right) &\propto p\left(\boldsymbol{x}|\boldsymbol{b},\boldsymbol{a},\boldsymbol{h},\sigma_n^2\right)p\left(\boldsymbol{h}\right) \\ &\propto \exp\left[-\frac{1}{2\sigma_n^2}\left\|\boldsymbol{x}-\boldsymbol{U}\boldsymbol{h}\right\|^2\right]\exp\left[-\frac{1}{2\sigma_h^2}\left\|\boldsymbol{h}\right\|^2\right] \\ &\propto \exp\left[-\frac{1}{2\sigma_n^2}\left[\left\|\boldsymbol{U}\boldsymbol{h}\right\|^2-2\boldsymbol{x}\boldsymbol{U}\boldsymbol{h}\right]\right]\exp\left[-\frac{1}{2\sigma_h^2}\left\|\boldsymbol{h}\right\|^2\right] \\ &\propto \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_n^2}\left\|\boldsymbol{U}\right\|^2+\frac{1}{\sigma_h^2}\right)\left\|\boldsymbol{h}\right\|^2+\frac{1}{\sigma_n^2}\boldsymbol{x}\boldsymbol{U}\boldsymbol{h}\right] \end{aligned}$$

Equivalently,

$$p\left(\boldsymbol{h}|\boldsymbol{b},\boldsymbol{a},\sigma_n^2,\boldsymbol{x}\right)=\mathcal{N}\left(\boldsymbol{\mu}_2,\boldsymbol{\sigma}_2^2\right)$$

with

$$\boldsymbol{\mu}_2=\frac{\boldsymbol{\sigma}_2^2\boldsymbol{U}^T\boldsymbol{x}}{\sigma_n^2},\quad \boldsymbol{\sigma}_2^2=\left(\frac{\boldsymbol{U}^T\boldsymbol{U}}{\sigma_n^2}+\frac{\boldsymbol{I}_{L+1}}{\sigma_h^2}\right)^{-1}.$$

Appendix B

Sampling distributions for the block Gibbs sampler

This appendix derives the expressions of the sampling distributions provided in Section 2.4.3.

T wave Indicators. The sampling distribution for T-wave Indicators $p\left(\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}|\boldsymbol{\theta}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{\theta}_{\mathrm{P}},\boldsymbol{\theta}_{\mathrm{cw}},\boldsymbol{x}\right)$ can be obtained as follows

$$\begin{aligned}
&p\left(\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}|\boldsymbol{b}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto p\left(\boldsymbol{x}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2\right)p\left(\boldsymbol{a}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}}\right)p\left(\boldsymbol{b}_{\mathrm{T}}\right)\\
&\quad\propto\left[\int p\left(\boldsymbol{x}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2\right)p\left(\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}}|\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}\right)\mathrm{d}\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}}\right]p\left(\boldsymbol{b}_{\mathrm{T}}\right)
\end{aligned}$$

Using the block constraint, there can only be one non-zero wave indicator within the block $J_{\mathrm{T},n}$. Let $k^{'}\in J_{\mathrm{T},n}$ denote this only non-zero indicator location, $\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}$ can be seen as two parts

$$\begin{cases} b_{k'}=1\\ b_m=0, m\in J_{\mathrm{T},n}\setminus\left\{k^{'}\right\}\end{cases}$$

where $J_{\mathrm{T},n}\setminus\left\{k^{'}\right\}$ denotes the set of locations within the block $J_{\mathrm{T},n}$ excluding $k^{'}$. The conditional distribution can be further developed by inserting all the prior distributions

$$\begin{aligned}
&p\left(\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}|\boldsymbol{b}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto\int\exp\Bigg[-\frac{1}{2\sigma_n^2}\Big\|\underbrace{\boldsymbol{x}-\boldsymbol{F}_{\mathrm{T},\sim J_{\mathrm{T},n}}\boldsymbol{B}_{\mathrm{T},\sim J_{\mathrm{T},n}}\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}}-\boldsymbol{F}_{\mathrm{P}}\boldsymbol{B}_{\mathrm{P}}\boldsymbol{a}_{\mathrm{P}}-\boldsymbol{M}\boldsymbol{\gamma}}_{\tilde{\boldsymbol{x}}}\\
&\quad-\boldsymbol{F}_{\mathrm{T},J_{\mathrm{T},n}}\boldsymbol{B}_{\mathrm{T},J_{\mathrm{T},n}}\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}}\Big\|^2-\frac{a_{\mathrm{T},k'}}{2\sigma_a^2}\Bigg]\prod_{m\in J_{\mathrm{T},n}\setminus k'}\delta\left(a_{\mathrm{T},m}\right)\mathrm{d}\boldsymbol{a}_{J_{\mathrm{T},n}}\;p\left(\boldsymbol{b}_{\mathrm{T}}\right)
\end{aligned}$$

$$
\begin{aligned}
&\propto \left[\int \exp\left[-\frac{1}{2\sigma_n^2}\left\|\tilde{\boldsymbol{x}} - f_{\mathrm{T},k'}a_{\mathrm{T},k'}\right\|^2 - \frac{a_{\mathrm{T},k'}}{2\sigma_a^2}\right]\mathrm{d}a_{\mathrm{T},k'}\right]p\left(\boldsymbol{b}_{\mathrm{T}}\right)\\
&\propto \left[\int \exp\left[-\frac{1}{2}\Bigg(a_{\mathrm{T},k'}\underbrace{\left(\frac{f_{\mathrm{T},k'}^T f_{\mathrm{T},k'}}{\sigma_n^2}+\frac{1}{\sigma_a^2}\right)}_{\frac{1}{\sigma_1^2}}a_{\mathrm{T},k'} - a_{\mathrm{T},k'}\underbrace{\frac{f_{\mathrm{T},k'}^T\tilde{\boldsymbol{x}}}{\sigma_n^2}}_{\frac{\mu_1}{\sigma_1^2}} - \frac{\tilde{\boldsymbol{x}}^T f_{\mathrm{T},k'}}{\sigma_n^2}a_{\mathrm{T},k'}\Bigg)\right]\mathrm{d}a_{\mathrm{T},k'}\right]p\left(\boldsymbol{b}_{\mathrm{T}}\right)\\
&\propto \left[\int \exp\left[-\frac{\left(a_{\mathrm{T},k'}-\mu_1\right)^2}{2\sigma_1^2}+\frac{\mu_1^2}{2\sigma_1^2}\right]\mathrm{d}a_{\mathrm{T},k'}\right]p\left(\boldsymbol{b}_{\mathrm{T}}\right)\\
&\propto \sigma_1\exp\left[\frac{\mu_1^2}{2\sigma_1^2}\right]p\left(\boldsymbol{b}_{\mathrm{T}}\right)
\end{aligned}
$$

where σ_1^2 and μ_1 contain information about $\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}$ and are defined as

$$
\begin{aligned}
\sigma_1^2 &= \left(\frac{\left\|f_{\mathrm{T},k'}\right\|^2}{\sigma_n^2}+\frac{1}{\sigma_a^2}\right)^{-1} = \left(\frac{\left\|\boldsymbol{F}_{\mathrm{T},J_{\mathrm{T},n}}\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}\right\|^2}{\sigma_n^2}+\frac{1}{\sigma_a^2}\right)^{-1}\\
\mu_1 &= \frac{\sigma_1^2 f_{\mathrm{T},k'}^T\tilde{\boldsymbol{x}}}{\sigma_n^2} = \frac{\sigma_1^2\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}^T\boldsymbol{F}_{\mathrm{T},J_{\mathrm{T},n}}^T\left(\boldsymbol{x}-\boldsymbol{F}_{\mathrm{T},\sim J_{\mathrm{T},n}}\boldsymbol{B}_{\mathrm{T},\sim J_{\mathrm{T},n}}\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}}-\boldsymbol{F}_{\mathrm{P}}\boldsymbol{B}_{\mathrm{P}}\boldsymbol{a}_{\mathrm{P}}-\boldsymbol{M}\boldsymbol{\gamma}\right)}{\sigma_n^2}.
\end{aligned}
$$

$\boldsymbol{F}_{\mathrm{T},J_{\mathrm{T},n}}$ denotes the columns of $\boldsymbol{F}_{\mathrm{T}}$ indexed by $J_{\mathrm{T},n}$, $\boldsymbol{F}_{\mathrm{T},\sim J_{\mathrm{T},n}}$ denotes $\boldsymbol{F}_{\mathrm{T}}$ without those columns, and $\boldsymbol{B}_{\mathrm{T},\sim J_{\mathrm{T},n}}$ denotes the diagonal matrix diag $\left(\boldsymbol{b}_{\mathrm{T},\sim J_{\mathrm{T},n}}\right)$.

