Digital implementation of a wavelet-based event detector for cardiac pacemakers

Rodrigues, Joachim; Olsson, Thomas; Sörnmo, Leif; Öwall, Viktor

Published in:
IEEE Transactions on Circuits and Systems Part 1: Fundamental Theory and Applications

DOI:
10.1109/TCSI.2005.857925

2005

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jan. 2019
Digital Implementation of a Wavelet-Based Event Detector for Cardiac Pacemakers

Joachim Neves Rodrigues, Member, IEEE, Thomas Olsson, Member, IEEE, Leif Sörnmo, Senior Member, IEEE, and Viktor Öwall, Member, IEEE

Abstract—This paper presents a digital hardware implementation of a novel wavelet-based event detector suitable for the next generation of cardiac pacemakers. Significant power savings are achieved by introducing a second operation mode that shuts down 2/3 of the hardware for long time periods when the pacemaker patient is not exposed to noise, while not degrading performance. Due to a 0.13-μm CMOS technology and the low clock frequency of 1 kHz, leakage power becomes the dominating power source. By introducing sleep transistors in the power-supply rails, leakage power of the hardware being shut off is reduced by 97%. Power estimation on RTL-level shows that the overall power consumption is reduced by 67% with a dual operation mode. Under these conditions, the detector is expected to operate in the sub-μW region. Detection performance is evaluated by means of databases containing electrograms to which five types of exogenic and endogenic interferences are added. The results show that reliable detection is obtained at moderate and low signal to noise-ratios (SNRs). Average detection performance in terms of detected events and false alarms for 25-dB SNR is $P_D = 0.98$ and $P_{FA} = 0.014$, respectively.

Index Terms—Cardiac pacemaker, event detector, gated supply, sleep transistors, wavelet filter.

I. INTRODUCTION

DEVICE longevity is a crucial design constraint in the evolving area of medical implants since replacement of implanted devices results in discomfort for the patient and high economical costs. Medical implants such as the cardiac pacemaker may last up to 20 years (projected longevity for one of the world’s longest lasting pacemakers Regency SC+, St. Jude Medical). At the same time, reliable detection performance, closely related to longevity, is essential as the number of devices that may interfere with the pacemaker is ever increasing.

A variety of event detectors for electrocardiograms (ECGs) have been proposed during the last three decades [1]–[3]. However, most of them are unsuitable for pacemaker applications since they do not operate in real time. Traditionally, event detectors for pacemakers are composed of a band-pass filter followed by a programmable threshold level, implemented in analog circuitry [4]. The proposed implementation is optimized for digital circuitry, and the wavelet-based structure offers a higher flexibility for different morphologies. Together with a low-power analog-to-digital converter (ADC) a single-chip solution becomes possible.

Digital hardware is feasible for today’s pacemaker generation due to recent development in low-power ADCs, e.g., the ADC in [5] operates at 2.2 μW. Moreover, shrinking technologies and advances in low-power digital circuitry makes a digital solution a competitive alternative to analog solutions. In favor to analog circuitry [6], a digital implementation has the advantage of accommodating more advanced signal processing such as features for morphology classification, e.g., in implantable cardioverter defibrillators (ICDs), and data compression for postanalysis [7], [8]. Postanalysis provides better knowledge of diseases and improves pacemaker/ICD parameter tuning [9], [10].

The proposed event detector is based on a wavelet filterbank that decomposes the input signal into subbands, followed by hypothesis testing [11]–[15] see Fig. 1. The threshold function of the hypothesis test determines whether the incoming beat is considered as cardiac activity or as noise. A dual operation mode of the detector is proposed by which major parts of the hardware can be shut down when the pacemaker patient is at rest or in a low-noise environment. Reliable detection performance is sustained by a noise detector that operates in supervision mode and reactivates the sleeping hardware when necessary, see Fig. 1.

Dynamic power savings are achieved using a gated clock to shut off parts of the deactivated detector. However, as the event detector is targeted to operate at a low frequency of 1 kHz, leakage is the main contributor to the total power figure. Therefore, leakage reduction techniques are required to efficiently address power dissipation. In the present implementation, gate transistors are used to effectively turn off the supply voltage and, thereby, reduce the leakage power [16]–[19].

In Section II, the detector principles are presented as well as the databases used for evaluation. Section III describes the implementation and optimization in digital hardware. Moreover, a hardware realization for a dual operation mode is presented. The performance of the event detector is discussed in Section IV.
Power optimization, an estimate for the core power consumption, and application-specific integrated circuit (ASIC) placement and routing are presented in Section V. Finally, conclusions are presented in Section VI.

II. MATERIALS AND METHODS

The electrical activity at the pacemaker electrode tip is reflected by the intracardiac electrogram (EGM) [20]–[23]. The depolarization and repolarization waves are decomposed into two perpendicular waves: one that propagates horizontally and another that propagates transversally to the myocardial wall [24]. Thus, the morphologies of these two waves differ significantly. The horizontal wave is composed of a large positive charge that rapidly changes to a negative charge resulting in a **biphasic wave**, whereas the transversal wave results in a **monophasic wave**, see Fig. 2.

Ventricular depolarization usually represents the cardiac event in an EGM and is referred to as the “R-wave” in this study, see Fig. 3; its duration is normally between 60 and 100 ms [4], [20].

