Low Power Analog and Digital (7,5) Convolutional Decoders in 65 nm CMOS

Meraji, Reza; Sherazi, S. M. Yasser; Anderson, John B; Sjöland, Henrik; Öwall, Viktor

Published in:
IEEE Transactions on Circuits and Systems Part 1: Regular Papers

DOI:
10.1109/TCSI.2015.2423792

2015

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.
Abstract—Targeting emerging energy constrained bio-
implantable or wearable wireless devices, this work presents
design space exploration of decoding circuits for (7,5),
convolutional codes in 65 nm CMOS for ultra-low power
operation. Decoders operating in digital and analog domains
are designed and measured for energy efficiency, Bit Error Rate
(BER) performance and throughput. For the analog decoders
which are sensitive to noise and device mismatch, the overall
effects of transistor dimensions on the output BER are also
investigated. The digital implementation with 0.11 mm² area
consumes minimum energy at 0.32 V supply, which gives 9 pJ/bit
energy efficiency at 125 kb/s and 2.9 dB coding gain. Likewise, in
analog domain, three decoding circuits are fabricated that share
the same topology and design, except for transistor dimensions.
The largest analog decoding core (AD1) takes 0.104 mm² and
the other two (AD2 and AD3) are 0.035 mm² and 0.015 mm²,
respectively. Consequently, coding gain in trade-off with silicon
area and throughput is presented. The analog decoders operate
with 0.8 V supply, and 2.3 dB coding gain with 10 pico-Joules
per bit (pJ/bit) energy efficiency is achieved at 2 Mbps.

Index Terms—circuits for wearable/implantable (bio) elec-
tronics, low power design, circuits and systems for coding in
communication systems, design space exploration.

I. INTRODUCTION

Increased attention to health care practices in recent years
has stimulated interest on battery-supplied small wireless
devices that can be worn or implanted in the human body
[1], [2]. One of the main challenges for these devices is to
maintain a long lifetime without having to recharge or replace
the batteries. While high data rates are not needed in most
scenarios, maintaining communication reliability is important.
In order to minimize the errors that occur during transmission
over a noisy channel, error correcting codes (ECCs) may be
utilized. Due to the amount of required computation to be
carried out, decoding circuits are usually power demanding.
Therefore, energy efficient implementation of decoders can
greatly improve the lifetime of the device.

For decoders implemented in the digital domain, technol-
yogy scaling has reduced both required chip area and power
consumption. In [3] the scaling trend of the energy efficiency
of analog and digital decoders has been investigated based
on published works over the last decade. There, also an
efficient digital LDPC decoder is presented for sub-threshold
(sub-$V_T$) operation, which is evaluated via simulations based
on the models described in [4] and [5]. While the dynamic
power quadratically decreases with voltage scaling, the leakage
power does not scale as much, which due to the reduced
speed of processing results in increased leakage energy per
operation. At a certain supply voltage the energy per operation
is minimized, which is referred to as the minimum energy
point. The speed of processing, however, becomes significantly
slower in sub-$V_T$ operating digital circuits.

The motivation to use analog circuits for decoding has been
based on faster analog parallel and continuous time processing
compared to digital designs [6]. Fewer transistors also
promised energy and area efficient analog decoding circuits.
Therefore, low power analog decoding circuits emerged and
have been existed for more than a decade [7], [8]. Early
analog decoders claimed to provide significant improvements
in consumed power from several times to more than two orders
of magnitude compared to their digital counterparts [9], [10].
Consequently, several analog decoding chips fabricated and
the results have been presented over the last few years [11]–
[16]. However, the benefits of analog decoders tend to degrade
with scaling, since device mismatch has a negative impact on
the bit error rate (BER), which imposes a lower bound on
the decoder’s physical size. As shown in [17], the number of
errors are more significant when the complexity of the decoder
is increased; however, small scale decoders were predicted to
be more resilient to mismatch errors.

While several low power decoder implementations have
been presented in recent years [11]–[16], [18]–[27], there has
been little in-depth investigation based on silicon measure-
ments to evaluate the relative performance and efficiency of
analog versus digital implementations. Especially for low data
rate systems, it is not clear what approach to take to address
the target specifications.