T-wave Amplitudes. The sampling distribution for $a_{\mathrm{T},k}$ can be obtained as follows

$$
\begin{aligned}
&p\left(a_{\mathrm{T},k}|b_{\mathrm{T},k}=1,\boldsymbol{b}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto \int p\left(\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},,\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)\mathrm{d}\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}\backslash k}\\
&\quad\propto \int p\left(\boldsymbol{x}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2\right)p\left(\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}}|\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}\right)\mathrm{d}\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}\backslash k}
\end{aligned}
$$

Consequently, similar to the conditional distribution of $\boldsymbol{b}_{\mathrm{T},J_{\mathrm{T},n}}$, the following results can be obtained

$$
\begin{aligned}
&p\left(a_{\mathrm{T},k}|b_{\mathrm{T},k}=1,\boldsymbol{b}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto \int \exp\left[-\frac{1}{2\sigma_n^2}\left\|\tilde{\boldsymbol{x}}-\boldsymbol{F}_{\mathrm{T},J_{\mathrm{T},n}}\boldsymbol{B}_{\mathrm{T},J_{\mathrm{T},n}}\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}}\right\|^2+\frac{a_{\mathrm{T},k}}{2\sigma_a^2}\right]\prod_{m\in J_{\mathrm{T},n}\backslash k}\delta\left(a_{\mathrm{T},m}\right)\mathrm{d}\boldsymbol{a}_{\mathrm{T},J_{\mathrm{T},n}\backslash k}
\end{aligned}
$$

$$\propto \exp\left[-\frac{1}{2\sigma_n^2}\left\|\tilde{\boldsymbol{x}}-f_{\mathrm{T},k}a_{\mathrm{T},k}\right\|^2+\frac{a_{\mathrm{T},k}}{2\sigma_a^2}\right]$$
$$\propto \exp\left[-\frac{1}{2\sigma_1^2}\left(a_{\mathrm{T},k}-\mu_1\right)^2\right]$$

that can be summarized as

$$p\left(a_{\mathrm{T},k}|b_{\mathrm{T},k}=1,\boldsymbol{b}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{a}_{\mathrm{T},\sim J_{\mathrm{T},n}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)=\mathcal{N}\left(\mu_1,\sigma_1^2\right).$$

T-Waveform coefficients. The sampling distribution for T-waveform coefficients $\boldsymbol{\alpha}_{\mathrm{T}}$ can be obtained as follows

$$\begin{aligned}
&p\left(\boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto p\left(\boldsymbol{x}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2\right)p\left(\boldsymbol{\alpha}_{\mathrm{T}}\right)\\
&\quad\propto \exp\left[-\frac{1}{2\sigma_n^2}\left\|\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{M}\boldsymbol{\gamma}\right)-\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right\|^2\right]\exp\left[-\frac{1}{2\sigma_\alpha^2}\left\|\boldsymbol{\alpha}_{\mathrm{T}}\right\|^2\right]\\
&\quad\propto \exp\left[-\frac{1}{2\sigma_n^2}\left[\left\|\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right\|^2-2\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{M}\boldsymbol{\gamma}\right)\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right]\right]\exp\left[-\frac{1}{2\sigma_\alpha^2}\left\|\boldsymbol{\alpha}_{\mathrm{T}}\right\|^2\right]\\
&\quad\propto \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_n^2}\left\|\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\right\|^2+\frac{\boldsymbol{I}_{L+1}}{\sigma_\alpha^2}\right)\left\|\boldsymbol{\alpha}_{\mathrm{T}}\right\|^2+\frac{1}{\sigma_n^2}\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{M}\boldsymbol{\gamma}\right)\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right]
\end{aligned}$$

Equivalently,

$$p\left(\boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2,\boldsymbol{x}\right)=\mathcal{N}\left(\boldsymbol{\mu}_3,\boldsymbol{\sigma}_3^2\right)$$

with

$$\boldsymbol{\mu}_3=\frac{\boldsymbol{\sigma}_3^2\boldsymbol{H}^T\boldsymbol{U}_{\mathrm{T}}^T\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{M}\boldsymbol{\gamma}\right)}{\sigma_n^2},\quad \boldsymbol{\sigma}_3^2=\left(\frac{\left\|\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\right\|^2}{\sigma_n^2}+\frac{\boldsymbol{I}_{L+1}}{\sigma_\alpha^2}\right)^{-1}.$$

Local baseline coefficients. Because $\boldsymbol{\gamma}$ is *a priori* Gaussian, it can be treated as one parameter in the Gibbs sampler without introducing excessive complexity. As a consequence, the sampling distribution for local baseline coefficients $p\left(\boldsymbol{\gamma}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\sigma_n^2,\boldsymbol{x}\right)$ can be obtained as follows

$$\begin{aligned}
&p\left(\boldsymbol{\gamma}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\sigma_n^2,\boldsymbol{x}\right)\\
&\quad\propto p\left(\boldsymbol{x}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\boldsymbol{\gamma},\sigma_n^2\right)p\left(\boldsymbol{\gamma}\right)\\
&\quad\propto \exp\left[-\frac{1}{2\sigma_n^2}\left\|\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right)-\boldsymbol{M}\boldsymbol{\gamma}\right\|^2\right]\exp\left[-\frac{1}{2\sigma_\gamma^2}\left\|\boldsymbol{\gamma}\right\|^2\right]\\
&\quad\propto \exp\left[-\frac{1}{2\sigma_n^2}\left[\left\|\boldsymbol{M}\boldsymbol{\gamma}\right\|^2-2\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right)\boldsymbol{M}\boldsymbol{\gamma}\right]\right]\exp\left[-\frac{1}{2\sigma_\gamma^2}\left\|\boldsymbol{\gamma}\right\|^2\right]\\
&\quad\propto \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_n^2}\left\|\boldsymbol{M}\right\|^2+\frac{\boldsymbol{I}_{5\times D}}{\sigma_\gamma^2}\right)\left\|\boldsymbol{\gamma}\right\|^2+\frac{1}{\sigma_n^2}\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right)\boldsymbol{M}\boldsymbol{\gamma}\right]
\end{aligned}$$

Equivalently,

$$p\left(\boldsymbol{\gamma}|\boldsymbol{b}_{\mathrm{T}},\boldsymbol{a}_{\mathrm{T}},\boldsymbol{b}_{\mathrm{P}},\boldsymbol{a}_{\mathrm{P}},\boldsymbol{\alpha}_{\mathrm{T}},\boldsymbol{\alpha}_{\mathrm{P}},\sigma_n^2,\boldsymbol{x}\right)=\mathcal{N}\left(\boldsymbol{\mu}_5,\boldsymbol{\sigma}_5^2\right)$$

with

$$\boldsymbol{\mu}_5=\frac{\boldsymbol{\sigma}_5^2\boldsymbol{M}^T\left(\boldsymbol{x}-\boldsymbol{U}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}-\boldsymbol{U}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\right)}{\sigma_n^2},\quad \boldsymbol{\sigma}_5^2=\left(\frac{\|\boldsymbol{M}\|^2}{\sigma_n^2}+\frac{\boldsymbol{I}_{5\times D}}{\sigma_\gamma^2}\right)^{-1}.$$

Appendix C

Sampling distributions for the beat-to-beat block Gibbs sampler

This appendix derives the expressions of the sampling distributions provided in Section 3.2.3. We recall the definitions $\boldsymbol{\theta} \triangleq (\boldsymbol{b}_{\mathrm{T}}^T\, \boldsymbol{b}_{\mathrm{P}}^T\, \boldsymbol{\alpha}_{\mathrm{T}}^T\, \boldsymbol{\alpha}_{\mathrm{P}}^T\, \boldsymbol{\gamma}^T\, \sigma_w^2)^T$, $\tilde{\boldsymbol{x}} \triangleq \boldsymbol{x} - \boldsymbol{B}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}} - \boldsymbol{B}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}}$, and $\tilde{\boldsymbol{x}}_{\mathrm{T}} \triangleq \boldsymbol{x} - \boldsymbol{B}_{\mathrm{P}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{P}} - \boldsymbol{M}\boldsymbol{\gamma}$. As before, the subscript n will be suppressed.