A. Wavelet Filterbank and GLRT

This section presents a brief theoretical background of the wavelet filterbank and GLRT needed to comprehend the hardware implementation. A more detailed description is to be found in [15] and [25].

The detector structure was developed with efficient digital hardware implementation in mind. The wavelet filterbank is a combination of a biphasic (antisymmetric) and a monophasic (symmetric) filter function that approximates biphasic and monophasic morphologies. The transfer function $h_{q,n}(t)$ of the biphasic wavelet filterbank is modeled as

$$
\begin{align*}
h_{1,n}(t) &= g_b(t) \\
h_{2,n}(t) &= f(t) \ast g_b(2t) \\
h_{3,n}(t) &= f(t) \ast f(2t) \ast g_b(4t) \\
&\vdots \\
h_{q,n}(t) &= f(t) \ast \cdots \ast f(2^{q-2}t) \ast g_b(2^{q-1}t) \\
\end{align*}
$$

(1)

where $q$ is the scale factor. An analysis has shown that three scales, $q = 2, 3, 4$, are sufficient to cover the frequency spectrum of an R-wave [15]. The case $q = 1$ is not considered in the design, as no prior filtering is defined. The functions $g_b(t)$ and $f(t)$ in (1) are defined as

$$
g_b(t) = [-1 \quad 1]
$$

(2)

and

$$
f(t) = [1 \quad 3 \quad 3 \quad 1]
$$

(3)

respectively. To achieve power-efficient hardware mapping, short filters with integer values are chosen, and, therefore, $g_b(t)$ in (2) is chosen as a first-order difference, and the impulse response $f(t)$ in (3) was chosen as a third-order binomial function [15]. The monophasic filterbank $g_m(n)$ is modeled by reusing $g_b(n)$ as

$$
g_m(n) = g_b(n) \ast g_b(n) = [1 \quad -2 \quad 1]
$$

(4)

such that the transfer function $h_{q,m}(n)$ of the monophasic filterbank is modeled as

$$
h_{q,m}(n) = f(n) \ast \cdots \ast f(2^{q-2}n) \ast g_m(2^{q-1}n).
$$

(5)

The output of the wavelet filterbank is defined as

$$
y(n) = x^T(n)H
$$

(6)

where

$$
x(n) = [x(n), \ldots x(n + N - 1)]^T
$$

(7)

is the input to the wavelet filterbank; $H$ is defined as

$$
H = [\tilde{H}_b \quad \tilde{H}_m]
$$

(8)

and can be efficiently implemented by Mallat’s algorithm [26]. The matrices $\tilde{H}_m$ and $\tilde{H}_b$ in (8) denote the reversals of $H_m$ and $H_b$, respectively, where the latter is defined as

$$
H_b = [h_{2,b} \quad h_{3,b} \quad h_{4,b}].
$$

(9)

The matrix $\tilde{H}_m$ in (8) is computed according to (5). Finally, the decision signal $T(n)$ is computed by a generalized likelihood ratio test (GLRT) as

$$
T(n) = x^T(n)H(H^T H)^{-1}H^T x(n)
$$

(10)
and compared to a threshold [11], [15]. Due to orthogonality of the mono- and biphasic functions, the matrix $(H^T H)^{-1}$ in (10) is symmetric and sparse with half of the elements equal to zero. Thus, half of the multiplications with the elements of the matrix in (10) do not need to be implemented. The threshold level determines the presence of an R-wave and controls the pulse generator as

$$T(n) \geq \beta T_{\text{max}}$$

(11)

where $\beta$ denotes a amplitude threshold fraction, and $T_{\text{max}}$ the average value of the maximum amplitudes of the previously detected events. If the condition in (11) is met, an R-wave is detected.

B. EGM Database

The database contains EGMs from 50 patients, recorded from ventricular pacemaker electrodes, and is used to evaluate the performance of the detector. The EGMs were recorded either during initial implantation or pacemaker replacement, throughout hospitals in Germany (coordinated by Justus-Liebig Universität, Gießen, Germany). The recordings were obtained from patients suffering from AV block and sick sinus syndrome [4]. Most signals were recorded from a unipolar electrode, however, a few signals were recorded with a bipolar electrode. The sampling rate was 44.1 kHz with a resolution of 16 bits. For this particular study the signals were decimated to 1 kHz, since frequencies above 400–500 Hz were judged to be less significant to detection. In order to be compliant with the ADC in [5] a resolution of 8 bits is chosen. The recordings were annotated with respect to a time reference of each R-wave, required for performance evaluation of the detection algorithm. The annotation of an event was defined as the steepest transition phase in the cardiac cycle. Three EGM recordings from different patients are displayed in Fig. 4, and illustrates the inter-patient variability in morphology.

C. Interference Database

The present event detector is tested with respect to sensitivity to exogenic and endogenic interferences, originating outside and inside the body, respectively [27]. The test is done in order to simulate situations when the pacemaker patient is subjected to electronic or magnetic noise. Exogenic interference is limited to sources in everyday life, e.g., caused by electronic household appliances or electronic article surveillance (EAS) systems. Endogenic interference is represented by muscular activity [28]. Fig. 5 presents examples of the different types of interference.
Household appliances represent a common source of interference, caused by electric and magnetic activity within the same frequency range as the R-wave. Furthermore, the magnetic field intensity is dependent on the signal transiency. In this study, recordings from an ac powered hand drill and an electric handmixer were used.