Therefore, in this paper, alternative ultra low power digital
and analog convolutional decoder chips are presented. The
proposed decoders are designed to embed in a custom low-
rate and low power transceiver that is previously presented in
[1]. The target for such a transceiver is to provide short-range
wireless connectivity for applications with low bit rate require-
ments, while consuming less power levels than the available
standards. More powerful codes have a higher decoding com-
plexity, making them less suited for these applications. In this
work, a low complexity convolutional decoder has therefore
been chosen. In the following sections, first the selection of
codes and the corresponding decoding algorithm are presented.
Then follows the design space exploration of critical design
factors for the presented digital and analog decoders. The

The authors are with the Department of Electrical and Informa-
tion Technology, Lund University, SE-221 00 Lund, Sweden (e-mail:
reza.meraji@eit.lth.se).

S. M. Yasser Sherazi is currently with IMEC, Kapeldreef 75 B-3001,
Heverlee, Belgium.
architecture of the decoders together with the fabricated chips are presented in Section V. The silicon measurement results are shown in Section VI, and finally the paper concludes with the evaluation of both digital and analog approaches in terms of chip area, power, BER performance and data rates.

II. TRELIS DECODING OF CONVOLUTIONAL CODES

The proposed decoders are designed for the familiar memory-2 (7,5) convolutional code defined by generator polynomial \( G(D) = [1 + D^2 + D + D^2] \), where \( D \) serves as a delay operator. The power of \( D \) represents the number of time units a bit is delayed. Choosing a convolutional code allows for a relatively short Block Length (BL), which results in small size decoding circuits suited for the mentioned target applications.

The relations between inputs, states and outputs of an encoder can graphically be illustrated by a state diagram referred to as the Trelis. For an encoder with memory \( m \), the trelis representation shows all \( 2^m \) states and all possible transitions between those. Every path in a trelis represents a codeword and the number of stages is the BL of the code. When a trelis is forced to start and end at the same states by proper encoder memory initializations, a circular trelis is formed. This structure, as shown in Fig. 1 for the codes used in this work, is a Tail-Biting (TB) trelis. It is known from theoretical studies that a TB trelis of only 14 sections is needed to decode the (7,5) convolutional code.

The work in [28], known as BCJR decoding after the names of its authors, and also referred to as the forward-backward algorithm, is an efficient procedure based on trelis representation to perform Maximum A Posteriori (MAP) estimations. The algorithm is rather complex, but has received increased practical popularity since the introduction of Turbo codes. TB convolutional codes can be decoded using the BCJR algorithm, in which two recursive clock-wise and counter clock-wise calculations along the trelis are performed to calculate the feedforward and feedback metrics referred to as \( \alpha \) and \( \beta \). The BCJR decoding algorithm estimates the original bit sequence \( u \) by computing the \textit{a posteriori} Log-Likelihood Ratio (LLR) \( L(u_k|y) \) for each single bit, a real number defined by the ratio

\[
L(u_k|y) = \ln \frac{p(u_k = +1|y)}{p(u_k = -1|y)},
\]

where \( y \) is a sequence of \( n \) real values at the input of the decoder. The numerator and denominator of Eq. 1 contain \textit{a posteriori} conditional probabilities; which are probabilities computed after the whole sequence \( y \) is received. The positive or negative sign of \( L(u_k|y) \) indicates which bit, +1 or -1, was coded at time instance \( k \). Its magnitude can be considered as a reliability measure on the decided bit: the larger the magnitude, the more confidence is implied on the estimated bit. This sign and magnitude information of \( L(u_k|y) \) provides \textit{soft information} for each bit that can be applied to the next decoding block, or converted to the corresponding information bit as a hard decision; i.e. if \( L(u_k|y) \) is negative, the decoder will output bit \( u_k = -1 \) and if \( L(u_k|y) \) is positive it will output \( u_k = +1 \).