Wave indicators. To derive expression (3.14) for the sampling distribution for $\boldsymbol{b}_{\mathrm{T}}$, we first note that

$$\begin{aligned}
&p(\boldsymbol{b}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x}) \\
&\quad = \int p(\boldsymbol{b}_{\mathrm{T}}, \boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{P}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_{\mathrm{P}}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x})\, \mathrm{d}\boldsymbol{\alpha}_{\mathrm{T}} \\
&\quad \propto \int p(\boldsymbol{\theta}|\boldsymbol{x}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1})\, \mathrm{d}\boldsymbol{\alpha}_{\mathrm{T}} \\
&\quad \propto \int p(\boldsymbol{x}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1})\, \mathrm{d}\boldsymbol{\alpha}_{\mathrm{T}}
\end{aligned}$$

In the second last and last steps, we used Bayes' rule and (3.13), respectively. Next, we insert (3.6) and (3.12) and drop factors that do not depend on $\boldsymbol{b}_{\mathrm{T}}$ or $\boldsymbol{\alpha}_{\mathrm{T}}$. This yields

$$\begin{aligned}
p(\boldsymbol{b}_{\mathrm{T}}|\ldots) \propto \int \exp\!\left(-\frac{\|\tilde{\boldsymbol{x}}_{\mathrm{T}} - \boldsymbol{B}_{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T}}\|^2}{2\sigma_w^2}\right) \\
\times p(\boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1})\, p(\boldsymbol{b}_{\mathrm{T}})\, \mathrm{d}\boldsymbol{\alpha}_{\mathrm{T}}\,. \qquad \text{(C.1)}
\end{aligned}$$

Now, we insert (3.9) for $p(\boldsymbol{\alpha}_{\mathrm{T}}|\boldsymbol{b}_{\mathrm{T}}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1})$, and (3.7) for $p(\boldsymbol{b}_{\mathrm{T}})$. For the case where $\|\boldsymbol{b}_{\mathrm{T}}\| = 0$ (no T wave in $\mathcal{J}_{\mathrm{T}}$), this yields

$$\begin{aligned}
p(\boldsymbol{b}_{\mathrm{T}}|\ldots) &\propto \exp\left(-\frac{\|\tilde{\boldsymbol{x}}_{\mathrm{T}}\|^2}{2\sigma_w^2}\right)\left[\int \delta(\boldsymbol{\alpha}_{\mathrm{T}} - \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1})\, \mathrm{d}\boldsymbol{\alpha}_{\mathrm{T}}\right] p_0 \\
&= \exp\left(-\frac{\|\tilde{\boldsymbol{x}}_{\mathrm{T}}\|^2}{2\sigma_w^2}\right) p_0\,. \qquad \text{(C.2)}
\end{aligned}$$

For the case where $\|\boldsymbol{b}_\mathrm{T}\| = 1$ (one T wave in $\mathcal{J}_\mathrm{T}$), we obtain

$$\begin{aligned} p(\boldsymbol{b}_\mathrm{T}|\ldots &\propto \frac{1}{(\sqrt{2\pi}\sigma_\alpha)^G}\left[\int \exp\left(-\frac{\|\tilde{\boldsymbol{x}}_\mathrm{T} - \boldsymbol{B}_\mathrm{T}\boldsymbol{H}\boldsymbol{\alpha}_\mathrm{T}\|^2}{2\sigma_w^2} - \frac{\|\boldsymbol{\alpha}_\mathrm{T} - \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}\|^2}{2\sigma_\alpha^2}\right)\mathrm{d}\boldsymbol{\alpha}_\mathrm{T}\right] p_1 \\ &= \frac{1}{(\sqrt{2\pi}\sigma_\alpha)^G}\left[\int \exp\left(-\frac{1}{2}(\boldsymbol{\alpha}_\mathrm{T} - \boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{\alpha}_\mathrm{T} - \boldsymbol{\mu}_1)\right)\mathrm{d}\boldsymbol{\alpha}_\mathrm{T}\right] \exp\left(-\frac{\|\tilde{\boldsymbol{x}}_\mathrm{T}\|^2}{2\sigma_w^2} + \boldsymbol{\mu}_1\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1\right) p_1 \\ &= \frac{\sqrt{|\boldsymbol{\Sigma}_1|}}{\sigma_\alpha^G}\ \exp\left(-\frac{\|\tilde{\boldsymbol{x}}_\mathrm{T}\|^2}{2\sigma_w^2}\right)\ \exp\left(\boldsymbol{\mu}_1\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1\right) p_1 \end{aligned} \tag{C.3}$$

with $\boldsymbol{\mu}_1$ and $\boldsymbol{\Sigma}_1$ as defined in (3.15) and (3.16), respectively. In the third case, $\|\boldsymbol{b}_\mathrm{T}\| > 1$, the prior in (3.7) is zero, hence (C.1) is also zero. Since the constant factor $\exp\left(-\frac{\|\tilde{\boldsymbol{x}}_\mathrm{T}\|^2}{2\sigma_w^2}\right)$ appears in both (C.2) and (C.3), it can be dropped. Thus, we obtain (3.14).

Waveform coefficients. To derive expression (3.17) for the sampling distribution for $\boldsymbol{\alpha}_\mathrm{T}$, we again use Bayes' rule as well as equations (3.13) and (3.12)

$$\begin{aligned} &p(\boldsymbol{\alpha}_\mathrm{T}|\boldsymbol{b}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \boldsymbol{\alpha}_\mathrm{P}, \boldsymbol{\gamma}, \sigma_w^2, \boldsymbol{x}) \\ &\quad\propto p(\boldsymbol{\theta}|\boldsymbol{x}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}) \\ &\quad\propto p(\boldsymbol{x}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}) \\ &\quad\propto\ p(\boldsymbol{x}|\boldsymbol{\theta})\, p(\boldsymbol{\alpha}_\mathrm{T}|\boldsymbol{b}_\mathrm{T}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1})\,. \end{aligned}$$

In the last expression, we dropped all factors that do not depend on $\boldsymbol{\alpha}_\mathrm{T}$. Inserting (3.6) for $p(\boldsymbol{x}|\boldsymbol{\theta})$ and (3.9) for $p(\boldsymbol{\alpha}_\mathrm{T}|\boldsymbol{b}_\mathrm{T}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1})$, we obtain for the case $\|\boldsymbol{b}_\mathrm{T}\| = 0$

$$\begin{aligned} p(\boldsymbol{\alpha}_\mathrm{T}|\ldots) &\propto \exp\left(-\frac{\|\tilde{\boldsymbol{x}}_\mathrm{T}\|^2}{2\sigma_w^2}\right)\delta(\boldsymbol{\alpha}_\mathrm{T} - \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) \\ &\propto\ \delta(\boldsymbol{\alpha}_\mathrm{T} - \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) \end{aligned}$$

and for the case $\|\boldsymbol{b}_\mathrm{T}\| = 1$

$$\begin{aligned} p(\boldsymbol{\alpha}_\mathrm{T}|\ldots) &\propto \frac{1}{(\sqrt{2\pi}\sigma_\alpha)^G}\exp\left(-\frac{\|\tilde{\boldsymbol{x}}_\mathrm{T} - \boldsymbol{B}_\mathrm{T}\boldsymbol{H}\boldsymbol{\alpha}_\mathrm{T}\|^2}{2\sigma_w^2} - \frac{\|\boldsymbol{\alpha}_\mathrm{T} - \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}\|^2}{2\sigma_\alpha^2}\right) \\ &\propto\ \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)\,. \end{aligned}$$

Thus, we have derived (3.17).

Baseline coefficients. The sampling distribution for $\boldsymbol{\gamma}$ in (3.18) is derived in analogy to that of the waveform coefficients

$$\begin{aligned} &p(\boldsymbol{\gamma}|\boldsymbol{b}, \boldsymbol{\alpha}_\mathrm{T}, \boldsymbol{\alpha}_\mathrm{P}, \sigma_w^2, \boldsymbol{x}) \\ &\quad\propto p(\boldsymbol{\theta}|\boldsymbol{x}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}) \\ &\quad\propto p(\boldsymbol{x}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \hat{\boldsymbol{\alpha}}_{\mathrm{P},n-1}) \\ &\quad\propto\ p(\boldsymbol{x}|\boldsymbol{\theta})\, p(\boldsymbol{\gamma})\,. \end{aligned}$$

Inserting (2.5) for $p(\boldsymbol{x}|\boldsymbol{\theta})$ and $p(\boldsymbol{\gamma}) = \mathcal{N}(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}_5)$, we obtain

$$\begin{aligned} p(\boldsymbol{\gamma}|\ldots) &\propto \frac{1}{(\sqrt{2\pi}\sigma_\gamma)^5} \exp\left(-\frac{\|\tilde{\boldsymbol{x}} - \boldsymbol{M}\boldsymbol{\gamma}\|^2}{2\sigma_w^2} - \frac{\|\boldsymbol{\gamma}\|^2}{2\sigma_\gamma^2}\right) \\ &\propto \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \end{aligned}$$

with $\boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}_2$ as defined in (3.19) and (3.20), respectively. Thus, we have derived (3.18).

