

This is the post peer-review accepted manuscript of:

Giovanni Rovere; Schekeb Fateh; Luca Benini; A 2.2- $\mu$ W Cognitive Always-On Wake-Up Circuit for Event-Driven Duty-Cycling of IoT Sensor Nodes, in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Year: 2018 , Volume: 8 , Issue: 3, Pages: 543 – 554, DOI: 10.1109/JETCAS.2018.2828505

The published version is available online at: <https://doi.org/10.1109/JETCAS.2018.2828505>

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

# A 2.2 $\mu$ W Cognitive Always-On Wake-Up Circuit for Event-Driven Duty-Cycling of IoT Sensor Nodes

Giovanni Rovere, *Student Member, IEEE*, Schekeb Fateh, *Member, IEEE*,  
and Luca Benini, *Fellow, IEEE*

**Abstract**—We report an always-on event-driven asynchronous wake-up circuit with trainable pattern recognition capabilities to duty-cycle power-constrained internet-of-things (IoT) sensor nodes. The wake-up circuit is based on a level-crossing analog-to-digital converter (LC-ADC) employed as feature-extraction block with automatic activity-sampling rate scaling behavior. A novel asynchronous digital logic classifier for sequential pattern recognition is presented. It is driven by the LC-ADC activity and trained to minimize classification errors due to falsely detected events. As proof-of-concept, a prototype of the wake-up circuit is fabricated in 130 nm CMOS technology within 0.054 mm<sup>2</sup> of active area, covering up to 2.6 kHz of input signal bandwidth. The prototype has been first validated by interfacing it with a commercial accelerometer to classify hand gestures in real-time, reaching 81% of accuracy with only 2.2  $\mu$ W at 1 V supply. To highlight the flexibility of the design, a second application, detecting pathologic ECG beats is also discussed.

**Index Terms**—Wake-up circuit, LC-ADC, ultra-low-power, event-driven, asynchronous, clockless, supervised learning, classifier, training.

## I. INTRODUCTION

INTERNET-OF-THINGS sensor nodes are key interfaces to the physical world enabling sense-making and insight-extraction from the sensed data. For the IoT sensor nodes to reach the foreseen 10<sup>12</sup> market volume units in the next decade, researchers and industry are urged to come up with innovative solutions that meet tight and multi-dimensional constraints [1]. High-precision and continuous acquisition are both desirable features of IoT sensor nodes, as well as low cost and integrated form factor to enable large scale deployment and ultimately commercial profitability [2]. However, as target data are often sparse in time, considerable energy is wasted in acquiring and processing uninformative data, preventing to reach the  $\mu$ W range of targeted power consumption. Even though tunable designs [3] may allow to trade power for precision on the fly, still the high-precision system (HPS) power consumption lower-bound is constrained by the hardware overhead for worst-case operating condition.

Alternatively, a common approach dynamically duty-cycles the HPS to a power-saving mode, e.g., to a sleep mode or

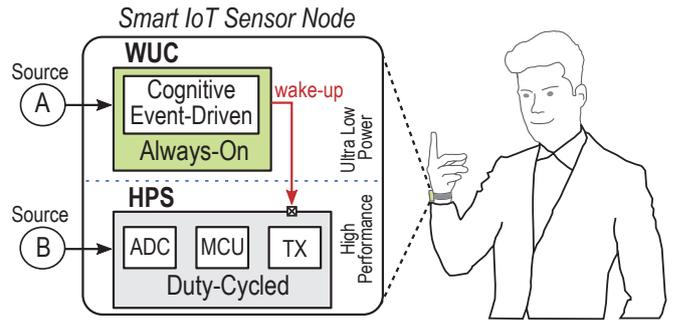


Fig. 1. The presented cognitive computing wake-up circuit (WUC) employed in the context of smart IoT sensor node for wearable applications. The high precision system (HPS) is normally in a power saving mode, and wake-up once an event of interest demanding high-precision is detected by the always-on wake-up circuit.

to power-gating mode, when high precision is not needed. The toggling of the HPS mode is controlled by an always-on wake-up circuit (WUC) responsible to wake-up the HPS from saving mode when full precision data is required, as illustrated in Fig. 1. However, a simple threshold-based WUC that only compares the input signal to a voltage level is often not suitable to discriminate the inception of data patterns. As commonly implemented in commercial products [4]–[7], threshold-based WUC may yield into many false positive (FP) events and ultimately in high power consumption as the HPS is woken up unnecessarily, as depicted in Fig. 2. Hence, more advanced signal processing must be embedded close to the sensor to correctly identify the events of interest, positive (P) events, while ignoring the others, negative (N) events. Being always-on, substantial effort is devoted to lower the power consumption of the WUC by exploiting approximate computation in the analog-to-digital conversion (ADC). The reduced precision, traded for high energy efficiency is recovered by the digital data-driven learning algorithm used to train the WUC parameters.

### A. Related Works

Circuits with advanced classification capabilities that could be used as WUCs only recently appeared in scientific literature in search of innovative ways to lower HPS based IoT sensing nodes power consumption [8]–[11]. Compared to this work, most implementations are based on alternative architectures and are designed to cover a single specific application only. For instance, the authors in [8] present a heartbeat detector for waking-up a wearable healthcare system. To increase prediction capabilities the computation is performed by a coarse-fine

Manuscript received December 15, 2017. Date of current version December 15, 2017.

This project was funded by the SNSF under the grants: 157048 *Transient Computing Systems* and 162524 *MicroLearn: Micropower Deep Learning* which are financed by the Swiss Confederation and scientifically evaluated by SNSF. The project is also financed by the *Oprecomp EU Project*.

G. Rovere and L. Benini are with the Dept. of Information Technology and Electrical Engineering (D-ITET), ETH Zurich, 8092 Zurich, Switzerland (e-mail: {rovere, benini}@iis.ee.ethz.ch). S. Fateh is with Miromico AG, Zurich, Switzerland (e-mail: fateh@miromico.ch).

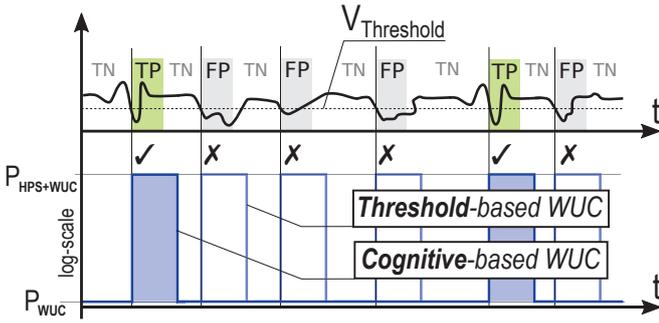


Fig. 2. Poor discrimination capabilities in simple threshold-based WUC may lead to several misdetected events (FP) that increase the IoT sensor node power consumption,  $P_{HPS}$ , by unnecessarily waking-up the HPS. By embedding cognition in the WUC, better discrimination allows isolating positive events (TP) only.

short-term autocorrelation and a template matching technique. However, although the algorithm is robust, it is application specific, and would require a complete revision to classify differently shaped signals. In [9] a speech/non-speech classification system for automatic voice-activity-detection (VAD) is proposed, band-filtered component energies are used as signal features and classified by a decision tree model with off-chip analog parameters. However, the feature extraction block is tailored to process speech signals, preventing to use the system in other application domains. The authors of [10] adopt a classic signal processing approach where a low-power ADC digitizes the microphone output signal, whereas a subsequent digital logic performs both feature extraction (DFT) and classification (support vector machine); by that only stationary signals can be distinguished, e.g., electrical generator, a small car and a truck. The architecture at most similar to the one presented in our work is reported in [11], where an event-driven circuit for wearable electrocardiogram (ECG) based on a "moving-window" LC-ADC implementation is presented. However, the training is not machine learning assisted and the classifier model supports ECG QRS-detection only, thus lacking the flexibility to cover different IoT applications [1].