According to the BCJR algorithm, the posteriori LLR \( L(u_k|y) \) can be written as

\[
L(u_k|y) = \ln \frac{\sum_{s'} T_R s' \gamma_k(s', s) \beta_k(s)}{\sum_{s'} T_R s' \gamma_k(s', s) \beta_k(s)}, \tag{2}
\]

In the above equation, \( s' \) and \( s \) refer to the trelis (encoder) previous state and current state, respectively. In the numerator \( TR1 \) indicates that the summation is carried out over all the state transitions from \( s' \) to \( s \) that are related to message bits \( u_k = +1 \). Similarly, \( TR0 \) in the denominator refers to the set of all transitions originated by message bits \( u_k = -1 \). The channel metric \( \gamma_k(s', s) \) is a conditional probability that is defined by the received signals from the channel. The \( \alpha \) and \( \beta \) metrics are computed recursively around the trelis, as

\[
\alpha_k(s) = \sum_{s'} \alpha_{k-1}(s') \gamma_k(s', s) \quad k = 1, 2, \ldots, BL - 1
\]

\[
\beta_{k-1}(s') = \sum_s \beta_k(s') \gamma_k(s', s) \quad k = BL, BL - 1, \ldots, 2
\]

For \( \alpha_k(s) \), the summation is over all converging branches from previous states \( s_{k-1} = s' \) linked to current state \( s \), while for \( \beta_{k-1}(s') \) the summation is over all states \( s_k = s \) that have links to state \( s \).

III. LOW POWER DIGITAL DECODER BASICS

In a digital decoder, quantized data are used and computations are performed in discrete time, where the speed is limited by the critical path. Multiplications in digital implementations are costly in terms of both area and power. The max-log-MAP algorithm is an approximate realization of the MAP algorithm that provides sub-optimum error performance compared to the MAP based BCJR algorithm. As shown in section VI, this sub-optimum performance is still sufficiently close to that of the original BCJR algorithm for most low power applications. In this max-log-MAP algorithm, the multiplications are replaced by additions,

\[
A_k = \ln[\alpha_k(s)] = \max_{s'}[A_{k-1}(s') + \Gamma_k(s)],
\]

\[
B_{k-1} = \ln[\beta_{k-1}(s')] = \max_s[B_k(s) + \Gamma_k(s')]. \tag{4}
\]
where the capital letters A, B and Γ correspond to the parameters α, β and γ of the BCJR algorithm, expressed in the logarithmic domain. This reduction in complexity reduces the power consumption and chip area significantly in a digital implementation. Memories are normally required to store the temporary data calculations. However, short BL helps to avoid using large memory blocks for temporary storage.

Aside from simplifications to the algorithm, decreasing the supply voltage to sub-$V_T$ is an effective method to lower the power consumption, since the dynamic power decreases quadratically with voltage [29], [30]. However, the circuit will then operate more slowly, increasing the critical path delay and the leakage energy per operation. In order to analyze energy dissipation and critical path delay of a given digital design, gate-level sub-$V_T$ characterization is required. The sub-$V_T$ energy model for standard cell based design presented in [4] has been used for this purpose. A benefit of the analysis is that it locates the energy minimum operating point ($E_{\text{min}}$). With the assumption of operating at maximum frequency at a given supply voltage, it is known that the dynamic energy ($E_{\text{dyn}}$) scales down quadratically with the supply voltage $V_{\text{DD}}$, while the leakage energy per operation increases exponentially. There is a sweet spot for the minimum total energy consumption $E_T$, where the sum of dynamic and leakage energy amounts to a minimum, which is called the energy minimum voltage point (EMV). The EMV is the optimum point in terms of energy per operation which can be used if the data rate requirements are satisfied.

IV. LOW POWER ANALOG DECODER BASICS

In an analog decoder, data is represented by voltage or current. The algorithmic computations for decoding are performed in continuous time [7], [8]; thus, there is no need for temporary storage of intermediate data. The speed of calculations is limited only by the speed of the transistors. Furthermore, the convergence in the iterative decoding algorithm is achieved by settling of transient voltage and currents values after presentation of each new set of received coded data. The final steady state of the currents or voltages in the circuit represents the decoded data, as shown in Fig. 2. The time between two pulses in Fig. 2 shows the allocated time for the circuit to reach a stable state, in which transient waves settle to a value for each output bit above or below the decision threshold.