Noise variance. The distribution in (3.21) is a well-known result (e.g., [LI06]).

Appendix D

Proposal distributions for the CPF

This appendix derives the expressions of the sampling distributions provided in Section 3.3.2.

Wave indicators. The proposal distribution for $\boldsymbol{b}_n$ can be obtained as follows

$$\begin{aligned} p(\boldsymbol{b}_n = \boldsymbol{\beta}_j | \boldsymbol{\alpha}_{n-1}, \boldsymbol{y}_n) &\propto p(\boldsymbol{y}_n | \boldsymbol{b}_n = \boldsymbol{\beta}_j, \boldsymbol{\alpha}_{n-1}) p(\boldsymbol{b}_n = \boldsymbol{\beta}_j) \\ &\propto p(\boldsymbol{y}_n | \boldsymbol{b}_n = \boldsymbol{\beta}_j, \boldsymbol{\alpha}_{n-1}) \\ &\propto \int p(\boldsymbol{y}_n | \boldsymbol{b}_n = \boldsymbol{\beta}_j, \boldsymbol{\alpha}_n) p(\boldsymbol{\alpha}_n | \boldsymbol{\alpha}_{n-1}) \mathrm{d}\boldsymbol{\alpha}_n \\ &\propto \int \exp\left(-\frac{\|\boldsymbol{y}_n - \boldsymbol{B}_{n,j} \boldsymbol{H} \boldsymbol{\alpha}_n\|^2}{2\sigma_w^2} \right) \exp\left(-\frac{\|\boldsymbol{\alpha}_n - \boldsymbol{\alpha}_{n-1}\|^2}{2\sigma_\alpha^2} \right) \mathrm{d}\boldsymbol{\alpha}_n \\ &\propto \exp\left(\boldsymbol{\mu}_{1,j} \boldsymbol{\Sigma}_{1,j}^{-1} \boldsymbol{\mu}_{1,j} \right) \end{aligned}$$

with

$$\begin{aligned} \boldsymbol{\mu}_{1,j} &= \boldsymbol{\Sigma}_{1,j} \left(\frac{\boldsymbol{H}^T \boldsymbol{B}_{n,j}^T \boldsymbol{y}_n}{\sigma_w^2} + \frac{\boldsymbol{\alpha}_{n-1}}{\sigma_\alpha^2} \right) \\ \boldsymbol{\Sigma}_{1,j} &= \left(\frac{\boldsymbol{H}^T \boldsymbol{B}_{n,j}^T \boldsymbol{B}_{n,j} \boldsymbol{H}}{\sigma_w^2} + \frac{\mathbf{I}_G}{\sigma_\alpha^2} \right)^{-1} \end{aligned}$$

where $\boldsymbol{B}_{n,j}$ is the $N_{\mathrm{T},n} \times (2L+1)$ Toeplitz matrix with first row $(b_{n,L+1} \cdots b_{n,1}\ 0 \cdots 0)$ and first column $(b_{n,L+1} \cdots b_{n,N_{\mathrm{T},n}}\ \ 0 \cdots 0)^T$ which corresponds to $\boldsymbol{b}_n = \boldsymbol{\beta}_j$.

Waveform coefficients. The proposal distribution for $\boldsymbol{\alpha}_n$ can be obtained as follows

$$p(\boldsymbol{\alpha}_n | \boldsymbol{b}_n, \boldsymbol{\alpha}_{n-1}, \boldsymbol{y}_n) \propto p(\boldsymbol{y}_n | \boldsymbol{b}_n, \boldsymbol{\alpha}_n) p(\boldsymbol{\alpha}_n | \boldsymbol{\alpha}_{n-1})$$

$$\begin{aligned} &\propto \exp\left(-\frac{\|\boldsymbol{y}_n - \boldsymbol{B}_n \boldsymbol{H} \boldsymbol{\alpha}_n\|^2}{2\sigma_w^2} \right) \exp\left(-\frac{\|\boldsymbol{\alpha}_n - \boldsymbol{\alpha}_{n-1}\|^2}{2\sigma_\alpha^2} \right) \\ &= \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \end{aligned}$$

with

$$\boldsymbol{\mu}_2 = \boldsymbol{\Sigma}_2 \left(\frac{\boldsymbol{H}^T \left(\boldsymbol{B}_n\right)^T \boldsymbol{y}_n}{\sigma_w^2} + \frac{\boldsymbol{\alpha}_{n-1}}{\sigma_\alpha^2} \right)$$

$$\boldsymbol{\Sigma}_2 = \left(\frac{\boldsymbol{H}^T \left(\boldsymbol{B}_n\right)^T \boldsymbol{B}_n \boldsymbol{H}}{\sigma_w^2} + \frac{\boldsymbol{I}_G}{\sigma_\alpha^2} \right)^{-1}$$

Appendix E

T waveform estimation using a beat-to-beat block Gibbs sampler

This appendix introduces the modified beat-to-beat Bayesian model which allows a T waveform estimation in each heartbeat in EGM signals.

Signal model for one T-wave search interval. The proposed method first detects QRS complexes and then defines T search intervals which is located to the right hand neighborhoods of the QRS ends. As shown in Fig. E.1, the T search interval $\mathcal{J}_{\mathrm{T},n}$ associated with the nth heartbeat is defined as the right hand neighborhood of the QRS end set. The temporal lengths of the non-QRS interval $\mathcal{J}_n$ and T search interval $\mathcal{J}_{\mathrm{T},n}$ are denoted as N_n and $N_{\mathrm{T},n}$, respectively. In this work, a value of $N_{\mathrm{T},n} = 0.6N_n$ is recommended by the cardiologist. Our goal is to estimate the locations, amplitudes, and shapes (waveforms) of the T waves within their search intervals $\mathcal{J}_{\mathrm{T},n}$. Note that unlike in Chapters 2 and 3, the P wave is no longer the concern of the problem. Thus the P search interval is not considered in the signal model.

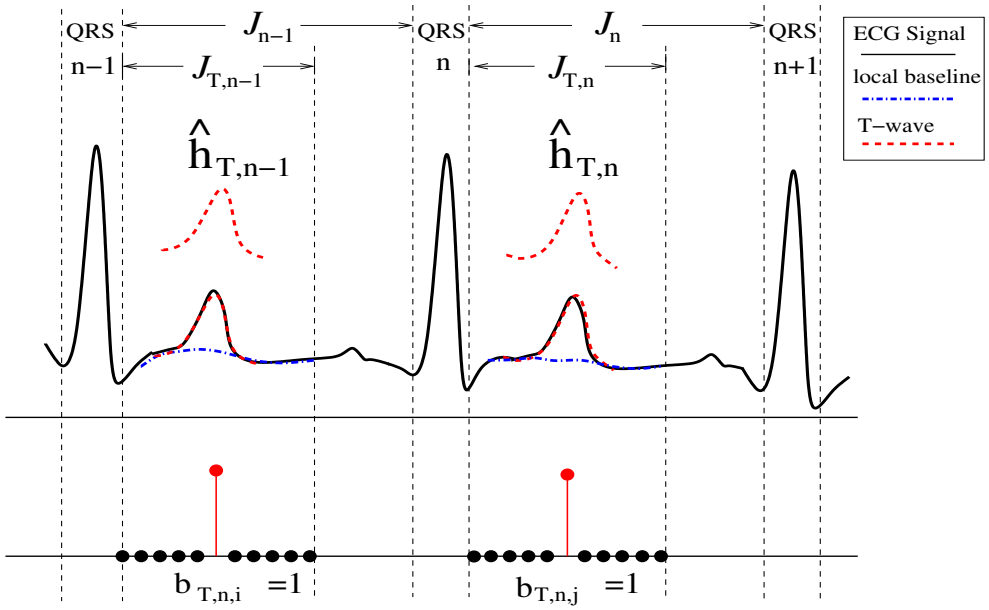


Figure E.1: Signal model for the beat-to-beat T waveform estimation scheme.

Similar to the signal model in Section 3.2, the signal in the T search interval $\mathcal{J}_{\mathrm{T},n}$ can be approximated by a pulse representing the T wave plus a local baseline (see Fig. E.1). The T wave within $\mathcal{J}_{\mathrm{T},n}$ is modeled by the convolution of an unknown binary "indicator sequence" $\boldsymbol{b}_{\mathrm{T},n} = (b_{\mathrm{T},n,1} \cdots b_{\mathrm{T},n,N_{\mathrm{T},n}})^T$ indicating the T wave location ($b_{\mathrm{T},n,i} = 1$ if there is a wave at location i, $b_{\mathrm{T},n,i} = 0$ otherwise) with unknown T waveforms $\boldsymbol{h}_{\mathrm{T},n} = (h_{\mathrm{T},n,-L} \cdots h_{\mathrm{T},n,L})^T$. Here, the waveform length $2L+1$ is chosen as a fixed percentage of $N_{\mathrm{T},n}$ that is large enough to accommodate the actual supports of the T waves. Note that at most one indicator is nonzero because at most one T wave maybe present in any given search interval.