### B. Contributions

In this paper, we propose an ultra-low-power (ULP), always-on, asynchronous WUC with flexible classification capabilities, demonstrated on real-world signals, to efficiently duty-cycle HPS in IoT applications. Due to the intrinsic real-world signal variability, e.g., caused by the source-to-source mismatch or same-source variations over time, embedding cognition in the WUC increases the classification capabilities in case of alike signals. An ad hoc training algorithm automatically configures the WUC based on the samples in the considered dataset, without requiring it to be explicitly programmed. This approach further relaxes analog precision requirements as analog non-idealities are included in the training loop and compensated for. To practically embed cognition in the WUC, we propose 1) a flexible and programmable classifier model which can be used to real-time discriminate the patterns of interest, 2) an efficient hardware implementation enabling ULP always-on operating modes, and 3) an ad hoc automatic data-driven training method. The proposed mixed-signal architec-

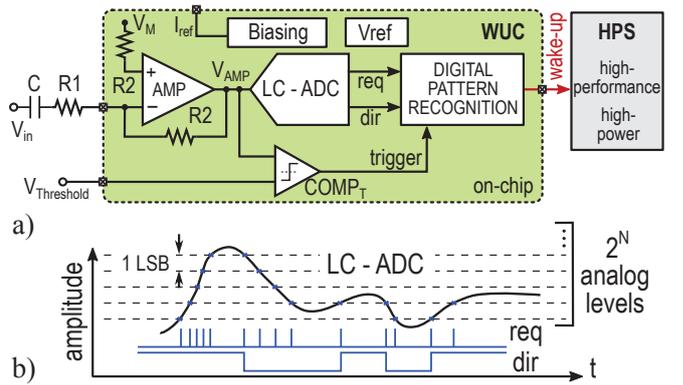


Fig. 3. (a) Architecture of the presented ultra-low-power pattern recognition WUC with on-chip references. The input signal is preamplified, (b) delta-encoded ( $req$ ,  $dir$ ) by a LC-ADC, and its matching pattern used to wake-up an external asleep HPS.

ture consists of a LC-ADC [12], [13], for energy-proportional signal preprocessing and feature extraction, followed by an asynchronous trainable digital pattern recognition circuit to achieve energy efficiency and activity-power proportionality. The presented WUC efficiently duty-cycles the HPS, waking it up once a pattern of interest has been detected, enabling full-precision processing in the HPS, otherwise remaining in ultra-low-power sleep mode. This approach breaks the IoT sensing node energy-quality trade-off, which is inherent in traditional duty-cycling.

This paper is based on and extends the work presented in [14]. The key contributions of this manuscript are: 1) the combination of a LC-ADC with a digital state machine to perform real-world signal feature extraction and pattern recognition classification; 2) increased energy efficiency of the LC-ADC by including on-chip references; 3) an efficient ad hoc supervised-learning algorithm to train the classifier; 4) the interfacing of the WUC with a commercial accelerometer to perform hand-gesture recognition and 5) the use of the WUC in a pathologic ECG contraction pattern recognition application. The paper is organized as follows. Section II reveals the system architecture of the WUC, whereas the circuit implementation is detailed in Section III. Section IV presents the ad hoc training algorithm for the classifier model. In Section V the chip measurement results are reported and discussed. In Section VI the two applications are presented. Conclusions and remarks are drawn in the last Section VII.

## II. COGNITIVE WAKE-UP CIRCUIT ARCHITECTURE

The WUC architecture is depicted in Fig. 3a, where the signal path consist of 1) a preamplifier, 2) a LC-ADC for non-uniform sampling analog-to-digital conversion and feature extraction, and 3) a digital pattern recognition (DPR) block for binary pattern classification. The LC-ADC outputs a delta-modulated sample, i.e., request and direction pair ( $req$ ,  $dir$ ), once one of the internal analog levels are crossed by the preamplified signal  $V_{amp}$ . The number of the internal analog levels is  $2^N$ , where  $N$  is the resolution of the LC-ADC. A pulse on the request line,  $req$ , marks a level-crossing (LC)

and the direction line,  $dir$ , is updated according to the  $V'_{amp}$  derivative sign, Fig. 3b. The LC-ADC can be seen as equivalent to a combination of multiple comparators with 1LSB spaced levels, similar to a flash-ADC architecture, but with higher efficiency given the lower number of required comparators, i.e., 2 instead of  $2^N$ . It is thus conceivable to expect better discrimination capabilities on the input signal from such a LC-ADC WUC compared to the simple single comparator threshold-based approach. Furthermore, as opposed to uniform sampling ADCs, the LC-ADC exhibits several interesting properties [15]–[17], such as: 1) no amplitude quantization noise, 2) alias-free output spectrum, and 3) intrinsic data-compression and frequency-information removal. The latter is desirable when the signal information resides in the shape only, for example in hand-gesture recognition where the gesture execution speed is often not relevant for the classification task. Similarly, detection of pathologic ECG contractions is independent on the actual heart pulse frequency.

A sequence of delta-modulated samples of the LC-ADC are segmented, digitally processed, and sequentially compared against programmable digital thresholds by the DPR block which implements the binary classifier. If the target pattern stored in the DPR block matches the pattern sought for in the input signal, then a wake-up line is asserted which can be used to awake a duty-cycled HPS, triggering high-precision signal processing. The distinctive WUC cognitive capabilities originate from the tight and seamless combination of the LC-ADC with a custom DPR model designed to process LC-ADC samples without requiring a periodic clock signal.

The comparator  $Comp_T$  triggers the DPR block only once there is sufficient signal energy at the input, synchronizing the inception of the event with the DPR state transitions. False negative (FN) events that may arise due to simple misalignment between the input pattern and the target pattern in the DPR block are reduced compared to always enabled DPR. A preamplifier has been added in front of the signal processing chain to allow easy interfacing of different sensors with wide amplitude range and to correctly drive the LC-ADC input load. The value of components  $R_1$  and  $C$  determines the preamplification gain and the high-pass cut-off frequency. Furthermore, contrary to other implementations [9], [15], all the needed analog voltage references have been integrated on-chip to reduce external dependencies, requiring only a single positive supply (0.9 V - 1.2 V) and a current-reference ( $\approx 100$  nA) for the prototype to operate properly. Even though this WUC architecture is generic and could be used in many applications, in this work we focus on always-on accelerometer-based hand gesture recognition and pathologic ECG classification. The WUC is designed to support up to 2.6 kHz of analog input bandwidth within  $\mu$ W of power consumption.