Analog implementation of a TB trellis decoder results in a circuit with a chip area directly proportional to the size of the trellis. To realize a compact analog decoding circuit a short BL should thus be chosen. Furthermore, transistor sizes have to be chosen carefully in a tradeoff between BER performance and total circuit area.

The analog decoder operates in current mode, making implementation of additions straightforward. Other mathematical operations are implemented based on the exponential relation between drain current and gate-source voltage of MOS transistors in weak inversion (sub-$V_T$). The exponential characteristic is used to convert the received LLR values, $L(u)$, to corresponding probabilities, $p(U)$, represented by currents, throughout the network:

\[
p_{U}(u = -1) = \frac{1}{1 + e^{-L(u)}}, \quad p_{U}(u = +1) = \frac{e^{-L(u)}}{1 + e^{-L(u)}}. \tag{5}
\]

Similarly, the probabilities can be converted back to the logarithmic domain LLR values via diode connected transistors in weak inversion. The analog vector multipliers that are required in the BCJR algorithm can be realized by the Gilbert topology with the transistors operating in weak inversion [31]. Low level of currents is not only necessary for proper weak inversion operation, but also maintain a low power consumption.

V. HARDWARE IMPLEMENTATIONS

A. Digital Decoding Circuit

1) Architecture: The architecture of the proposed digital decoder is presented in Fig. 3. By using the max-log-MAP algorithm, multiplications in the BCJR algorithm are replaced by adders in the logarithmic domain. The max-log-MAP decoding requires calculation of $\Gamma$, $A$, and $B$ parameters and storage over the entire data block due to the forward and backward recursions.

Proper selection of BL is important since it directly affects the error correcting capability of decoder, as well as its chip area and power consumption. Simulations in [32] show that increasing the BL from 8 to 14 in steps of 2, improves the coding gain at BER=$10^{-3}$ by 0.6 dB, 0.25 dB and 0.1 dB respectively. Further increment of BL up to 20 and higher only demands more hardware and higher power consumption, but returns negligible improvement in coding gain; therefore, BL=14 was chosen. For large scale decoding circuits the BL is usually long and there is a need of memory blocks corresponding to it. However, for this design the target is a small scale decoder with a short BL=14; so, register files are used for data storage. The number of iterations around the circular TB trellis and the minimum required word-lengths were determined by high level simulations. It was concluded that by starting at all-zero initial values for A and B metrics, at least two iterations along the trellis are needed to successfully decode the received data. Also, each received soft LLR value is represented by 4 digital bits. Simulations also showed that to benefit from the
full error correcting capability of the algorithm, $\Gamma$ has to be represented by at least 7 bits. Consequently, at least 11-bits are required to cover the full range of $A$ and $B$ values after the iterative decoding calculations. The operation of the decoder is described in the following sub-sections.

- **Input Section**: The digital decoder operates on received blocks of 28 coded soft bits. Hence, the decoding of each block starts with buffering into allocated input registers. After this, all data is moved to another register file for calculation of $\Gamma$ metrics. Simultaneously, buffering of the next block of incoming data starts. For the forward and backward metrics ($A$ and $B$) to be calculated concurrently, the $\Gamma$ calculations are performed from both directions by $\Gamma$-Low and $\Gamma$-High calculation blocks. The allocated $\Gamma$ registers are filled gradually as the computations are performed. Since $BL$ is equal to 14 and each trellis stage has 4 states, $14 \times 4 = 56$ registers are dedicated for $\Gamma$ storage.

- **Iterative Forward-Backward Calculations**: While the $\Gamma$ registers are getting filled in parallel, the calculation of $\alpha$ and $\beta$ starts, as illustrated in Fig. 5. For these calculations to initiate, the start and end $\Gamma$ parameters at the dedicated register block addresses 0-4 and 52-55 have to be available. The rest of the calculations continue step by step. After 14 clock periods, the second iteration starts. On clock cycle 17, all $\Gamma$ values are already calculated and updating the values are no longer required. ’A/B load’ signal refers to the loading time for the $A$ and $B$ metrics to the next stage before decoding of the next block.