From the above discussion, it follows that the nth T search interval can be expressed by the convolution relation

$$x_{n,k} = \sum_{j=1}^{N_{\mathrm{T},n}} h_{\mathrm{T},n,k-j} b_{\mathrm{T},n,j} + c_{n,k} + w_{n,k}\,, \quad k \in \mathcal{J}_{\mathrm{T},n}. \tag{E.1}$$

Here, $c_{n,k}$ denotes the baseline sequence and $w_{n,k}$ is a white Gaussian noise with unknown variance $\sigma^2_{w,n}$.

By applying the Hermite decomposition technique as in the previous sections, the T waveform is represented by a basis expansion using discrete-time versions of Hermite functions such that

$$\boldsymbol{h}_{\mathrm{T},n} = \boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n} \tag{E.2}$$

where $\boldsymbol{H}$ is a $(2L+1) \times G$ matrix whose columns are the first G Hermite functions (with $G \leq 2L+1$), suitably sampled and truncated to length $2L+1$, and $\boldsymbol{\alpha}_{\mathrm{T},n}$ is an unknown coefficient vector of length G. Similarly, the residual local baseline within the T wave search interval is modeled by using a polynomial of degree 4, i.e.,

$$c_{n,k} = \sum_{i=1}^{5} \gamma_{n,i} k^{i-1}, \quad k \in \mathcal{J}_n\,. \tag{E.3}$$

In vector-matrix form, (E.3) reads as $\boldsymbol{c}_n = \boldsymbol{M}\boldsymbol{\gamma}_n$, with the known $N_{\mathrm{T},n} \times 5$ Vandermonde matrix $\boldsymbol{M}$ and the unknown coefficient vector $\boldsymbol{\gamma}_n = (\gamma_{n,1} \cdots \gamma_{n,5})^T$.

By applying (E.2) and (E.3), we obtain the following vector representation of the T search interval signal in (E.1)

$$\boldsymbol{x}_n = \boldsymbol{B}_{\mathrm{T},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n} + \boldsymbol{M}\boldsymbol{\gamma}_n + \boldsymbol{w}_n \tag{E.4}$$

where $\boldsymbol{x}_n = (x_{n,1} \cdots x_{n,N_{\mathrm{T},n}})^T$ and $\boldsymbol{w}_n = (w_{n,1} \cdots w_{n,N_{\mathrm{T},n}})^T$ and $\boldsymbol{B}_{\mathrm{T},n}$ is the $N_{\mathrm{T},n} \times (2L+1)$ Toeplitz matrix with first row $(b_{\mathrm{T},n,L+1} \cdots b_{\mathrm{T},n,1}\ 0 \cdots 0)$ and first column $(b_{\mathrm{T},n,L+1} \cdots b_{\mathrm{T},n,N_{\mathrm{T},n}}\ 0 \cdots 0)^T$.

Bayesian inference for T wave parameter estimation

Due to the previous parametrization, the unknown parameters for the nth T search interval $\mathcal{J}_{\mathrm{T},n}$ can be summarized in the random vector $\boldsymbol{\theta}_n \triangleq (\boldsymbol{b}^T_{\mathrm{T},n}\, \boldsymbol{\alpha}^T_{\mathrm{T},n}\, \sigma^2_{w,n})^T$. Note, in particular,

that the noise variance $\sigma^2_{w,n}$ is modeled as a random parameter. Bayesian detection/estimation relies on the posterior distribution

$$p(\boldsymbol{\theta}_n|\boldsymbol{x}_n) \propto p(\boldsymbol{x}_n|\boldsymbol{\theta}_n)p(\boldsymbol{\theta}_n) \tag{E.5}$$

where $p(\boldsymbol{x}_n|\boldsymbol{\theta}_n)$ 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.

Likelihood Function. Using (E.4) and the fact that $w_{n,k}$ is white and Gaussian with variance $\sigma^2_{w,n}$, the likelihood function is obtained as

$$p(\boldsymbol{x}_n|\boldsymbol{\theta}_n) = \mathcal{N}(\boldsymbol{B}_{\mathrm{T},n}\boldsymbol{H}\boldsymbol{\alpha}_{\mathrm{T},n} + \boldsymbol{M}\boldsymbol{\gamma}_n,\, \sigma^2_{w,n}\mathbf{I}_{N_n}) \tag{E.6}$$

Prior distributions. 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, $\|\boldsymbol{b}_{\mathrm{T},n}\| = 1$) 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

$$p(\boldsymbol{b}_{\mathrm{T},n}) = \begin{cases} p_0 & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 0 \\ p_1 & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{E.7}$$

where $p_1 = (1-p_0)/N_{\mathrm{T},n}$ and p_0 is chosen very small. Note that there are $N_{\mathrm{T},n}$ vectors satisfying $\|\boldsymbol{b}_{\mathrm{T},n}\| = 1$ whereas the zero vector is the only vector satisfying $\|\boldsymbol{b}_{\mathrm{T},n}\| = 0$. Thus the probabilities in (E.7) sum to one. The wave indicator vectors $\boldsymbol{b}_{\mathrm{T},n}$ for different search intervals (different n) are assumed to be statistically independent.

The waveform coefficient vectors $\boldsymbol{\alpha}_{\mathrm{T},n}$ for the nth T-wave search interval $\mathcal{J}_{\mathrm{T},n}$ are supposed to depend on the respective estimates obtained in the $(n-1)$th T-wave search interval $\mathcal{J}_{\mathrm{T},n-1}$. The prior of $\boldsymbol{\alpha}_{\mathrm{T},n}$ is defined as

$$p(\boldsymbol{\alpha}_{\mathrm{T},n}|\boldsymbol{b}_{\mathrm{T},n}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) = \begin{cases} \delta(\boldsymbol{\alpha}_{\mathrm{T},n} - \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 0 \\ \mathcal{N}(\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}, \sigma^2_\alpha \mathbf{I}_G) & \text{if } \|\boldsymbol{b}_{\mathrm{T},n}\| = 1 \end{cases} \tag{E.8}$$

where $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$ is the estimate of the T waveform coefficient vector associated with $\mathcal{J}_{n-1}$ and $\mathbf{I}_G$ is the identity matrix of size $G \times G$. For the variance σ^2_α, we choose a value that allows for 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 $\hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}$, i.e., the waveform coefficients are assigned the same values as those estimated in $\mathcal{J}_{\mathrm{T},n-1}$. These definitions of the priors associated with $\boldsymbol{\alpha}_{\mathrm{T},n}$ introduce a memory in the statistical model for the T waveforms and, in turn, result in a sequential type of processing.

Concerning the baseline coefficients and the noise variance, they are assigned similar priors as in the beat-to-beat Bayesian model in Section 3.2. The baseline coefficient vectors $\boldsymbol{\gamma}_n$ for different n are assumed to be independent. Moreover, the baseline coefficients $\gamma_{n,i}$ for a given

n are modeled as independent and identically distributed (iid) zero-mean Gaussian, i.e.,

$$p(\boldsymbol{\gamma}_n) = \mathcal{N}(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}_5), \tag{E.9}$$

with a fixed variance σ_γ^2. The noise variances $\sigma_{w,n}^2$ for different n are modeled as independent and distributed according to the inverse gamma distribution

$$p(\sigma_{w,n}^2) = \mathcal{IG}(\xi, \eta), \tag{E.10}$$

where ξ and η are fixed hyperparameters providing a vague prior.

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

$$p(\boldsymbol{\theta}_n | \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) = p(\boldsymbol{\alpha}_{\mathrm{T},n} | \boldsymbol{b}_{\mathrm{T},n}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) \, p(\boldsymbol{b}_{\mathrm{T},n}) p(\boldsymbol{\gamma}_n) \, p(\sigma_{w,n}^2) \tag{E.11}$$

Posterior distribution. The posterior distribution of the parameter vector $\boldsymbol{\theta}_n$ can be derived using Bayes' rule, i.e.,

$$p(\boldsymbol{\theta}_n | \boldsymbol{x}_n, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) \propto p(\boldsymbol{x}_n | \boldsymbol{\theta}_n) p(\boldsymbol{\theta}_n | \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1}) \tag{E.12}$$

where the different densities have been defined in (E.6), (E.11), and the individual prior distributions have been discussed previously.