Electronic article surveillance systems have been identified as a common interference source [29]–[32]. Such systems use widely different transmission techniques which makes it difficult to generalize results to how such systems interfere with the pacemaker. In this study, two systems that operate within the R-wave frequency spectra have been tested. The EAS 1 system uses a 16.6-Hz triangular wave modulated with 5 or 7.5 kHz. The EAS 2 system transmits 3-ms long bursts of 58-kHz acoustomagnetic signals with a high amplitude at an interval of 27 ms. The pulse period of 30 ms (33 Hz) is considered to be the reason for possible interference with pacemakers.

Muscular activity is an endogenic interference source which spectrally overlaps with heart signals. In this study, signals recorded pressing the palms together have been considered. The effect of muscular noise on pacemaker performance was one of the first studies on pacemaker interference [28].

Fig. 6 shows the morphology of an EGM with interferences originating from a hand drill and muscular activity, respectively.

III. DIGITAL HARDWARE MAPPING AND OPTIMIZATION

This section describes how the wavelet filterbank and the GLRT are implemented in digital hardware. The proposed structure has been optimized with respect to wordlength and numerical strength reduction to reduce area and power consumption [33], [34].

A. Structure of the R-Wave Detector

The implemented wavelet filterbank consist of three branches, $q = 2, 3, 4$, that scale and filter the signal from
are elements, see Fig. 8. The delays for each branch are presented and the monophase filter output are represented as $Y_{q,d}(z)$ and $Y_{q,m}(z)$, respectively.

Reusing $G_b(z)$, see Fig. 8. The delays for each branch are presented in Table I, and the impulse responses of the filterbank are presented in Fig. 9.

\[
F(z) = 1 + 3z^{-(q-1)} + 3z^{-(2q-2)} + z^{-(2q-1)}
\]

\[
G_b(z) = -1 + z^{-q}.
\]

Reusing $G_b(z)$ according to (4) realizes the monophasic filterbank using a single branch for one scale factor, see Fig. 8, and the output of the filterbank realizes (6). However, in order to center the functions to the longest propagation delay, which is in the third branch, it is necessary to introduce additional delays in $G_b(z)$, see Fig. 8. The delays for each branch are presented in Table I, and the impulse responses of the filterbank are presented in Fig. 9.

Since $x^T(n)H = H^T x(n)$, the remaining part of (10) to be implemented is the multiplication by (12), shown at the bottom of the page. A matrix which is symmetric and sparse with half of its elements equal to zero. The multiplication of $y(n)$ with the first column of $(H^T H)^{-1}$ and the first element of $H^T x(n)$ is carried out as depicted in Fig. 10, where $c_{3k+j}$ are elements of $(H^T H)^{-1}$ and $y_{3k+j}(n)$ the output of the filterbank, with $k = 0, 1$ and $j = 1, 2, 3$.

\[
(H^T H)^{-1} = \begin{bmatrix}
4.25 & -2.81 & 0.71 & 0 & 0 & 0 \\
-2.81 & 4.47 & -1.75 & 0 & 0 & 0 \\
0.71 & -1.75 & 1.49 & 0 & 0 & 0 \\
0 & 0 & 0 & 4.84 & -2.31 & 0.6 \\
0 & 0 & 0 & -2.31 & 4.29 & -1.49 \\
0 & 0 & 0 & 0.6 & -1.49 & 1.77
\end{bmatrix}
\]
B. Hardware Optimization

The aim of hardware optimization is the reduction of silicon area as well as the reduction of power dissipation.

1) Optimization of the Wavelet Filterbank: The internal wordlength of the wavelet filterbank and the output signal $y(n)$ are bit-optimized in order to reduce complexity. In a theoretically worst case scenario mathematical operations in the filterbank leads to an extended dynamic range and therefore wordlength has to increase accordingly, in order to be sure to avoid overflow. However, this is a very pessimistic approach leading to large overhead due to excessive number of bits. In order to determine the maximum number of required bits for a more realistic scenario, all recordings in the EGM database have been analyzed using the filterbank. The internal wordlength at $F(z)$ and the output $Y_1(z), \ldots, Y_6(z)$ was traced in order to determine the number of bits required to represent the largest occurring number. This analysis has shown that overflow can be avoided if the dynamic range is increased by two bits compared to the input wordlength $N$. Thus, the wordlength, $N_{wc}$, for a worst case scenario could be reduced significantly from $N + 15$ to $N + 2$, see Table II, and the target implementation has a wordlength of ten bits at $y_1(n), \ldots, y_6(n)$. Saturation logic guarantees that the signal at $y(n)$ is upwards limited to values representable by ten bits if overflow should occur. This optimization leads to significant reductions in the filterbank and results in narrower multipliers and adders in the following GLRT.

Furthermore, the fixed multipliers in each branch of the filterbank are implemented as shift-add instructions, carried out during one clock cycle, referred to as numerical strength reduction [33]. This optimization results in area and power reduction.

The straight-forward block diagram of the filterbank in Fig. 8 has an excessive number of delays. Thus, the number of registers in $G_6(z)$ is minimized by reusing the registers needed to center the impulse responses in Fig. 9. This results in a reduction of approximately 300 (1-bit) registers.