Before using the WUC for a classification task, as depicted in Fig. 4, preliminary training is required to "teach" the DPR digital thresholds to identify the targeted signal. During the training phase, labeled positive and negative signals are fed through both the preamplifier and the LC-ADC. The generated delta-modulated samples are acquired and fed to the off-chip training algorithm that infers a set of optimum DPR parameters, i.e., digital thresholds. In a second step, efficient real-

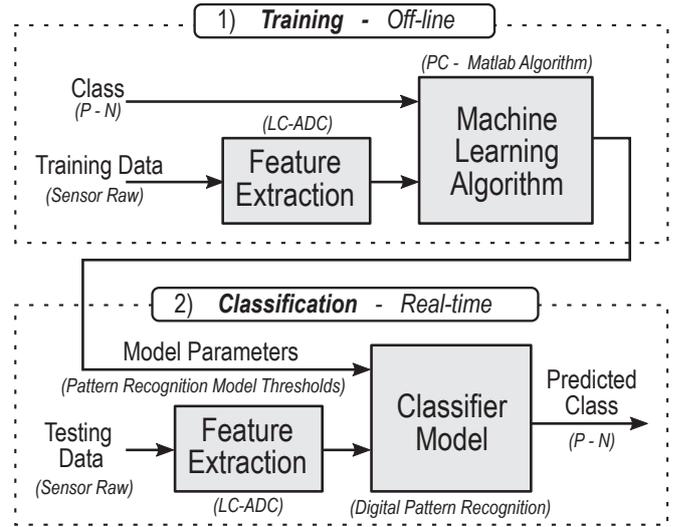


Fig. 4. A supervised-learning approach is used to increase the WUC discrimination capabilities. During the training phase, raw data are first LC-ADC digitized and collectively off-line post-processed in Matlab to optimally choose DPR digital thresholds. During the classification phase, new unlabeled data are real-time processed by the trained WUC to discriminate incoming signals.

time classification is performed by the DPR on unlabeled, LC-ADC digitized signals. The hardware required for this phase is fully contained in the WUC prototype, hence no additional computational units are required, allowing deployment of the system even in IoT constrained environments.

### III. CIRCUIT IMPLEMENTATION

#### A. Level-Crossing ADC

The conversion efficiency of early LC-ADC implementation [11] is mainly limited 1) by the N-bit resistor-ladder digital-to-analog converter (DAC) and 2) by comparators with rail-to-rail input stage capabilities. To achieve higher energy efficiency, a pseudo-asynchronous LC-ADC based on dynamic comparators and a 40 pF analogue memory cell has been proposed in [18]. However this architecture assumes the availability of a clock signal for correct operation, which is not desirable in an always-on self-contained WUC. In this paper the employed LC-ADC architecture is inspired by [15] as it combines highest energy efficiency (219 - 565 fJ/conv.) and clock-less implementation. This architecture is known in the literature as "fixed window" LC-ADC and greatly contributes to increase the energy efficiency by constraining the comparator input voltage swing by 2LSB. We further improved the design by adding on-chip charge-sharing voltage references to lower the system power consumption and reduce the wake-up circuit dependencies to external components. The proposed architecture is depicted in Fig. 5a and consists of a 1-bit DAC tracking the input signal  $V_{amp}$ , two continuous-time comparators for level crossing detection and an asynchronous digital logic for state control.

The continuous-time 1-bit DAC tracks the input signal  $V_{amp}$  on  $V_{DAC}$ . A  $\pm 1$ LSB offset is injected by charge-sharing on  $V_{DAC}$  from one of the two precharged branches  $V_{BL}$  and  $V_{BR}$

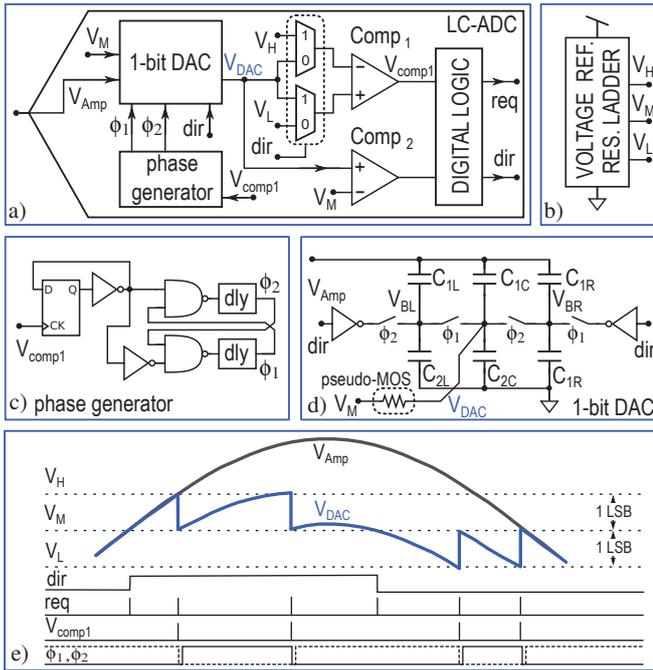


Fig. 5. (a) LC-ADC block diagram; (b) on-chip voltage reference generation; (c) asynchronous driven non-overlapping phase generator controlling the 1-bit DAC; (d) charge-sharing 1-bit DAC with integrated references, pseudo-MOS large resistor sets  $V_{DAC}$  DC value; (e) a realistic LC-ADC timing diagram, where the  $req$  signal marks any of the thresholds crossing, while the  $dir$  signal indicates the derivative of the input signal.

once an analog level ( $V_L$ ,  $V_H$ ) is crossed. The offset direction controlled by  $Comp_2$  is opposite to the derivative sign of the input signal. For example, during phase  $\phi_2$ ,  $V_{amp}$  is tracked on the right and central branches ( $C_{1R}$ ,  $C_{2R}$ ,  $C_{1C}$ ,  $C_{2C}$ ), while the left branch intermediate node  $V_{BL}$  is charged to digital logic levels according to the  $\overline{dir}$  signal. For every  $V_{comp1}$  pulse, as a consequence of  $V_H$  or  $V_L$  level-crossing, the  $\phi_1$  and  $\phi_2$  states toggle disconnecting the left branch  $V_{BL}$  and connecting the right branch  $V_{BR}$  to the central one,  $V_{DAC}$ , eventually altering its value by  $\pm 1$  LSB due to charge conservation in the 1-bit DAC. This mechanism is then repeated alternatively for each branch and the  $V_{DAC}$  node is thus always constrained to be within  $\pm 1$  LSB of  $V_M$  via feedback.