Due to the complexity of the resulting posterior distribution, we propose a block Gibbs sampler that generates samples asymptotically distributed according to $p(\boldsymbol{\theta}_n | \boldsymbol{x}, \hat{\boldsymbol{\alpha}}_{\mathrm{T},n-1})$. This sampler has a very similar structure to the block Gibbs sampler proposed in Section 3.2.3, except that the sampling distributions for certain parameters are different since the steps to generate P wave parameters are omitted.

Bibliography

[AAC81] D. R. Adam, S. Akselrod, and R. J. Cohen, "Estimation of ventricular vulnerability to fibrillation through T-wave time series analysis," *Comput. Cardiol.*, vol. 8, pp. 307–310, 1981.

[AAC+04] A. A. Armoundas, C. M. Albert, R. J. Cohen, T. Mela, and TOVA investigators, "Utility of implantable cardioverter defibrillator electrograms to estimate repolarization alternans preceding a tachyarrhythmic event," *J. Cardiovasc. Electrophysiol.*, vol. 15, no. 5, pp. 594–597, 2004.

[Ala] "Image: Marquette electronics," *Alan E. Lindsay ECG Learning Center in Cyberspace*. [Online]. Available: http://interstices.info

[AM79] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall, 1979.

[AMG02] S. Arulampalam, S. Maskell, and N. Gordon., "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process*, vol. 50, no. 2, pp. 174–188, 2002.

[AMRL67] R. Almeida, J. P. Martinez, A. P. Rocha, and P. Laguna, "Characterization of an adaptative filter for the analysis of variable latency neuroelectric signals," *Med. and biol. Eng. Comp.*, vol. 5, no. 6, pp. 539–553, 1967.

[AMRL09] ——, "Multilead ECG delineation using spatially projected leads from wavelet transform loops," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 8, pp. 1996 –2005, Aug. 2009.

[ARA09] F. A. Afsar, M. S. Riaz, and M. Arif, "A comparison of baseline removal algorithms for electrocardiogram (ECG) based automated diagnosis of coronory heart disease," in *Proc. of the 3rd Int. Conf. on Bioinformatics and Biomed. Eng. (ICBBE)*, june 2009, pp. 1–4.

[AS85] J. A. V. Alste and T. S. Schilder, "Removal of base-line wander and power-line interference from the ecg by an efficient fir filter with a reduced number of taps," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 12, pp. 1052–1060, dec. 1985.

[ATNL99] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 192–202, Feb. 1999.

[BG98] S. P. Brooks and A. Gelman, "General methods for monitoring convergence of iterative simulations," *J. of Comput. Graph. Stat.*, vol. 7, no. 4, pp. 434–455, 1998.

[BNP10] R. C. Barr, L. W. Nolte, and A. E. Pollard, "Bayesian quantitative electrophysiology and its multiple applications in bioengineering," *IEEE Reviews in Biomed. Eng.*, vol. 3, pp. 155 –168, 2010.

[BNY01] E. D. A. Botter, C. L. Nascimento, and T. Yoneyama, "A neural network with asymmetric basis functions for feature extraction of ECG P waves," *IEEE Trans. on Neural Networks*, vol. 12, no. 5, pp. 1252–1255, 2001.

[BPSN82] P. O. Borjesson, O. Pahlm, L. Sornmo, and M.-E. Nygards, "Adaptive QRS detection based on maximum a posteriori estimation," *IEEE Trans. Biomed. Eng.*, vol. 29, no. 5, pp. 341–351, May 1982.

[BVCRGLB10] M. Blanco-Velasco, F. Cruz-Roldán, J. I. Godino-Llorente, and K. E. Barner, "Nonlinear trend estimation of the ventricular repolarization segment for T-wave alternans detection," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 10, pp. 2402–2412, 2010.

[BZB06] L. Burattini, W. Zareba, and R. Burattini, "Automatic detection of microvolt T-wave alternans in holter recordings: Effect of baseline wandering," *Biomed. Signal Process. Control*, vol. 1, no. 1, pp. 162–168, 2006.

[BZM99] L. Burattini, W. Zareba, and A. J. Moss, "Correlation method for detection of transient T-wave alternans in digital Holter ECG recordings," *Ann. Electrocardiol.*, vol. 4, no. 4, pp. 416–426, 1999.

[Cab08] A. Cabasson, "Estimation et analyse des intervalles cardiaques," Ph.D. dissertation, Université de Nice - Sophia Antipolis, France, Dec. 2008.

[CAM06] G. Clifford, F. Azuaje, and P. McSharry, *Advanced Methods And Tools for ECG Data Analysis*. Boston (MA): Artech House Publishers, 2006.

[CCL96] Q. Cheng, R. Chen, and T.-H. Li, "Simultaneous wavelet estimation and deconvolution of reflection seismic signals," *IEEE Trans. Geosci. and Remote Sensing*, vol. 34, no. 2, pp. 377–384, 1996.

[CJO01] J. Carlson, R. Johansson, and S. B. Olsson, "Classification of electrocardiographic P-wave morphology," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 4, pp. 401–405, 2001.

[CLK06] P.-C. Chen, S. Lee, and C.-D. Kuo, "Delineation of T-wave in ECG by wavelet transform using multiscale differential operator," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1429–1433, 2006.

[CM07] V. S. Chouhan and S. S. Mehta, "Total removal of baseline drift from ECG signal," in *Proc. of the Int. Conf. on Computing: Theory and Applications*, Kolkata, India, March 2007, pp. 512–515.

[CM08] ——, "Threshold-based detection of P and T-wave in ECG using new feature signal," *Int. J. Comp. Science Net. Security*, vol. 8, no. 2, pp. 144–152, 2008.

[CPD09] S. Chaudhuri, T. D. Pawar, and S. Duttagupta, *Ambulation Analysis in Wearable ECG*. New York: Springer, 2009.

[CR90] M. K. Chung and M. W. Rich, "Introduction to the cardiovascular system," *Alcohol Health and Research World*, vol. 14, no. 4, pp. 269–276, 1990.

[CSCB90] D. A. Coast, R. M. Stern, G. G. Cano, and S. A. Briller, "An approach to cardiac arrhythmia analysis using hidden Markov models," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 9, pp. 826–836, sept. 1990.

[CSH+03] D. J. Christini, K. M. Stein, S. C. Hao, S. M. Markowitz, S. Mittal, D. J. Slotwiner, S. Iwai, M. K. Das, and B. B. Lerman, "Endocardial detection of repolarization alternans," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 855–862, 2003.

[CV06] G. D. Clifford and M. Villarroel, "Model-based determination of QT intervals," in *Proc. of Comput. in Cardiol.*, vol. 33, Valencia, Spain, Sept. 2006, pp. 357–360.

[DdFG01] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York, NY: Springer-Verlag, 2001.

[DG02] P. E. Dilaveris and J. E. Gialafos, "Future concepts in P wave morphological analysis," *Cardiac Electrophysiology Review*, no. 6, pp. 221–224, 2002.

[DGA00] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.

[DGK01] A. Doucet, N. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 613–624, 2001.

[DHC10] J. Dumont, A. I. Hernandez, and G. Carrault, "Improving ECG beats delineation with an evolutionary optimization process," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 607–615, 2010.

[DHS00] R. O. Duda, P. E. Hart, and D. H. Stork, *Pattern Classification (2nd ed.)*. Wiley Interscience, 2000.

[DTD07] N. Dobigeon, J.-Y. Tourneret, and M. Davy, "Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a Bayesian sampling approach," *IEEE Trans. Signal Process*, vol. 55, no. 4, pp. 1251–1263, April 2007.

[DTS07] N. Dobigeon, J.-Y. Tourneret, and J. D. Scargle, "Joint segmentation of multivariate astronomical time series: Bayesian sampling with a hierarchical model," *IEEE Trans. Signal Process*, vol. 55, no. 2, pp. 414–423, 2007.

[EG96] Z. Elghazzawi and F. Gehed, "A knowledge-based system for arrhythmia detection," in *Proc. of Comput. in Cardiol.*, Indianapolis, IN, Sept. 1996, pp. 541–544.

[EZ79] W. A. H. Engelse and C. Zeelenberg, "A single scan algorithm for QRS-detection and feature extraction," in *Proc. of Comput. in Cardiol.*, 1979, pp. 37–42.

[FJJ+90] G. M. Friesen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint, and H. T. Nagle, "A comparision of the noise sensitivity of nine QRS detection algorithms," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 1, pp. 85–98, Jan. 1990.