2) Optimization of the GLRT: Substitution of the real values in $(H^T H)^{-1}$ with their respective rounded integer values reduces the computational cost of the GLRT, see Table III. As multiplication is a more complex operation than addition, complexity reduction is achieved by trading multipliers against adders [33]. Therefore, all the multiplications with the elements of $(H^T H)^{-1}$ are replaced with shift and add operations performing the same operation during one clock cycle. The only multiplication that remains in the GLRT is the one by $H^T x(n)$, which is the output of the wavelet filterbank and already computed in (6). The result of this operation is implemented by the rightmost multiplier in the schematic in Fig. 10.

This optimization achieves a reduction of 77% in multipliers while the number of adders is increased by 41%, see Table III. Thus, power and area is reduced in the filterbank hardware.

A typical decision signal using Matlab floating point and integer coefficients, respectively, is presented in Fig. 11. It is shown that the deviation of the GLRT output after numerical strength reduction is minor. Detection performance using the database, between optimized and original structure was compared, and remained unchanged.

C. Dual Operation Mode

The R-wave detector in [12] is designed to assure good detection performance when the EGM is corrupted with noise, see Fig. 13 shown later. However, during long periods the pacemaker patient is not exposed to noise, e.g., during sleep (rest), low physical activity, etc. Therefore, it is highly desirable to automatically shut off parts of the R-wave detector during such periods to save power. At the same time the R-wave detector must be able to operate with full noise suppression performance whenever necessary. Operation when the patient is not exposed to noise is referred to as normal mode as this is the case most of the time. Alert mode is when the entire filterbank and GLRT are active.

One approach is to shut off one or two branches in the wavelet filterbank during normal mode. Thus, it is also possible to shut off parts of the GLRT with respect to the inactivated branches in the filterbank. Section IV-C presents a performance analysis of how the branches are activated or deactivated in alert and normal mode, respectively. The parts of the R-wave detector being shut off in normal mode are triggered by a gated clock [35]. This clock tree is enabled by the noise detector.

### Table I

<table>
<thead>
<tr>
<th>$q$</th>
<th>$b_{1,q}$</th>
<th>$m_{s,q}$</th>
<th>$b_{q}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th>$N_{wc}$</th>
<th>$N + 6$</th>
<th>$N + 7$</th>
<th>$N + 11$</th>
<th>$N + 12$</th>
<th>$N + 14$</th>
<th>$N + 15$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{imp}$</td>
<td>$N + 1$</td>
<td>$N + 1$</td>
<td>$N + 2$</td>
<td>$N + 1$</td>
<td>$N + 2$</td>
<td>$N + 1$</td>
</tr>
</tbody>
</table>

### Table III

<table>
<thead>
<tr>
<th></th>
<th>not optimized</th>
<th>optimized</th>
<th>savings %</th>
</tr>
</thead>
<tbody>
<tr>
<td>MULT</td>
<td>26</td>
<td>6</td>
<td>-77</td>
</tr>
<tr>
<td>ADD</td>
<td>32</td>
<td>45</td>
<td>+41</td>
</tr>
</tbody>
</table>

Fig. 10. Data flow diagram of a block in the GLRT. The multipliers with shaded background are realized by shift-add operations.
alternative to consider is a time-multiplexed architecture, however, this leads to a more complex control structure and a higher operating frequency.

In order to make the R-wave detector resilient to noise, a noise detector has been supplemented, see Fig. 7. The noise detector operates in supervision mode and guarantees full noise suppression performance by reactivating the hardware that has been shut off during normal mode. The power savings gained by deactivating parts of the R-wave detector must not be dissipated by the noise detector since the proposed modification would then lack significance. Therefore, it is necessary to design a low-complexity noise detector.

D. Noise Detector

In this study, noise quantification is based on a zero-crossing rate measurement $Z_S(n)$ [36]; the number of zero crossings is the number of times a sequence changes sign. The $Z_S(n)$ measurement on all the recordings in the database show that an upper bound for a patient ranges from 5 to 7 zero-crossings during 100 ms.

If the input signal has a dc component, a zero-crossing measurement cannot be carried out correctly. Therefore, any dc component of an EGM is filtered out before $Z_S(n)$ is estimated by a differencing filter as

$$d(n) = x(n) - x(n - 1)$$

where

$$\text{sgn}(d(n)) = \begin{cases} +1, & d(n) \geq 0 \\ -1, & d(n) < 0 \end{cases}$$

In order to define a short-term $Z_S(n)$, (14) can be used as

$$Z_S(n) = \frac{1}{2N} \sum_{m=-N+1}^{N} |\text{sgn}(d(m)) - \text{sgn}(d(m-1))|$$

where $N$ is the length of the short-term interval. Since most zero-crossings occur during an R-wave, the length of $N$ needs to be longer than this interval, usually no longer than 100 ms [4]. However, to achieve a flexible implementation $N$ is a programmable parameter in the target implementation.