Ultimately, a short-pulse on the  $req$  line is asserted every time one of the three thresholds ( $V_H$ ,  $V_M$  or  $V_L$ ) is crossed, while the  $dir$  encodes the signal derivative, providing a delta-modulating encoding of the input signal. A representative time evolution of the LC-ADC is illustrated in Fig. 5e. To increase the 1-bit DAC matching, capacitances have been implemented as parallel connection of several  $C_U = 40$  fF unit-capacitors, of which multiplication factor satisfies the precharge time condition  $t_{precharge} = 1/(2^N \pi f_{max})$  on nodes  $V_{BL}$  and  $V_{BR}$ . Where  $N$  is the LC-ADC resolution and  $f_{max}$  is the maximum input signal frequency. The  $V_{DAC}$  node voltage  $V_{amp}$  is attenuated by 14/15 whereas its DC level is set at  $V_M$  by a pseudo-MOS resistor ( $\approx 70$  G $\Omega$ ). This finite resistance leaks charge from  $V_{DAC}$  resulting in a constant signal drift over time and thus in a signal distortion. To guarantee charge conservation in the 1-bit DAC, a phase generator has been

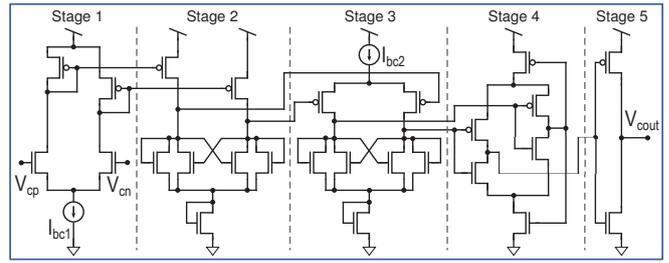


Fig. 6. LC-ADC ultra-low-current comparator schematic. Post-layout simulations show propagation delay of 5.2  $\mu$ s given 108 nA current consumption at 1 V supply.

implemented to derive non-overlapping phase control signals triggered by  $V_{comp1}$  activity, Fig. 5c.

The 1-bit DAC right and left branch nodes,  $V_{BL}$  and  $V_{BR}$ , are alternatively precharged to either VDD or GND levels by digital inverters as in [19] to enable efficient on-chip reference instead of by analog buffers as in [15]. The total harmonic distortion (THD) limits the LC-ADC precision due to 1) 1-bit DAC capacitor array mismatch, 2) comparator non-idealities and hysteresis, 3) reference offsets, and 4)  $V_M$ - $V_{DAC}$  pseudo-MOS connection. Higher power consumption is usually required to reduce THD by employing larger components or by active cancellation techniques. However, at system level these effects are not an issue since distortion errors can be compensated during the training process.

## B. Comparator

The comparator speed is the dominant limitation of the LC-ADC input bandwidth that can be handled without incurring in signal slewing, i.e., slope overload [15]. Assuming a full scale sine wave, the maximum input signal frequency  $f_{MAX}$  is defined as follows:

$$f_{MAX} = [2^N \cdot (t_{DAC} + t_{Comp} + t_{Logic})]^{-1} \quad (1)$$

where  $t_{DAC}$ ,  $t_{Comp}$ , and  $t_{Logic}$  respectively represents the propagation delay of the 1-bit DAC, the comparator and the digital logic; the sum of all these delay components defines the LC-ADC loop delay. Note that for input signals with smaller amplitudes, the maximum input frequency can be proportionally higher than  $f_{MAX}$ . Simulation in the target technology shows that  $t_{DAC}$  and  $t_{Logic}$  are both in the nano-second range. Hence  $t_{Comp}$  dominates in the  $f_{MAX}$  equation mandating  $t_{Comp} < 5$   $\mu$ s to let the LC-ADC handle 2 kHz range input signals. Whereas the comparator output swing has to be rail-to-rail to guarantee digital logic compatibility and avoid short-circuit currents of poorly driven digital stages, the input signal range is sufficient to be  $V_M \pm 1$  LSB, greatly helping satisfying the 150 nW of power consumption allocated budget.

To achieve these specifications the multi-stage architecture shown in Fig. 6 is employed. The first stage provides only little gain but it is essential to reduce decision time and further helps isolating the  $V_{DAC}$  sensitive node from kickback noise. The second and third stages provide high gain by cross-connected NMOS transistors, whose ratios also define the amount of hysteresis the comparator exhibits. The fourth stage requires

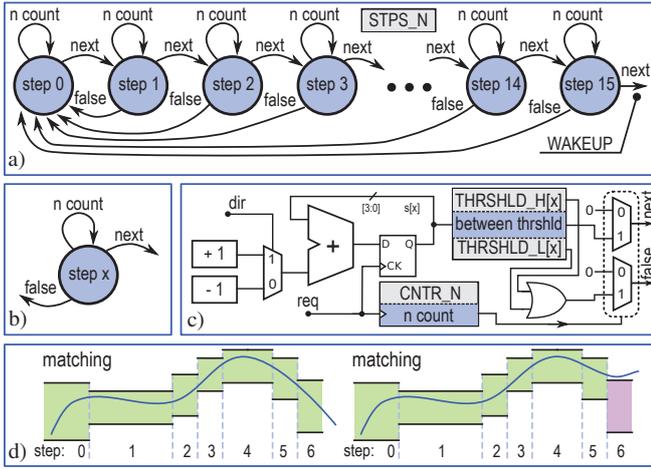


Fig. 7. (a) The full 16-states of the digital pattern recognition implemented as finite state machine; (b) abstraction of the pattern recognition single step; (c) implementation detail of the pattern recognition step.  $dir$  signals are summed  $CNTR\_N$  times and compared against two threshold levels; (d) a successful (left) and non-successful (right) pattern matching plausible scenarios are depicted, where different steps lengths are due to different input slopes.

no static bias as the input signal is already in the hundred of millivolts range. It consist of two parallel digital inverters: one drives the next stage while the other disables the current in this stage paths from VDD or to GND, reducing static power. The fifth stage is required to fully restore logic levels to be compliant with the subsequent digital logic.

In this work a comparator hysteresis value of 45 % is chosen to effectively suppress noise-induced fluctuations when the input is close to the threshold. High values of hysteresis will cause signal distortion but only when the input signal derivate changes sign. Even comparator delay dispersion resulting from different input amplitudes will contribute to distortion since the propagation time is dependent on the comparator differential voltage. However, as already emphasized, distortion of the input signal is not a concern in this work WUC as long as a data-driven learning mechanism is employed. Extensive PVT post-layout transient simulations have been conducted, confirming  $t_{Comp}$  to range from 4.3  $\mu$ s to 7.7  $\mu$ s.

### C. Digital Pattern Recognition

To efficiently process the asynchronous LC-ADC delta-modulated data without requiring expensive time quantization, an event-driven DPR block with clockless operation is desirable. Analog/mixed-signal and full-digital asynchronous processors have been reported in [20] and [21], both targeting general-purpose high-performance computation resulting in excessive power (4 mW and 65 mW) and area consumption (51.4 mm<sup>2</sup> and 4.3 cm<sup>2</sup>) and are thus not suitable for integration in the proposed WUC. To enable always-on operation, a  $\mu$ W-range power consumption is required, hence a minimalistic design based on an asynchronous FSM that matches the incoming signal with a stored pattern is proposed.