- **Decision Section**: The final stage is where the hard decision on the value of each bit is made. Each decision consists of addressing the corresponding register locations, then addition, comparison, and selection operations are performed. A flag signal precedes the starting of each block of decoded bits in the output.

2) **Hardware Mapping of Digital Decoder**: The digital decoder was fabricated in 65 nm CMOS and takes 0.11 mm$^2$ silicon area excluding pads, see die photo in Fig. 4. It has been synthesized with low power standard threshold voltage (LP-SVT) standard cells. LP-SVT proved favorable in a study presented in [4], where the main constraints were maximum throughput, lowest energy dissipation, and a single power domain. Furthermore, tight synthesis constraints were set to achieve minimum area, minimum leakage, and a short critical path at nominal voltage. During place and route, the digital decoder core (DDC) was placed as a separate block together with a peripheral communication core (PCC). The purpose of the PCC is to provide communication between the DDC and the external test environment. The benefit of using PCC is that the DDC can operate at very low voltages, while the outputs remain strong enough for measurements. The connections between these blocks are realized without using level-shifters; rather, buffers are placed in between the two domains for appropriate translation of signal voltages.

**B. Analog Decoding Circuit**

In a conventional digital receiver, baseband processing such as synchronization, filtering and demodulation precedes the channel decoding process. The decoder is designed to be seamlessly embedded in such a receiver.

1) **Architecture**: The detailed architecture of the implemented analog decoders, together with a customized simulation method aiding the choice of design parameters, are
presented in [32], [33] and [34]. The top-level architecture is provided in Fig. 6 for convenience.

The architecture of the analog decoder includes a digital interface as well as the required data converting circuits to fit in a digital receiver. The design consists of the analog decoding core for BL=14, a simple digital interface, an array of 2xBL=28 low resolution current steering digital to analog converters (CS-DACs), and an array of 14 current comparators. The digital circuitry buffers the 2xBL=28 received soft information symbols for each coded block. As an alternative to the digital decoder, the analog architecture is designed to operate on similar input streams. Consequently, each soft information symbol is represented by four bits, hence, a total of 28x4-bit registers are needed for the buffer.

When a complete block has been buffered, it is applied in parallel to the decoding core via an array of 4-bit CS-DACs, for which details are given in [35]. The decoding core works on these data, represented by currents, and generates 14 differential decoded soft output bits. The comparator array translates the soft decoded bits into hard decided bits. The level of the currents in the decoding core can be adjusted by an off-chip variable resistor.

A sample α or β analog multiplying circuit is shown in Fig. 7. This circuit is a hardware representation of a selected butterfly section of the trellis, the one highlighted in Fig. 1. Copies of these blocks are connected together in accordance with the trellis connections. This procedure forms two separate circular circuit arrangements for calculating α and β concurrently. After the current levels have converged to steady state, the outputs are compared by the current comparators to make a hard decision on the value of every bit, i.e. 0 or 1.

A similar multiplying circuit handles γ calculations from the differential logarithmic LLR values as input to the decoder. The γ calculation block takes the values from the CS-DACs, and converts these to the corresponding probabilities by a set of diode-connected transistors operating in weak inversion. The Gilbert configuration calculates these probabilities and applies them to the α and β calculation circuits after duplicating them by a set of PMOS current mirrors. A reference current \( I_{\text{ref}} \), also adjustable by an off-chip resistor, acts as a normalizer to adjust the level of outgoing currents. In addition, \( I_{\text{ref}} \) forces the transistors to remain in weak inversion region, and controls the total power consumption of the decoding core.

The digital interface takes the outputs from the comparators, coordinates the serial streaming of the decoded bits, and handles all the required timing signals, including the time period for the analog core to converge. Except input buffering, which only takes 28x4-bit registers, no other storage is required; i.e. no analog memory is involved.

2) Area: Besides the selection of BL, another important design factor for the chip area is the transistor dimensions. It is thus desirable to find the smallest required device size before the combined effects of mismatch and flicker noise start to deteriorate the BER performance.