E. Noise Detector Implementation

A zero crossing can be identified by comparing the signs of two successive samples computed in (13). Using digital hardware and two’s complement representation the comparison can be carried out by analyzing the most-significant-bit (MSB) which indicates the sign of a number. A zero crossing has occurred if

$$\text{MSB}(d(n)) \oplus \text{MSB}(d(n - 1)) = 1$$

where $\oplus$ is the XOR function. The number of zero-crossings, indicated by a low to high transition at the XOR gate, is accumulated for the time $N$, see Fig. 12. Noise is detected when $Z_S(n)$ exceeds an upper bound $Z_{\text{acc}}$, causing the R-wave detector to switch to alert mode. $Z_{\text{acc}}$ differs for every patient and needs to be programmed during pacemaker surgery or check-up. The $Z_S$-accumulator is reset after $N$ samples are processed to start a new $Z_S$ determination for the next input sequence. Thus, a counter that provides a reset signal after $N$ samples is needed in addition to the schematic in Fig. 13. The $Z_S$-accumulator is implemented by a register and an adder. The simple noise detector
structure results in very little area and power overhead when implemented in digital hardware. More sophisticated alternatives can be considered but would result in higher complexity.

IV. DETECTION PERFORMANCE

The performance of the implemented detector is analyzed by adding various interferences to the EGM recordings such that different signal-to-noise ratios (SNRs) are obtained. Detection performance is measured by computing the probability of missed detection \( P_D \) and false alarms \( P_{FA} \) as

\[
P_D = \frac{N_T}{N_T + N_M} \quad \text{and} \quad P_{FA} = \frac{N_{FA}}{N_T + N_{FA}}
\]

where \( N_T \) is the number of true detections, \( N_M \) the number of missed detections, and \( N_{FA} \) the number of false alarms. A true detection is defined as an event that occurs within 50 ms of the annotation, whereas events outside this interval are declared as false alarms.

A. SNR Definition

The analyzed signal consists of the nonstationary EGM, \( x(n) \), to which a noise signal, \( v(n) \), has been added. The SNR of \( y(n) \) is defined as

\[
\text{SNR} = 20 \cdot \log \left( \frac{V_x}{\sigma_V} \right),
\]

where \( V_x \) is the average peak-to-peak amplitude of all the R-waves in one EGM recording and \( \sigma_V \) the standard deviation of the noise to be added. The peak-to-peak amplitude \( V_x \) is calculated according to

\[
V_x = \frac{1}{N_x} \sum_{i=1}^{N_x} \left[ \max_{-50 \leq m \leq 20} \{ x(R_i + m) \} \right] + \left[ \min_{-50 \leq m \leq 20} \{ x(R_i + m) \} \right]
\]

where \( x(R_i) \) is a vector containing an R-wave positioned at \( R_i \), and \( N_x \) is the number of R-wave templates. The standard deviation \( \sigma_V \) is calculated according to

\[
\sigma_V = \sqrt{\frac{1}{N_v} \sum_{i=1}^{N_v} (v(i) - \overline{v})^2}
\]

where \( \overline{v} \) is the mean of the noise signal and \( N_v \) the number of discrete samples.

B. Detection Performance for Noisy EGMs

Although the pacemaker patient is mostly exposed to low-noise environment, the importance of handling heavily disturbed EGMs must nonetheless be addressed. The detection performance for the circumstance when the pacemaker patient is exposed to various interferences is analyzed in this section. The EGMs are disturbed with recordings of the interference database, see Sections II-B and C.

Shape and body constitution of humans vary considerably and, therefore, it is not possible to find one single estimate of how much noise can interfere with the pacemaker. However, a SNR of 20 dB corresponds to a very high interference level which should include the worst case situation in real life. Thus, noise levels are chosen which results in SNRs of 20 and 25 dB to assure that this situation can be handled by the detector. A typical EGM that is disturbed by interference and the corresponding decision signal \( T(n) \) is shown in Fig. 13.

The threshold level \( \beta \) is varied from 0.3 to 0.5. A low value for \( \beta \) produces high rates of \( P_D \), however, \( P_{FA} \) will also increase as more false events will exceed the threshold. Contrarily, a high \( \beta \) leads to a lower value for \( P_D \) and \( P_{FA} \).

It can be observed that the detector is more sensitive to interference that originates from the EAS2 system and muscular activity than from other sources, see Fig. 14. This is the case for \( P_D \) and \( P_{FA} \) for both noise levels. Interference that originates from muscle contractions is the most difficult noise to suppress in all the tested EGMs [28]. The results show that the detector attains reliable detection performance at moderate to low SNRs. For 20-db and 25-db SNR the average performance for all noise sources is \( P_D = 0.88 \) and \( P_{FA} = 0.13 \) and \( P_D = 0.98 \) and \( P_{FA} = 0.014 \), respectively. Moreover, a threshold level that strikes a good balance between \( P_D \) and \( P_{FA} \) is \( \beta = 0.4 \).
The pacemaker is mostly operating in a low-noise environment and, therefore, a performance analysis is carried out where no additional noise is added to the EGM. The threshold level $\beta$ is 0.4. The total number of 3200 events for all the recordings in the EGM database has been analyzed. In order to find out which branches of the filterbank can be shut off without performance degradation for all recordings in the EGM database, the branches have been activated in different combinations, see Table IV. The mode of a single branch, active or inactive, is binary coded. If a branch is active it is coded as 1 in a 3 digit word, e.g., 110 indicates that branch one and two ($q = 2, 3$) are active while three ($q = 4$) is inactive. However, as the branches are connected in series it is not possible to entirely shut off the first branch if the second or third branches is active, see Fig. 7. For such cases, only the filtering part $G_3(z)$ of the preceding branches are shut off, see Fig. 8.