The implemented DPR block consists of  $STPS\_N$  identical steps that are feed-forward connected as in Fig. 7a, each processing a  $CNTR\_N$  long segment of the delta-modulated

samples. State transitions from  $step\_k$  lead the system to either advance to  $step\_k+1$  if the segment matches, or to the initial state  $step\_0$ . The system evolution is asynchronously triggered by the  $req$  signal falling edges and the wake-up line is eventually asserted when all segments correspond in all the  $STPS\_N$  states, indicating a full pattern match. The pattern recognition state is detailed in Fig. 7c, consisting of a block segmenting a sequence of  $CNTR\_N$  elements and computing the algebraic sum:

$$s[x] = \sum_{n=1}^{CNTR\_N} [2 \cdot dir(n) - 1]$$

Signal  $s[x]$  is then compared against two digital thresholds  $THRSHLD\_L[x]$  and  $THRSHLD\_H[x]$ , both individually programmable for each step. If the condition  $THRSHLD\_L[x] \leq s[x] \leq THRSHLD\_H[x]$  is met, the segment matches and trigger the transition to  $state\_k+1$ , otherwise the system resets to  $state\_0$ . The pattern recognition parameters, i.e.,  $STPS\_N$ ,  $CNTR\_N$ ,  $THRSHLD\_L[STPS\_N-1:0]$  and  $THRSHLD\_H[STPS\_N-1:0]$ , can be independently set to maximize the pattern recognition accuracy. In the current implementation the limit is 16 for  $STPS\_N$  and  $CNTR\_N$ , and both  $THRSHLD\_L$  and  $THRSHLD\_H$  ranges from -15 to +15 allowing odd-only values. Since the pattern is evaluated in a sequential fashion and energy efficiency is paramount in always-on circuits, pattern recognition states share most of the hardware depicted in Fig. 7b, where  $THRSHLD\_L[STPS\_N-1:0]$  and  $THRSHLD\_H[STPS\_N-1:0]$  values are sourced from a look-up table. Even though the pattern recognition is asynchronous, limited complexity and slow operation speed,  $\mu$ s, allowed us to still use standard digital EDA tools for circuit synthesis and physical implementation of the DPR block.

### D. Preamplifier

An on-chip amplifier is often required to handle small amplitude signals and to drive the 1-bit DAC input capacitance (280 fF), though not directly contributing to the WUC classification functionality. As long as distortion is not a concern, a single-ended architecture is employed, not requiring additional common-mode feedback regulation circuit. A Gain-Bandwidth-Product of 200 kHz is required to provide significant gain (40 dB) in the 2 kHz bandwidth while power consumption must be within few  $\mu$ W.

A power-efficient push-pull rail-to-rail output architecture is presented in Fig. 8, where both the amplification and the biasing structures are reported. The first stage is a cascode structure ( $M_1-M_7$ ), followed by a class AB output stage ( $M_8-M_9$ ). Transistors  $M_{10}-M_{13}$  set a stable bias current for the output stage, of which  $M_{12}-M_{13}$  implement a floating current source biased with  $M_{11}$ ,  $M_{13}$ ,  $M_{15}$ ,  $M_{16}$ ,  $M_{10}$ ,  $M_{12}$ ,  $M_{17}$ ,  $M_{18}$ . A positive injected current from the first stage on the  $V_A$  node causes both  $M_8-M_9$  gates voltages to increase, sinking current from the  $V_{amp}$  output. Similarly, a negative current on the  $V_A$  node charges the load. Even though the floating current source transistors  $M_{12}-M_{13}$  have source connections on nodes  $V_A$  an  $V_B$ , the loading effect is limited due to

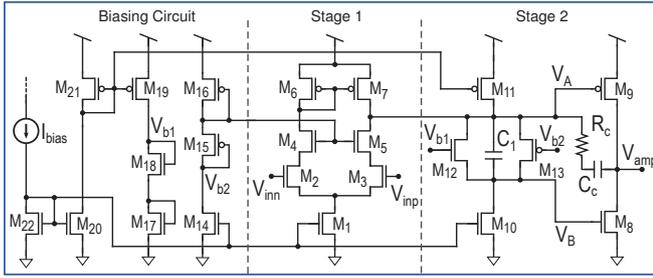


Fig. 8. On-chip class-AB preamplifier schematic including the biasing circuit.

feedback in  $M_{12}$ - $M_{13}$  [22], thus high DC gain is preserved. The low frequency open-loop gain of the preamplifier,  $A_{ol}$ , can thus be approximated as:

$$A_{ol} \approx [g_{m3} ((g_{m5} r_{ds5} r_{ds3}) || r_{ds7})] [(g_{m8} + g_{m9}) \cdot (r_{ds8} || r_{ds9})]$$

Capacitor  $C_1$  contributes to keep the class-AB output stage biased when the effect of the floating current source is reduced due to bandwidth limitations. Adding a compensation capacitor  $C_c$  helps to stabilize the amplifier at high frequency while  $R_c$  moves the zero to the left-half-plane.

#### IV. DIGITAL PATTERN RECOGNITION TRAINING ALGORITHM

A data-driven algorithm is usually employed to achieve optimal classification performance by automatically configuring the DPR parameters to 1) model the target signal traits while 2) ignoring sample-to-sample intrinsic changes due to source variability or random noise. As the DPR classifier model presented in this work is a novel design, we developed an ad hoc off-line algorithm to train the classifier. First a theoretical description of the classifier model is presented in Subsection IV-A, then the actual two-phase training algorithm is detailed in Subsection IV-B and in Subsection IV-C.

##### A. Digital Pattern Recognition (DPR) Classifier Model

After the LC-ADC performs feature extraction by LSB delta-modulating the input signal, a sequence of  $CNTR\_N$  samples ( $req, dir$ ) are algebraically summed,  $s[0]$ , in the DPR step 0, as described in Section III-C. Similarly, a  $s[x]$  feature is repeatedly computed for each of the  $1 \dots STPS\_N$  steps of the DPR, encoding a different time portion of the input signal. The resulting features are grouped in a vector  $\bar{s} = (s_1, s_2, \dots, s_{STPS\_N}) \in \mathbb{S}$  mapping each element of the dataset from the time domain to the features set.

The space  $\mathbb{S}$  of all possible vectors modeled by the DPR is the  $STPS\_N$ -dimensional subset of odd numbers defined as:

$$\mathbb{S} = \left\{ \begin{array}{l} \{2m - 1 : m \in \mathbb{Z}\}^{STPS\_N} \\ -\frac{CNTR\_N-1}{2} < m < \frac{CNTR\_N-1}{2} \end{array} \right.$$

where similar input signals are encoded by similar vectors. The digital thresholds of the DPR model partition  $\mathbb{S}$  in a positive class set defined as:

$$\mathbb{S}_P = \{\bar{s} : THRSHLD\_L[x] \leq s_x \leq THRSHLD\_H[x]\}$$

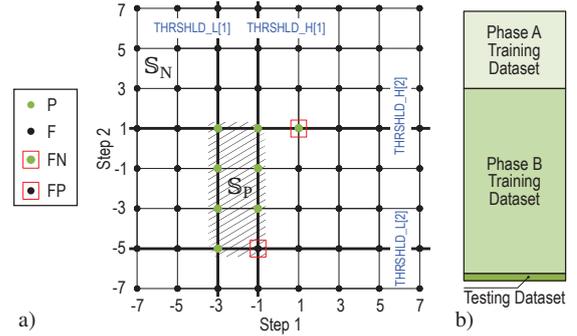


Fig. 9. (a) Illustration of the possible vectors space. Thick lines represent the DPR thresholds defining the  $\mathbb{S}_P$  hyperrectangle (highlighted fill), complementary to the  $\mathbb{S}_N$  set. Green dots are positive events while black dots are negative events. (b) Dataset partition.

and in a negative class set,  $\mathbb{S}_N = \mathbb{S} - \mathbb{S}_P$ . The  $\mathbb{S}_P$  is a hyperrectangle wherein belonging vectors are labeled as positive class,  $\bar{s}^* \in \mathbb{S}_P$ , otherwise as negative class. A graphical representation of the classifier model for  $STPS\_N = 2$  and  $CNTR\_N = 3$  is depicted in Fig. 9a.