To investigate the effects of device mismatch on the performance and accordingly determine the minimum device size required for successful operation, a series of estimations based on statistical simulations were performed in [32]. Following these simulations, NMOS transistor dimensions (W/L) for the three cores were mainly selected as 6.0/0.5, 2.0/0.5 and 1.0/0.5 μm. Selection of shorter length (L=0.2 μm) for the core transistors resulted in severe estimated BER degradation. Therefore, the analog decoding core was fabricated with the three sets of transistor dimensions in Table I. However, as included in Table I, at a few critical places larger transistor were used for better matching. In all three cores, one and two dimensional common centroid layout techniques have been used to improve the matching of current mirrors. The layout of all individual computational blocks was done by hand, whereas all inter-block and higher level routing was performed using automatic routing tools.

3) Timing: The timing of the operations for the decoder is shown in Fig. 8. The total time allocated for the currents in the analog decoding core to settle to their final values, which represent the decoded data, is 24 clock periods, which can easily be adjusted by changing the clock frequency. During this time, the decoder processes the current block of received data, while the input interface buffers the next block of 28 data, one at each clock cycle. Decision time is the time at which the hard decisions are taken based on the output currents. Right after the decision, an output shift register is loaded to stream out the decoded data in serial form.

4) Hardware Mapping of Analog Decoder: The three analog decoders with different decoding core dimensions, AD1, AD2 and AD3, together with the interface circuits were fabricated in CMOS 65 nm process. The design of the interface and mixed signal circuits were kept identical for all the three circuits to support a valid comparison of the decoding cores.

The corresponding chip photos are provided in Fig. 9. For the first chip, Fig. 9(a), the silicon area excluding pads is 0.27 mm\(^2\), of which the analog decoding circuitry, AD1, occupies 0.104 mm\(^2\). The second chip, Fig. 9(b), occupies 0.30 mm\(^2\) without the pads, of which AD2 and AD3 take 0.035 mm\(^2\) and 0.015 mm\(^2\), respectively.
Fig. 7. A sample of Gilbert computation blocks for $\alpha$ and $\beta$ calculations.

Fig. 8. Timing diagram of the decoder.

VI. MEASUREMENT RESULTS

The presented digital and analog circuits were evaluated by measurements of power consumption, BER performance and throughput. The results are provided for the decoding cores only, the focus of this work, since the interfacing blocks are not particularly optimized for power or area. For the analog decoders, however, the power and energy figures include the current of the CS-DACs, since this passes through the decoding cores. Power consumption of the digital interface of the analog decoding circuits is provided in an earlier publication [35], where measurement results of the first analog decoder, AD1, are also presented.

A. Measurement Setup

To generate the required test data, a communication system with BPSK modulation and AWGN channel was considered. A measurement setup including logic analyzer, digital pattern generator, power supplies and high precision digital multimeters was used. Test files were generated in MATLAB for signal-to-noise ratios (SNR) from 1 dB to 6 dB in steps of 1 dB. Measurements were performed in a climate chamber at both room and body temperature to consider the operational environment for the target applications. For each of the designs, either analog or digital, three chip samples were measured. It was noticed that the variation over samples was less significant in case of the analog designs. Digital samples show more variations over chip samples during the measurements, which can be seen in the plots in section VI-B.

B. Digital Decoder

The measured performance of the digital decoder is presented by the four plots in Fig. 10. Minimum energy dissipation is 9 pJ/b at room temperature ($23^\circ C$), whereas it improves slightly to 8 pJ/b at body temperature. Although circuits operated at higher temperature have higher leakage currents [36], they are also faster. Therefore, for a given throughput, the supply voltage at body temperature can be reduced by 30 mV. This reduction in supply voltage results in a slight improvement in energy dissipation.

Furthermore, throughputs are successfully measured from 5 kbps up to 2 Mbps corresponding to supply voltages from 0.25 V to 0.52 V. The corresponding power consumption span is from 0.10 $\mu$W to 25 $\mu$W. Minimum energy dissipation at room temperature, 9 pJ/b, is reached at 0.32 V for a throughput of 125 kbps. Maximum measured throughput, however, is 20 Mbps, which is reached at nominal voltage 1.2 V.