The average detection and false alarm rate $P_D$ and $P_{FA}$, respectively, is presented in Table IV. In none of the simulated cases, the detection performance drops below 0.97 whereas the highest $P_{FA}$ rate is less than 0.021, for a threshold level $\beta = 0.4$. The highest detection and lowest false alarm rates, $P_D$ and $P_{FA}$, are obtained if all branches, i.e., 111, in the filterbank are operating, >0.99 and <0.001, respectively. However, the difference between 111 and 100 is negligible as only a minor difference for $P_{FA}$ is measurable. Furthermore, it is possible to shut off block three to six in the GLRT, whereas block one and two operate partially, see Fig. 7. As the first branch has to operate partially in all the combinations it is of further advantage to inactivate filterbank two and three since the amount of hardware that can be shut off is higher compared to other combinations.

In the target implementation, not all presented modes will be implemented. Only two modes will be considered, 100 for the normal mode and 111 for the alert mode.

1) Noisy Signal in Normal Mode: The noise detector activates branch two and three if the EGM is corrupted by noise, and, thereby, sustains filtering performance. Nevertheless, it is of interest to analyze detection performance for the circumstance that noise is present but not indicated by the noise detector. To evaluate detection performance for such a circumstance, noise is added (25-dB SNR) to the recordings in the EGM database. The situation that no noise is detected is simulated by a forced normal operation mode, i.e., branch two and three are permanently off. Detection performance is analyzed by computing $P_D$ and $P_{FA}$ for $\beta = 0.4$ for all the recordings in the database, see Table V. It can be seen that performance degrades
if the EGM is disturbed and the R-wave detector continues the normal mode operation. The $P_D$ rate for EAS2 drops to $\sim0.52$ which is unacceptable. Thus, it is necessary to switch to alert mode and to reactivate branch two and three to sustain reliable performance.

V. POWER CONSUMPTION

The power consumption of a digital ASIC is defined as

$$P = P_{\text{dyn}} + P_{\text{dp}} + P_{\text{leak}}$$  \hspace{1cm} (21)

where $P_{\text{dyn}}$ is the switching power, $P_{\text{dp}}$, the direct-path power, and $P_{\text{leak}}$ the leakage power [37]. For a long time, dynamic power consumption has been the dominant source whereas leakage power has been ignored for most applications. However, with shrinking technology and decreasing threshold voltage $V_T$, $P_{\text{leak}}$ represents a substantial or dominant share of the total power. Leakage power is consumed as long as the supply voltage is switched on, regardless of the switching activity. For the presented design leakage power is the main contributor due to the low clock frequency and correspondingly low switching activity.

A. Gated Supply Lines

The R-wave detector is operating at a low clock frequency of 1 kHz, implying that leakage power will have a large share of the total power consumption. To reduce the leakage current effectively three approaches can be applied: a multithreshold CMOS (MTCMOS) process, transistor stacking or the combination of transistor stacking and MTCMOS [38], [16], [39]. Using MTCMOS the design is implemented using high $V_T$ devices for the noncritical and low-$V_T$ devices for the critical path. Transistor stacking, on the other hand, cuts off one of the supply supply rails and thereby reduces the leakage power. This can be implemented by using a standard CMOS process without dual $V_T$. Combining the two techniques by using high-$V_T$ transistors for transistor stacking will lead to a higher leakage reduction but requires extra process steps with corresponding costs [38].

The presented design has been implemented in a United Microelectronics Corporation (UMC) low-leakage process that provides high-$V_T$ devices. Furthermore, transistor stacking is applied to achieve substantial leakage power reduction. An extra gate transistor is introduced in the leakage path and can either be placed between the power supply, $VDD$, and the cells or between ground (GND) and the cells, see Fig. 15 [37]. The gate transistor is turned on and off in the alert and normal mode, respectively, and thus the cell supply voltage is gated which achieves significant leakage reduction. An extra sleep transistor reduces leakage current by orders of ten and the gate transistor can be shared among multiple cells which amortizes the area overhead of an extra transistor [37].

1) Transistor Sizing: The gate transistor must be large enough to sink the current flowing through the cells during alert mode. However, a too large transistor degrades the stacking effect and introduces an area overhead. The dimensions of a transistor that match the needs to gate a single cell can be determined analytically. Unfortunately, this is not the case if several cells with different properties are driven by one gate transistor, which is the case in the proposed design. Therefore, Spice level simulations are carried out to find out how leakage current is influenced by the chosen size of the gate transistor, see Fig. 15. The static resistance of a transistor can be computed as

$$R_{(P_N)} = \frac{L_G}{Wk_f(VGS-V_T-VDS)}$$  \hspace{1cm} (22)

where $L_G$ is the channel length, $W$ is the channel width, and $k_f$ the process transconductance, which for the provided cell library results in $R_{(P_N)} < 1 \Omega$ for a $33 \mu m$ wide nMOS and a $69-\mu m$ wide pMOS transistor [37]. Thus, $R_{(P_N)}$ is negligible compared to the wire resistance for power routing.