Since the  $\mathbb{S}_P$  hyperrectangle boundaries have high impact on the classifier discrimination performance their selection is of paramount importance. Even though systematic exhaustive-search in the whole feature-space parameter guarantees optimum hyperrectangle boundary selection, it is often practically not feasible to iterate over

$$N_{search} = \left[ \frac{(CNTR\_N + 1) \cdot (CNTR\_N + 2)}{2} \right]^{STPS\_N}$$

possible combinations. However, most of the  $N_{search}$  configurations either have no meaning in the model, e.g.,  $THRSHLD\_H[x] < THRSHLD\_L[x]$ , or are by far off from the optimum, thus could be easily discarded without risking overlooking the optimal configuration point. Aiming both solution coverage of exhaustive-search and computational efficiency, a cascaded two phases (*A* and *B*) training algorithm has been developed. *Phase A* identifies a subset of potential configurations from the dataset, marking them for subsequent use and ignoring the others. *Phase B* iterates over all the previously found configurations to find the best with an exhaustive search within a short time, thanks to the reduced number of possible combinations. The training dataset is thus partitioned as well in *phase A training data* and *phase B training data*, as shown in Fig. 9b.

##### B. Phase A – Training Algorithm

From the training set, only positive-labeled vectors (P) are used in *phase A* to preselect a set of potential optimal hyperrectangles. From this set, Fig. 10a, the number of occurrences of each vector element is annotated, indicating what the most likely values are that each vector element should have to be identified as positive vector. This can be graphically depicted with a histogram plot, Fig. 10b; the values and distributions herein are chosen for illustration purposes, assuming a 20 vector dataset. The number of occurrences, i.e., the histogram bars, can be either higher (solid fill) or equal-lower (line fill)

than the `OUTLIER_LEVEL`, indicating that vector elements whose occurrences are above this level must be included in the hyperrectangles, otherwise they may be omitted, as potential outliers. For instance, the first vector element,  $s_1$ , assumes 16 times the value  $-3$ , twice the value  $-1$  and only once the values  $+1$  and  $+3$ ; indicating that the target vector is likely to have the element  $s_1 = -3$ . To account for this consideration in the DPR, thresholds must be selected such as the  $THRSHLD\_L[1] \leq -3 \leq THRSHLD\_H[1]$  condition holds. However, as this may be a too restrictive condition yielding to poor generalization performance, one must adapt the thresholds to consider also less probable occurring values. A set of potential low and high thresholds respectively is thus the following:  $(-3;-3)$ ,  $(-3;-1)$ ,  $(-3;+1)$ ,  $(-3;+3)$  as shown in Fig. 10c; reducing the number of possible thresholds combinations from 10 to 4. A specular conclusion can be drawn for the second vector element,  $s_2$ , as well. However, if more than one vector element is above the `OUTLIER_LEVEL`, as for  $s_3$ ,  $s_4$  and  $s_5$  in the depicted example, the hyperrectangles must be chosen such that it includes all of these elements. As a result, at the end of *phase A*, a set of thresholds pairs is defined for each vector element, listing potential DPR threshold levels for each DPR step.

### C. Phase B – Training Algorithm

During *phase B*, a set of hyperrectangles is first generated by iterating over all the possible threshold combinations found in *phase A*. For each hyperrectangle, all the *phase B* vectors are fed to the DPR and the hyperrectangle that shows the best classifier performance is selected. To quantify classifier performance during the training, several metrics can be computed from the raw data confusion matrix, including sensitivity, specificity, precision and accuracy (all defined in Table III). In this work the Matthews Correlation Coefficient (MCC) is adopted to allow classifier performance comparison in presence of unbalanced-class dataset [23]. The MCC is defined as follows:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{P \cdot N \cdot (TP + FP)(TN + FN)}} \quad (2)$$

where a MMC of 1 represents a perfect prediction, 0 is equivalent to a random guess and  $-1$  an always wrong prediction. To practically evaluate the classifier model generalization performances, a Leave-One-Out-Cross-Validation (LOOCV) statistical technique is adopted, where the classifier has been tested with a dataset element not used during the training phase.

## V. PROTOTYPE MEASUREMENT RESULTS

The proposed WUC has been fabricated in a 130 nm 1P8M mixed-signal CMOS technology with MiM-capacitor option running from a single supply (0.9 V-1.2 V). The prototype is encapsulated in a QFN56 package and mounted on a PCB with auxiliary circuits, i.e., a linear supply regulator, a current reference and a MCU, to allow convenient prototype configuration and low-bandwidth data readout, e.g., the DPR wake-up output line. The used nano-ampere current source

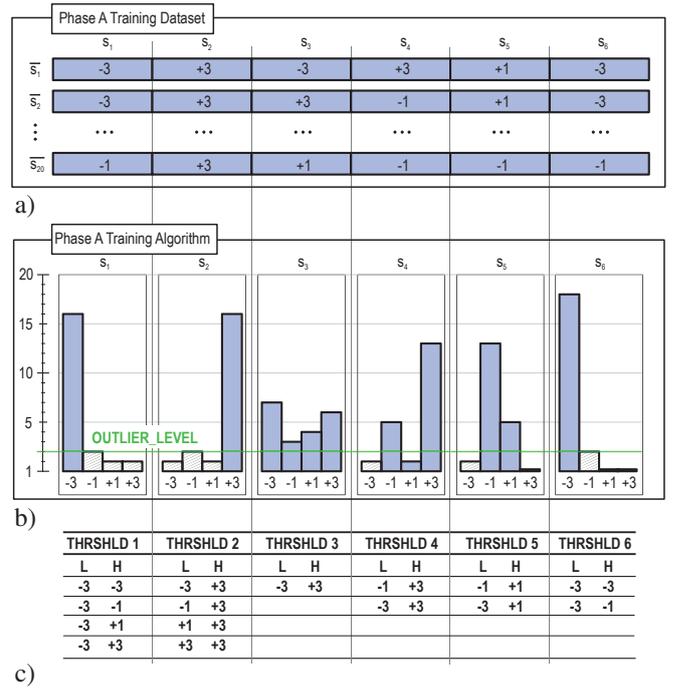


Fig. 10. (a) Positive-labeled dataset vectors used in *phase A* training; (b) histogram representing the number of occurrences for each vector element; (c) for each vector element a set of thresholds has been selected which includes high frequent vector element but may neglect low frequent ones.

design is described in [24] and employed to externally control the bias current of the prototype during testing. However, as absolute precision of the current reference is not required, an on-chip current source can be integrated with negligible (2.6 nW) power consumption overhead [25].