[FS91] K. Freeman and A. Singh, "P wave detection of ambulatory ecg," in *Proc. of the Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society (EMBC)*, vol. 13, Oct. 1991, pp. 647 –648.

[FT83] G. S. Furno and W. J. Tompkins, "A learning filter for removing noise interference," *IEEE Trans. Biomed. Eng.*, vol. 30, no. 4, pp. 234 –235, 1983.

[Fuk90] K. Fukunaga, *Introduction to statistical pattern recognition (2nd ed.).* San Diego: Academic Press Professional, 1990.

[GCF09] D. Ge, E. L. Carpentier, and D. Farina, "Unsupervised Bayesian decomposition of multi-unit EMG recordings using Tabu search," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 12, pp. 1–9, Dec. 2009.

[GCIF11] D. Ge, E. L. Carpentier, J. Idier, and D. Farina, "Spike sorting by stochastic simulation," *IEEE Trans. Neur. Sys. Rehab. Eng.*, vol. 19, pp. 249–259, 2011.

[GG84] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721 –741, nov. 1984.

[GG89] F. Gritzali and F. Gehed, "Detection of the P and T waves in an ECG," *Comput. and Biomed. Res.*, vol. 22, pp. 83–91, 1989.

[GHM07] A. Gruetzmann, S. Hansen, and J. Müller, "Novel dry electrodes for ECG monitoring," *Physiol. Meas.*, vol. 28, no. 11, pp. 1375–1390, 2007.

[Gib02] R. J. Gibbons, "Acc/aha 2002 guideline update for exercise testing," *American College of Cardiology Foundation and American Heart Association*, 2002.

[GIC11] D. Ge, J. Idier, and E. L. Carpentier, "Enhanced sampling schemes for MCMC based blind Bernoulli-Gaussian deconvolution," *Signal Processing*, vol. 91, pp. 759–772, 2011.

[Ham96] P. S. Hamilton, "A comparison of adaptive and nonadaptive filters for reduction of power line interference in the ecg," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 1, pp. 105 –109, jan. 1996.

[Har78] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. of IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[HB97] R. Haas and J.-C. Belfiore, "A time-frequency well-localized pulse for multiple carrier transmission," *Wireless Personal Comm.*, vol. 5, pp. 1–18, 1997.

[hea] "Image: Conduction system of the heart." [Online]. Available: http://http://en.wikipedia.org/wiki/File:Conductionsystemoftheheart.png

[HTUA93] Y.-H. Hu, W. J. Tompkins, J. L. Urrusti, and V. X. Afonso, "Applications of artificial neural networks for ECG signal detection and classification," *J. Electrocardiology*, vol. 26, pp. 66–73, sept. 1993.

[JOLC93] R. Jane, S. Olmos, P. Laguna, and P. Caminal, "Adaptive Hermite models for ECG data compression: performance and evaluation with automatic wave detection," in *Proc. of Comput. in Cardiol.*, London, UK, Sept. 1993, pp. 389–392.

[KHO07] B.-U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *IEEE Eng. in Med. and Biology Magazine*, pp. 42–57, Jan.-Feb. 2007.

[KMBB99] S. Kadambe, R. Murray, and G. F. Boudreaux-Bartels, "Wavelet transform-based QRS complex detector," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 7, pp. 838–848, July 1999.

[KNHH09] G. Kail, C. Novak, B. Hofer, and F. Hlawatsch, "A blind Monte Carlo detection-estimation method for optical coherence tomography," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, April 2009, pp. 493–496.

[KPP+09] J. W. Kim, H. N. Pak, J. H. Park, G. B. Nam, S. K. Kim, H. S. Lee, J. K. Jang, J. I. Choi, and Y. H. Kim, "Defibillator electrogram T wave alternans as a predictor of spontaneous ventricular tachyarrhythmias in defibrillator recipients," *Circ. J.*, vol. 73, pp. 55–62, 2009.

[KTHD10] G. Kail, J.-Y. Tourneret, F. Hlawatsch, and N. Dobigeon, "A partially collapsed Gibbs sampler for parameters with local constraints," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Dallas, USA, March 2010, pp. 3886–3889.

[KTHD12] ——, "Blind deconvolution of sparse pulse sequences under a minimum distance constraint: A partially collapsed Gibbs sampler method," *IEEE Trans. Signal Process*, vol. 60, no. 6, pp. 2727–2743, 2012.

[LB97] 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, april 1997.

[LBMT11] C. Lin, M. Bugallo, C. Mailhes, and J.-Y. Tourneret, "ECG denoising using a dynamical model and a marginalized particle filter," in *Proc. of the 44th IEEE Asilomar Conf. Signals, Systems and Computers (Asilomar)*, Pacific Grove, CA, Nov. 2011, pp. 364–368.

[Leh97] E. L. Lehmann, *Testing Statistical Hypotheses*. New York: Springer, 1997.

[LI06] C. Labat and J. Idier, "Sparse blind deconvolution accounting for time-shift ambiguity," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006, pp. 616–620.

[LJC94] P. Laguna, R. Jané, and P. Caminal, "Automatic detection of wave boundaries in multilead ECG signals: Validation with the CSE database," *Comput. Biomed. Res.*, vol. 27, no. 1, pp. 45–60, 1994.

[LJM+92] P. Laguna, R. Jané, 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.

[LMGM97] P. Laguna, R. Mark, A. Goldberger, and G. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," in *Proc. of Comput. in Cardiol.*, Lund, Sweden, Aug. 1997, pp. 673–676.

[LMT10] C. Lin, C. Mailhes, and 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.*, vol. 57, no. 12, pp. 2840–2849, Dec. 2010.

[LTC+90] P. Laguna, N. V. Thakor, P. Caminal, R. Jané, H.-R. Yoon, A. B. Luna, V. Marti, and J. Guindo, "New algorithm for QT interval analysis in 24-hour Holter ECG: Performance and applications," *Medical and Biological Engineering and Computing*, vol. 28, no. 1, pp. 67–73, Jan. 1990.

[LZT95] C. Li, C. Zheng, and C. Tai., "Detection of ECG characteristic points using Wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 1, pp. 21–28, Jan. 1995.

[MAO+04] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A wavelet-based ECG delineator: Evaluation on standard databases," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, 2004.

[Mar90] R. Mark, "Biological measurement: Electrical characteristics of the heart," in *Systems Control Encyclopedia*, Singh, M.G.(ed.), Oxford,, 1990, pp. 450–456.

[MCTS03] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289 –294, march 2003.

[MHI+05] A. Maybhate, S. C. Hao, S. Iwai, J. U. Lee, A. B. Guttigoli, K. M. Stein, B. B. Lerman, and D. J. Christini, "Detection of repolarization alternans with an implantable cardioverter defibrillator lead in a porcine model," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 7, pp. 1188–1194, 2005.

[MK77] C. R. Meyer and H. N. Keiser, "Electrocardiogram baseline noise estimation and removal using cubic splines and state-space computation techniques," *Comput. Biomed. Res.*, vol. 10, no. 5, pp. 459–470, 1977.

[MKS06] G. Moody, H. Koch, and U. Steinhoff, "The Physionet/Computers in Cardiology Challenge 2006: QT interval measurement," in *Proc. of Comput. in Cardiol.*, Valencia, Spain, Sept. 2006, pp. 313–316.

[MLM09] V. Monasterio, P. Laguna, and J. P. Martínez, "Multilead analysis of T-wave alternans in the ECG using principal component analysis," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 7, pp. 1880–1890, 2009.

[MN92] I. S. N. Murthy and U. C. Niranjan, "Component wave delineation of ECG by filtering in the fourier domain," *Med. & Bio. Eng. Comput.*, vol. 30, pp. 169–176, 1992.

[MO04] J. P. Martínez and S. Olmos, "Methodological principles of T wave alternans analysis: A unified framework," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 599–613, 2004.

[MOL00] J. P. Martínez, S. Olmos, and P. Laguna, "Simulation study and performance evaluation of T-wave alternans detectors," in *Proc. of the 22nd Annu. Int. Conf. of the IEEE Eng. in Med. and Biology Society (EMBC)*, vol. 3, Chicago, IL , USA, juil. 2000, pp. 2291–2297.

[MOWL06] J. P. Martínez, S. Olmos, G. Wagner, and P. Laguna, "Characterization of repolarization alternans during ischemia: Time-course and spatial analysis," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 4, pp. 701–711, 2006.

[MP92] I. S. N. Murthy and G. S. S. D. Prasad, "Analysis of ECG from pole-zero models," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 7, pp. 741–751, 1992.