B. Leakage Reduction Estimation

In order to find an estimate for leakage reduction, the schematic in Fig. 15 is used, since gate transistors cannot be included in a gate level power-simulation. The number of equivalent inverter cells with respect to the R-wave detector hardware that is shut off needs to be determined. Therefore, the number of cells to shut off in normal mode and their corresponding leakage power are identified. A brief overview of these cells is presented in Table VI. This overview represents 88% of the hardware being shut off in normal mode, and is considered to be sufficient to determine the sleeping hardware leakage in this simulation. In order to be consistent with the schematic in Fig. 15, the leakage power of the cells in Table VI is interpolated to the leakage power using only inverters, e.g., an AND2 cell has 20% higher leakage than an INV cell, thus 91 AND2 cells are approximated by 109 INV cells. Thus, an equivalent leakage power to the shut down part of the detector would result in $\sim4500$ minimum sized inverter cells. To simplify simulation, the maximum dimension for a transistor,
restricted by the simulation tool, is used, which is 625 times the minimum size. The number of maximum sized transistors is indicated by the load factor \( L \). With the equivalent number of inverters estimated in Table VI the load can be approximated by 7. The width of the gate transistor in Fig. 15 is set to 33 \( \mu \text{m} \) and loaded with an increasing number of balanced inverters with the transistor dimension of \( L_P/W_P = 0.13/100 \) and \( L_N/W_N = 0.13/33 \). The width of the pMOS transistor is thrice that of the nMOS width to balance the inverter cell. The leakage current \( I_L \) is estimated for having either nMOS/pMOS of the inverter gate in on or off mode, with equal likelihood. The average of these measurements is presented in Table VII, where \( \text{IL}_{\text{cbl}} \) and \( \text{IL}_{\text{ll}} \) is the gated leakage current and the non-gated leakage current, respectively. Simulation on \textit{Spice} level results show that the leakage reduction rate increases with the number of inverters driven by the gate transistor. As presented in Table VII, gated-ground is more suitable if a large number of cells will be shut off, and the area overhead for a nMOS gate transistor is 52% smaller compared to the pMOS solution. Thus, gated ground is chosen for this silicon implementation. The expected leakage reduction is approximately 96% according to the simulation results presented in Table VII.

### C. Gate Level Power Estimation

In order to confirm the expected power savings, the power dissipation of the two modes is estimated on gate level [40]. Since sleep transistors cannot be included in a gate level simulation, results of the leakage reduction estimation in Table VII are used. The power estimation in normal mode is carried out by removing the hardware that is shut off from the netlist, i.e., the hardware neither leaks nor switches. However, leakage is included in the estimate by adding a fraction, according to Table VII. Using gated ground and an estimated load of 7 results in a leakage reduction of approximately 96%. The power analysis confirms that leakage power is the dominant power source, using the provided technology and a clock frequency of 1 kHz, see Table VIII. Using solely a gated clock would not result in equal power savings since this only effects the dynamic power consumption and leakage current still exists as long as the cells are connected to the supplies. Thus, the power saving would only be 29% if no gated-supply is used. The presented estimates only serve for identification of the dissipation sources and to provide a relative measurement of the power savings. A more accurate estimate will be available on the fabricated ASIC level. Due to a large time slack in the critical path it is possible to gain further power savings by lowering \( V_{\text{ddl}} \). Furthermore, as a scaled supply voltage increases the threshold voltage, a reduction in leakage is expected [41]–[43]. Such low supply voltage eliminates short-circuit power as \( V_{\text{ddl}} \) is expected to be below \( V_{\text{th}} + V_{\text{ddp}} \) [37].

However, such estimates cannot be obtained in the current gate-level simulation. The ASIC core is expected to operate in the 100 nW region in the target implementation.

### D. Placement and Routing

The ASIC core consists of two regions with shared \( V_{\text{ddl}} \) and two independently routed GND supplies. The sleep-transistors that gate GND supply are implemented between the GND core-ring and power-supply pad. The transistor gates are connected to the alert signal and thereby shut off the transistors in normal mode. The total chip area, inclusive pads, is 1 mm\(^2\) with 7842 gates using a 0.13-\( \mu \text{m} \) UMC process, see Fig. 16. The ASIC has been sent for fabrication and further power analysis will be carried out in the future.

<table>
<thead>
<tr>
<th>Number</th>
<th>( I_{\text{leak}} )</th>
<th>INV equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-latch</td>
<td>395</td>
<td>46</td>
</tr>
<tr>
<td>NAND</td>
<td>365</td>
<td>11</td>
</tr>
<tr>
<td>NOR</td>
<td>329</td>
<td>11</td>
</tr>
<tr>
<td>FA</td>
<td>633</td>
<td>43</td>
</tr>
<tr>
<td>XOR</td>
<td>61</td>
<td>11</td>
</tr>
<tr>
<td>AND2</td>
<td>91</td>
<td>18</td>
</tr>
<tr>
<td>INV</td>
<td>791</td>
<td>15</td>
</tr>
<tr>
<td>total</td>
<td>2665</td>
<td>-</td>
</tr>
</tbody>
</table>

#### TABLE VI

**Brief Summary of Cells That Are Shut Off in Normal Mode. Summary Represents Approximately 88% of Implemented Hardware**