C. Analog Decoder

During measurements of analog decoding circuits the clock frequency was varied from 250 kHz to 1 MHz in steps of 250 kHz and from 1 MHz to 4 MHz in steps of 1 MHz, which corresponds to throughputs from 125 kb/s to 2 Mb/s due to the half rate code. The BER performance of the decoders was measured at a supply voltage of 0.8 V and different power profiles were set by adjusting the current. The current of each decoding core was provided by a dedicated current source array that was adjustable by an off-chip variable resistor.

Figures 11(a) and 11(b) shows the measured energy per decoded bit versus coding gain at BER=$10^{-3}$ for the three
analog decoding cores, AD1 to AD3 at 500 kbps and 2 Mbps, respectively. The coding gains should be compared to the maximum of 3.1 dB for an ideal implementation with BL=14, simulated in MATLAB®.

As shown in Fig. 11(a), for the largest decoding core, AD1, at least 10.5 μW is needed to reach to its maximum 2.3 dB gain at 500 kbps. The energy is dissipated by 20 pJ/b to perform the decoding algorithm at this power. The gain is reduced to 1.75 pJ/B for AD2 due to the overall effects of noise and mismatch errors which are shown to be more destructive in case of AD3.

For a higher data rate of 2 Mbps in Fig. 11(b) it is shown that for AD1, at least 20.6 μW is needed to provide 2.3 dB gain. The coding gain, however, is reduced to 2.0 dB and 1.2 dB for AD2 and AD3 respectively at the same power level. This power level corresponds to about 10 pJ/b energy dissipation. It can be seen in Fig. 11(b) that more or less the same energy is enough for AD2, to reach to its maximum coding gain. Therefore, AD2 with 0.038 mm² silicon area might be a better choice if area has to be traded for reduction in gain from 2.3 dB to 1.9 dB. AD3 is pushed for even smaller area, in which the minimum energy required to reach to 1.2 dB gain is more than 20.6 pJ/b.

The decreasing coding gain trend from AD1 to AD3 relates to the increased mismatch errors for smaller transistors. Degraded gains at lower power levels generally relates to the increased effects of noise on computations.

D. Digital vs. Analog: an analysis

Figure 12 shows the BER performances of AD1, which offers the best coding gain among the three analog decoding cores, together with the performance of the digital decoder for two test cases of 125 kbps and 2 Mbps. This figure also includes software (MATLAB®) simulation performance of the decoder with long BL=100, the performance of the short BL=14 used for implementations in this work, and the expected performance of the uncoded system. At the lower throughput of 125 kbps, the digital decoder offers the desired BER performance with only 1.2 μW power consumption. As presented earlier, this can be achieved at 0.32 V which corresponds to the minimum energy point for the digital decoder. For this rate, the performance of the analog implementation is 0.6 dB degraded compared to that of its digital counterpart, while the power consumption is also significantly higher. At 2 Mbps, consuming only 15.6 μW, AD1 can function as an error correcting block, even though the gain is not at its maximum and degrades to 1.9 dB. At 2 Mbps the minimum required power for the digital circuit is about twice that
power, i.e. 32.4 µW at 0.52 V. Below this supply voltage, the digital implementation is not functional at this rate. While the proposed digital decoder offers a somewhat superior 2.9 dB coding gain at BER=10^{-3}, the analog decoder offers the option of full control over power consumption in trade-offs with the coding gain.

AD1 has an area comparable to the area of the DDC, has higher processing speed for the same power budget, but shows degraded BER performance. The degradation in performance comes from using non-ideal analog multipliers that have limited range of operation, and from effects as device mismatch errors and noise. Following the performance improvement trend from AD3, AD2 to AD1 suggests that increasing the sizes for transistors even further may possibly improve the BER performance. However, in that case while providing higher throughput and coding gain for a lower power, the decoding core area will become larger than the DDC circuitry.