[MSV95] S. S. Mehta, S. C. Saxena, and H. K. Verma, "Recognition of P and T waves in electrocardiograms using fuzzy theory," in *Proc. of the First Regional Conference, IEEE Eng. in Med. and Bio. Society*, New Delhi, India, Feb. 1995, pp. 2/54–2/55.

[Nar06] S. M. Narayan, "T-wave alternans and the susceptibility to ventricular arrhythmias," *J. Am. Coll. Cardiol.*, vol. 47, pp. 269–281, 2006.

[NS99a] S. M. Narayan and J. M. Smith, "Differing rate dependence and temporal distribution of repolarization alternans in patients with and without ventricular tachycardia," *J. Cardiovasc. Electrophysiol.*, vol. 10, pp. 61–77, 1999.

[NS99b] ——, "Spectral analysis of periodic fluctuations in electrocardiographic repolarization," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 203–212, 1999.

[NV02] B. D. Nearing and R. L. Verrier, "Modified moving average analysis of T-wave alternans to predict ventricular fibrillation with great accuracy," *Appl. Physiol.*, vol. 92, no. 2, pp. 541–549, 2002.

[Pan96] S. Pandit, "ECG baseline drift removal through STFT," in *Proc. of the 18th Annu. Int. Conf. of the IEEE Eng. in Med. and Biology Society (EMBC)*, vol. 4, nov. 1996, pp. 1405–1406.

[Phy08] Physionet, "Physionet challenge 2008: Detection and quantifying T wave alternans," Sep. 2008. [Online]. Available: http://physionet.org/challenge/2008/

[PLY98] K. L. Park, K. J. Lee, and H. R. Yoon, "Nonlinear trend estimation of the ventricular repolarization segment for T-wave alternans detection," *Med. and Biologi. Eng. and Comp.*, vol. 36, pp. 581–586, 1998.

[PT85] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230–236, 1985.

[PZG+06] O. Paz, X. Zhou, J. Gillberg, H. J. Tseng, E. Gang, and C. Swerdlow, "Detection of T-wave alternans using an implantable cardioverter-defibrillator," *Heart Rhythm.*, vol. 3, pp. 791–797, 2006.

[RAC96] D. S. Rosenbaum, P. Albrecht, and R. J. Cohen, "Electrical alternans and vulnerability to ventricular arrhytimias," *J. Cardiovasc. Electrophysiol.*, no. 7, pp. 1095–1111, 1996.

[RC04] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer, 2004.

[RSS+94] D. S. Rosenbaum, L. E. Smith, J. M. Smith, H. Garan, J. N. Ruskin, and R. J. Cohen, "Electrical alternans and vulnerability to ventricular arrhytimias," *N. Engl. J. Med.*, vol. 330, no. 4, pp. 235–214, 1994.

[SBNP81] L. Sörnmo, P. Börjesson, M. Nygårds, and O. Pahlm, "A method for evaluation of QRS shape features using a mathematical model for the ECG," *IEEE Trans. Biomed. Eng.*, vol. 28, no. 10, pp. 713–717, 1981.

[SCD+11] C. Swerdlow, T. Chow, M. Das, A. M. Gillis, X. Zhou, A. Abeyratne, and R. N. Ghanem, "Intracardiac electrogram T-wave alternans/variability increases before spontaneous ventricular tachyarrhythmias in implantable cardioverter-defibrillator patients: a prospective, multi-center study," *Circulation*, vol. 113, no. 10, pp. 1052–1060, 2011.

[SCV+94] J. M. Smith, E. A. Clancy, C. R. Valeri, J. N. Ruskin, and R. J. Cohen, "Electrical alternans and cardiac electrical instability," *Circulation*, vol. 77, no. 1, pp. 110–121, 1994.

[SF83] A. V. Sahakian and G. F. Furno, "An adaptive filter for distorted line-frequency noise," *Biorricd. Sci. Iristrum.*, vol. 19, pp. 47–52, 1983.

[SGN05] T. Schön, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2279–2289, 2005.

[SLKG02] T. Srikanth, D. Lin, N. Kanaan, and H. Gu, "Presence of T wave alternans in the statistical context-a new approach to low amplitude alternans measurement," *Comput. Cardiol.*, vol. 29, pp. 681–684, 2002.

[SPN+07] R. J. Selvaraj, P. Picton, K. Nanthakumar, S. Mak, and V. S. Chauhan, "Endocardial and epicardial repolarization alternans in human cardiomyopathy: evidence for spatiotemporal heterogeneity and correlation with body surface T-wave alternans," *J. Am. Coll. Cardiol.*, vol. 49, pp. 338–346, 2007.

[SS07] O. Sayadi and M. B. Shamsollahi, "ECG baseline correction with adaptive bionic Wavelet transform," in *Proc. of the 9th Int. Symp. on Signal Processing and Its Applications (ISSPA)*, feb. 2007, pp. 1–4.

[SS09] ——, "A model-based Bayesian framework for ECG beat segmentation," *J. of Physiol. Measurement*, vol. 30, pp. 335–352, 2009.

[SSJ05] R. Sameni, M. B. Shamsollahi, and C. Jutten, "Filtering electrocardiogram signals using the extended Kalman filter," in *Proc. of the 27th Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc. (EMBS)*, Shanghai, China, Sep. 2005, pp. 5639–5642.

[SSJC07] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A nonlinear Bayesian filtering framework for ECG denoising," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2172–2185, dec. 2007.

[STB97] J. S. Sahambi, S. N. Tandon, and R. K. P. Bhatt, "Using Wavelet transforms for ECG characterization. An on-line digital signal processing system," *IEEE Eng. in Med. and Biology Magazine*, vol. 16, no. 1, pp. 77–83, Jan-Feb 1997.

[Str02] P. Strumillo, "Nested median filtering for detecting T-wave offset in ECGs," *Electronics Letters*, vol. 38, no. 14, pp. 682–683, 2002.

[Suz95] Y. Suzuki, "Self-organizing QRS-wave recognition in ECG using neural networks," *IEEE Trans. on Neural Networks*, vol. 6, no. 6, pp. 1469–1477, Nov. 1995.

[TCWP85] The CSE Working Party, "Recommendations for measurement standards in quantitative electrocardiography," *Eur. J. Heart*, vol. 6, no. 10, pp. 815–825, 1985.

[TRC06] J. Thomas, C. Rose, and F. Charpillet, "A multi-HMM approach to ECG segmentation," in *Proc. of 18th IEEE Int. Conf. on Tools with Art. Intel.*, Arlington, VA, Nov. 2006, pp. 609–616.

[TS90] P. Trahanias and E. Skordalakis, "Syntactic pattern recognition of the ECG," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 12, no. 7, pp. 648–657, 1990.

[TZ91] N. V. Thakor and Y. S. Zhu, "Application of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 8, pp. 785–793, 1991.

[TZP90] N. V. Thakor, Y. Zhu, and K. Pan, "Ventricular tachycardia and fibrillation detection by a sequential hypothesis testing algorithm," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 9, pp. 837–3, Sept. 1990.

[UM90] J. K. Udupa and I. Murthy, "Syntactic approach to ECG rhythm analysis," *IEEE Trans. Biomed. Eng.*, vol. 27, no. 7, pp. 370–375, July 1990.

[Van68] H. L. Van Trees, *Detection, Estimation and Modulation Theory.* New York: John Wiley & Sons Inc., 1968.

[VGP+00] J. A. Vila, Y. Gang, J. Presedo, M. Delgado, S. Barro, and M. Malik, "A new approach for TU complex characterization," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 6, pp. 746–772, 2000.

[VP08] D. A. Van Dyk and T. Park, "Partially collapsed Gibbs samplers: Theory and methods," *J. Acoust. Soc. Amer.*, vol. 103, pp. 790–796, 2008.

[WHO11] World Health Organization, "Cardiovascular diseases (CVDs)," Sep. 2011. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs317/en/index.html

[YLL98] Z. Yang, L. Li, and J. Ling, "Approach to recognition of ECG P waves based on approximating functions," *J. Biomed. Eng.*, vol. 15, no. 2, pp. 120–122, 1998.

[Zha05] D. Zhang, "Wavelet approach for ECG baseline wander correction and noise reduction," in *Proc. of the 27th Annu. Int. Conf. of the IEEE Eng. in Med. and Biology Society (EMBC)*, avril 2005, pp. 1212–1215.

[ZMJ03] W. Zong, G. B. Moody, and D. Jiang, "A robust open-source algorithm to detect onset and duration of QRS complexes," in *Proc. of Comput. in Cardiol.*, Thessaloniki, Greece, Sept. 2003, pp. 737–740.