#### TABLE VII

<table>
<thead>
<tr>
<th>PMOS</th>
<th>NMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L )</td>
<td>( I_{\text{gcl}} ) [nA]</td>
</tr>
<tr>
<td>1</td>
<td>46.4</td>
</tr>
<tr>
<td>2</td>
<td>27.6</td>
</tr>
<tr>
<td>5</td>
<td>18.4</td>
</tr>
<tr>
<td>10</td>
<td>12.5</td>
</tr>
</tbody>
</table>

#### TABLE VIII

**Core Power Estimation at Gate Level Using a 0.13-\( \mu \text{m} \)**

**Low-Leakage Library.** \( F = 1 \text{ kHz}, V_{\text{ddl}} = 1.2 \) nV

<table>
<thead>
<tr>
<th></th>
<th>Alert [nW]</th>
<th>Normal [nW]</th>
<th>Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>dynamic</td>
<td>24.6</td>
<td>6.8</td>
<td>72%</td>
</tr>
<tr>
<td>short-circuit</td>
<td>22.1</td>
<td>6.1</td>
<td>72%</td>
</tr>
<tr>
<td>leakage</td>
<td>67.8</td>
<td>&lt; 25</td>
<td>&gt; 63%</td>
</tr>
<tr>
<td>total</td>
<td>114</td>
<td>&lt; 38</td>
<td>&gt; 67%</td>
</tr>
</tbody>
</table>

Fig. 16. Layout of the routed ASIC.
The design is pad-limited as various control signals are fed to the IOs for verification purpose. These control signals are not needed in the target implementation and therefore it is possible to reduce the chip size significantly. Moreover, the sleep transistors can be accommodated in the supply pads which results in further area reduction.

VI. CONCLUSION

The implementation of a waveform-based R-wave detector in 0.13-μm low-leakage UMC technology has been presented in this paper. The design has been power optimized by applying strength reduction, as well as wordlength and register minimization. The inclusion of a noise detector facilitates a dual operation mode. Thus, 2/3 of the hardware can be shut down if the pacemaker patient is at rest or not exposed to interferences (normal mode). Reliable detection performance is sustained by reactivating the sleeping hardware whenever necessary.

Gate transistors in the GND supply lines are used to address the dominating leakage power and to effectively reduce the power consumption when operating in normal mode. Gate-level power estimation predicts 67% power savings if operating in normal mode; no performance degradation is measurable for such cases. The total chip area is 1 mm² and is expected to operate in the 100 nW range.

The ASIC has been fabricated in 0.13-μm UMC low-leakage technology and further performance measures as well as power optimization will be carried out.

ACKNOWLEDGMENT

The authors are grateful to Dr. M. Åström for contributing his expertise and to St. Jude Medical AB, Järfälla, Sweden, for providing the data for this study.

REFERENCES


Joachim Neves Rodrigues (S’00–M’05) was born in Wiesbaden, Germany, in 1971. He received the electrical engineering and computer science degree from the University of Applied Sciences, Kaiserslautern, Germany, and the Ph.D. degree from the Department of Electroscience, Lund University, Lund, Sweden, in 2000 and 2005, respectively.

His research interest is in the application of digital hardware to realize real-time biomedical signal processing as well as low-energy high-performance deep submicrometer system-on-chip design.

Thomas Olsson (S’98–M’04) was born in Karlshamn, Sweden, in 1973. He received the M.S. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden, in 1998 and 2004, respectively.

His research interests include clock generators, power-supply management, and on-chip signalling for system-on-chip application-specific integrated circuits. He is currently with Ericsson Mobile Platforms, Lund, Sweden.

Leif Sörnmo (S’80–M’85–SM’02) received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden, in 1978 and 1984, respectively.

He held a research position with the Department of Clinical Physiology, Lund University, from 1983 to 1995, where he worked on computer-based ECG analysis. Since 1990, he has been with the Signal Processing Group, Department of Electroscience, Lund University, where he now holds a position as a Professor in biomedical signal processing. His main research interests include statistical signal processing and the modeling of biomedical signals. Current research projects include methods in ischemia monitoring, time-frequency analysis of atrial fibrillation, power efficient signal processing in pacemakers, hemodialysis, and detection of otoacoustic emissions. He is, together with Pablo Laguna, the author of Bioelectrical Signal Processing in Cardiac and Neurological Applications (Elsevier, 2005).

Dr. Sörnmo has been an Associate Editor of Computers in Biomedical Research (1997–2000). He is currently on the editorial boards of the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING and the Journal of Electrocardiology.

Viktor Öwall (S’90–M’95) received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden, in 1988 and 1994, respectively.

During 1995 to 1996, he joined the Electrical Engineering Department, the University of California at Los Angeles as a Visiting Researcher on a Postdoctoral grant from the Swedish Research Council for Engineering Sciences where he mainly worked in the field of multi-media simulations. Since 1996, he has been with the Department of Electroscience, Lund University, where he is currently an Associate Professor. His main research interest is in the field of digital hardware implementation, especially algorithms and architectures for wireless communication, image processing and biomedical applications. Current research projects include combining theoretical research with hardware implementation aspects in the areas of pacemakers, channel coding, video processing, and digital holography.

Dr. Öwall was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: ANALOG AND DIGITAL SIGNAL PROCESSING from 2000–2002.