To guarantee faithful signal acquisition for the LC-ADC characterization, the delta-modulated signals  $req$  and  $dir$  are recorded with a Rohde & Schwarz RTO1024 oscilloscope with digital probes at 10 MS/s and subsequently post-processed in MATLAB to reconstruct the input signal. The same setup has been used for acquiring samples of the training dataset employed in Section IV. However, as the time-stamp information is not relevant for the classification and thus for training, a simple general-purpose MCU operating at low clock frequency, could be alternatively employed for delta-modulated signal acquisition.

The WUC active area is approximately  $0.054 \text{ mm}^2$ , of which 36% are occupied by the LC-ADC, 23% by the voltage references ( $V_H$ ,  $V_M$ ,  $V_L$ ), 21% by the DPR, 11% by the pre-amplifier, and 4% by the comparator  $COMP_T$ . The remaining area is used for biasing circuitry, and signal/power routing.

The delta-modulated signals  $req$  and  $dir$  are shown in Fig. 11 for a sinusoidal test signal (blue line). The signal can thus be reconstructed via *spline* interpolation and used for the LC-ADC characterization by computing the FFT. The type of interpolation has been shown in [15] to have only little impact on the SNDR performance. Out-of-band signals cause *slope overload* in the LC-ADC due to excessive signal slope between two 1LSB-spaced levels and eventually causing distortion. Even though high distortion is not a concern in our WUC

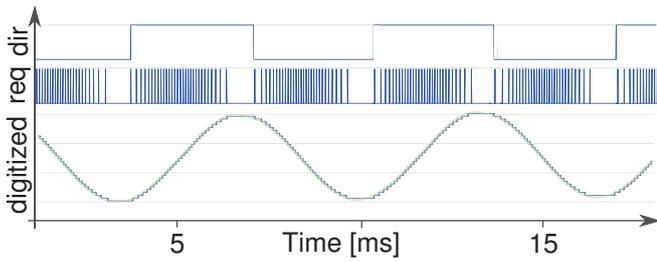


Fig. 11. Measured delta-encoded (*req*, *dir*) LC-ADC output signals (top) and the "spline" interpolated reconstructed signal (bottom). The digitized signal is slope-overload free up to 2.6 kHz.

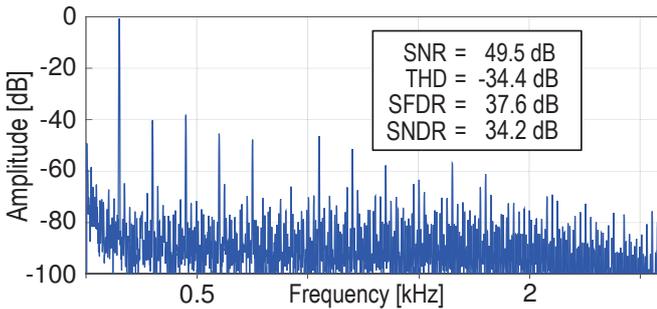


Fig. 12. FFT of the measured LC-ADC output for a full-scale (1 V) 150 Hz reconstructed sinusoidal signal.

because its negative effect can be counteracted by the training algorithm, the recover time from slope overload is directly proportional to the out-of-band signal frequency, hence it must be avoided to prevent missing any relevant signal nuances. It is also worth noting that for LC-ADC the bandwidth is not an appropriate metric, as the tracking capability is proportional to the slope of the signal, which depends on both frequency and amplitude. Hence, to characterize the LC-ADC, the performance loop delay would be a more appropriate metric. However, conversion between loop-delay and frequency is straightforward assuming a full scale sinusoidal signal, Eq. 1.

From the reconstructed signal, a standard FFT can be computed and performance metrics calculated as shown in Fig. 12. The SNDR of the LC-ADC is clearly limited by the THD whose sources are highlighted in Subsection III-A. By including the preamplifier in the signal path, the in-band noise floor rises by 35 dB as shown in Fig. 13. The preamplifier gain plot and the power-precision trade-off are shown respectively in Fig. 14a-b. A comparison of our LC-ADC implementation with state-of-the-art (SOA) designs is reported in Table I. To allow a fair comparison, the on-chip references contribution has been excluded from the current consumption of this work, as other authors assumed off-chip voltage references are available.

Always-on operation mandates ultra-low-current which in our case is 310 nA, the lowest compared to SOA, achieved by optimizing the comparator design for low current operation as well as introducing switch-capacitor dynamic references in the 1-bit DAC. The measured distortion in our prototype is higher than [15], however besides different optimization goals, their implementation requires off-chip analog thresholds which have been hand trimmed to compensate for the LC-ADC

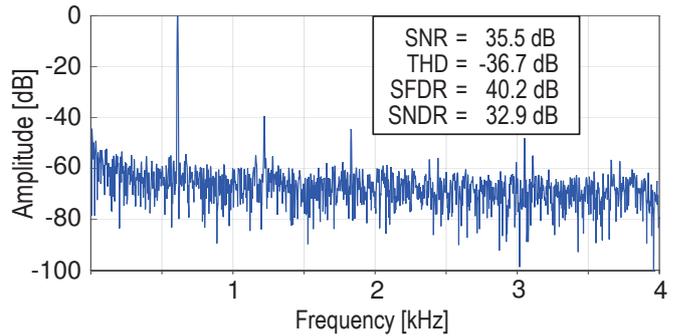


Fig. 13. FFT of the signal path (LC-ADC and preamplifier) output for a full-scale 600 Hz reconstructed sinusoidal signal.

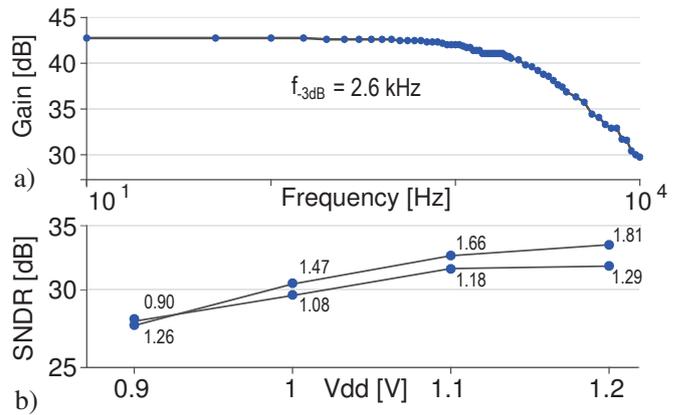


Fig. 14. Measured preamplifier (a) frequency response and (b) precision-power trade-off plots. Power consumption, in  $\mu\text{W}$ , is annotated next to data points for two bias conditions.

on-chip non-linearities. As we target stand-alone operation mode, individual trimming of analog external components is not desirable, as it makes it difficult to deploy the solution in real-life applications. To fairly compare different LC-ADC designs a figure-of-merit is introduced in [18] defined as:

$$\text{FoM} = \frac{\text{Power} \cdot \text{Area}}{f_{\text{Nyquist}} \cdot 2^{\text{ENOB}}}$$

where  $f_{\text{Nyquist}}$  is the Nyquist sampling frequency for the LC-ADC input bandwidth. Comparing different FoM values reported in Table I, it emerges that "fixed window" LC-ADC architectures, like the one implemented in our WUC and in [15], yield to high power efficiency within a given area and are thus suitable for always-on operation.