Table II summarizes the performance of the presented decoders together with the previously published analog and digital decoders. Since decoders are usually designed for different applications, energy efficiency in terms of pJ/b has been considered in Table II as a rough indicator to compare the power efficiency of decoders. Figure 13 illustrates the energy efficiency of reported measured decoders in the literature versus technology node of implementation. Both analog and digital implementations, with a variety of code selections, complexity and decoding algorithms are represented in the figure. It is hard to draw a solid conclusion due to the variety of the decoders, but following the trend together with the results from this work suggest an energy efficiency meeting point between analog and digital implementations at about 10 pJ/b in 65 nm CMOS.

**VII. CONCLUSION**

Considering wireless bio-implants and wearable devices, an exploration and comparison of digital and analog implementation alternatives for ultra low power (7,5) convolutional decoding circuits was pursued. While the main focus has been on low power and energy efficiency, other important criteria such as silicon area, throughput, BER performance and temperature variations were studied based on the implemented and measured chips. To push the analog decoder to occupy less silicon area, three versions with different transistor sizes were fabricated and compared, investigating silicon area versus coding gain trade-offs. The digital decoder presented operates at 125 kb/s with 0.32 V supply and dissipates minimum 9 pJ/b energy. The analog decoder chips can function with even smaller supply voltages, operating at 0.3 V and dissipating only 3 pJ/b.

**Fig. 11.** Measured coding gains for AD1,2,3 at 500 kb/s (a), and at 2 Mb/s (b).

**Fig. 12.** BER performance comparisons of the analog (AD1) and digital decoders at (a) 125 kbs and (b) 2 Mbs.

**Fig. 13.** Measured normalized energy per decoded bit evolution vs. technology node for analog: [11]–[16], [20], [37] and digital [18], [19], [23]–[27], [38] decoders.
less than 9 pJ/b while processing faster, but the corresponding coding gains are degraded by more than 1 dB compared to the digital implementation. Considering the complexity in the design process and additional power and area overhead of the presented analog decoders due to the interface circuitry, the sub-V<sub>T</sub> digital approach seems more promising for the tailbiting codes considered in this study.

REFERENCES


Henryk Sjöland received the M.Sc. degree in electrical engineering from Lund University, Sweden, in 1994, and the Ph.D. degree from the same university in 1997. In 1999 he was a Postdoc at UCLA on a Fulbright scholarship. He has been an associate professor at Lund University since year 2000, and a full professor since 2008. Since 2002 he is also part time employed at Ericsson Research, where he is currently a Research Fellow. He is heading the research group in Radio Frequency Integrated Circuit Design at Lund University, and he has authored or co-authored more than 100 international peer reviewed journal and conference papers and holds patents on more than 10 different inventions. Henryk Sjöland is an associate editor of IEEE Transactions on Circuits and Systems II, and he is a member of the Technical Program Committee of the European Solid-State Circuits Conference (ESSCIRC). He is a Senior Member of IEEE. He has co-supervised 9 PhD-students and received their degrees, and is currently the main supervisor of 6 PhD students. His research interests include design of radio frequency, microwave, and mm wave integrated circuits, primarily in CMOS technology.

John B. Anderson was born in New York State in 1945. He received the Ph.D. in electrical engineering from Cornell University in 1972. During 1972-80 he was on the electrical engineering faculty at McMaster University in Canada, and during 1981-98 he was on the faculty at Rensselaer Polytechnic Institute. Since 1998 he has held the Ericsson Chair in Digital Communication at Lund Univ., Sweden. He has held visiting professorships at Berkeley, Chalmers Univ. (Sweden), Queens Univ. (Canada), Deutsche Luft- und Raumfahrt (Germany), and Tech. Univ. of Munich. He was Director of the Swedish Strategic Research Foundation Center for High Speed Wireless Communication at Lund during 2005-11. His research work is in coding and communication algorithms, bandwidth-efficient coding, and data compression. He has served widely as a consultant in these fields.

Viktor Öwall was an Associate Editor of the IEEE TRANS. CIRCUITS AND SYST.-II: ANALOG AND DIGITAL SIGNAL PROCESSING from 2000-2002 and of the IEEE TRANS. CIRCUITS AND SYST.-E: REGULAR PAPERS from 2007-2009. He was Guest Editor for the Special issue on ISCAS 2010 of the IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.