## VI. WUC APPLICATION EXAMPLES

### A. Accelerometer-Based Hand-Gesture Pattern Recognition

To validate the proposed WUC architecture the fabricated prototype has been evaluated in the context of an accelerometer-based hand-gesture recognition application. The *finger snapping* gesture is selected as target class of interest (positive class), because of its potential use as trigger event for data acquisition in wearable form factors, e.g., healthcare devices like smartwatches. *Finger snapping* can be easily discriminated against low energy gestures, e.g., writing, keyboard typesetting, handshaking, touchscreen navigation, even by a

TABLE I  
COMPARISON OF STATE-OF-THE-ART ULTRA-LOW-POWER LC-ADC

|   | TBCAS 13<br>[15]         | TBCAS 14<br>[11] | TBCAS 17<br>[18]         | <b>This<br/>Work</b>    |
|---|--------------------------|------------------|--------------------------|-------------------------|
| Process [ $\mu\text{m}$ ]                                 | 0.18                     | 0.13             | 0.35                     | <b>0.13</b>             |
| Supply [V]  | 0.8                      | 0.3              | 1.8 - 2.4                | <b>1.0</b>              |
| Current [ $\mu\text{A}$ ]                                 | 0.39 - 0.70 <sup>a</sup> | 0.73             | 0.33 - 0.83 <sup>b</sup> | <b>0.31<sup>c</sup></b> |
| Resolution [bits]   | 6                        | 5                | 4 - 8                    | <b>5</b>                |
| Freq. [kHz]   | 3.3                      | 1                | 1                        | <b>2.6</b>              |
| SNDR [dB]   | 40 - 49                  | 28.3             | 37 - 48                  | <b>34.2</b>             |
| ENOB [bit]  | 6.35 - 7.84              | 4.4              | 5.85 - 7.68              | <b>5.39</b>             |
| Area [ $\text{mm}^2$ ]                                    | 0.045                    | 0.360            | 0.037                    | <b>0.020</b>            |
| FoM [ $\frac{\text{fJ}}{\text{conv}} \cdot \text{mm}^2$ ] | 10 - 25                  | 1876             | 42                       | <b>28</b>               |

<sup>a</sup> Off-chip voltage references excluded.

<sup>b</sup> Off-chip digital logic and voltage references excluded, measured at different supply.

<sup>c</sup> Excluding on-chip references.

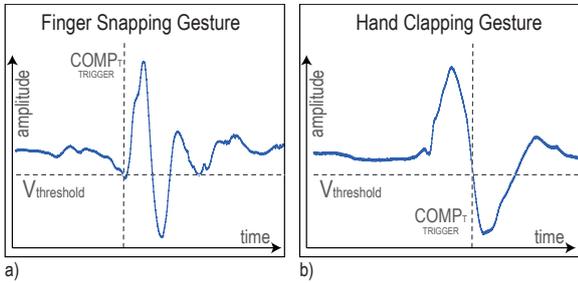


Fig. 15. Analog output of a wrist-mounted accelerometer while performing (a) finger snapping gestures and (b) hand clapping gesture.

simple thresholds-based approach. However, it becomes more challenging to keep the number of FP events low when comparable energy gestures happens, e.g., hand clapping or other activities that may normally occur during the subject daily activity. Examples of preamplified ( $V_{\text{amp}}$  node), accelerometer based hand-gestures waveforms for *finger snapping* and *hand clapping* are depicted in Fig. 15.

To generate a dataset useful for the WUC training and testing a commercial accelerometer (*Analog Devices ADXL326* [26]) is firmly attached on the right wrist of the subject asked to perform the defined gestures. The analog output of the accelerometer is connected to the WUC prototype for signal amplification, digitization and classification, as described in Sec. II. Two different gestures with similar energy content, i.e., *finger snapping* and *hand clapping*, are considered. For each of those, a 100-gestures sample dataset has been acquired before and after digitization (signals  $V_{\text{amp}}$  and  $req, dir$ ), together with the gesture label. Even though the subject was instructed to perform identical gestures, same-class and class-to-class amplitude differences are still visible. Note that these have not been normalized to replicate a realistic use-case scenario. In the randomly reshuffled dataset 20 samples are used for the Training Phase A, whereas 79 samples for the Training Phase B. This division has been empirically chosen to yield high classification performance within just few minutes of training time, though exact partition is not critical. The remaining sample (1 out of 100) has been

TABLE II  
CONFUSION MATRIX OF THE DPR BINARY CLASSIFIER

|            |              | Ground Truth |              |
|------------|--------------|--------------|--------------|
|            |              | Positive (P) | Negative (N) |
| Prediction | Positive (P) | 87 (TP)      | 25 (FP)      |
|            | Negative (N) | 13 (FN)      | 75 (TN)      |

TABLE III  
PERFORMANCE SUMMARY OF THE DPR BINARY CLASSIFIER

| Metric      | Definition            | Value |
|-------------|-----------------------|-------|
| Sensitivity | TP / P                | 87%   |
| Specificity | TN / N                | 75%   |
| Precision   | TP / (TP + FP)        | 78%   |
| Accuracy    | (TP + TN) / (P + N)   | 81%   |
| F1          | 2TP / (2TP + FP + FN) | 82%   |
| MCC         | Equation (2)          | 62%   |

used in the classifier testing phase. To implement the LOOCV, the classifier has been trained and tested 100 times each with a different partitioned dataset allowing averaging the classifier generalization performance over the whole dataset.

The classification accuracy is 81%, and performance details are reported in Tables II and III. The power consumption of the operating WUC is partitioned as follow: 877 nW for the LC-ADC, references and bias distribution, 173 nW for the DPR, 107 nW for the  $\text{COMP}_T$  and 1.05  $\mu\text{W}$  for the preamplifier, yielding approximately 2.2  $\mu\text{W}$  of overall chip power consumption. The employed analog-output accelerometer has a current consumption of 350  $\mu\text{A}$ . To further decrease the IoT node power consumption, a low-power accelerometer can be alternatively employed, e.g., the 832-0025 of TE Connectivity [27]; with 4  $\mu\text{W}$  of power consumption but with lower sensitivity (50 mV/g) as well. The latter can be recovered by increasing the closed loop gain of the preamplifier by reducing the resistor value  $R_1$  shown in Fig. 3.

To compare the cognitive-based WUC presented in this work with respect to a threshold-based WUC, a numeric example is here presented. An estimate of the power saving achieved by duty-cycling the HPS is computed assuming the following conditions: 1) the HPS consumes 10 mW [28] and 2) it's woke up for 60 s whenever a predicted positive event is identified. 3) For every hour, there are 1 positive event and 5 negative events and 4) a timespan of 24 h is considered. Given that, the average system power consumption is 1 mW for the case of a threshold-based WUC, as it labels all the events as positive, whereas 377  $\mu\text{W}$  if employing the cognitive-based WUC, yielding to an overall power saving of approximately 2.65x.

### B. Pathologic ECG Pattern Recognition

To practically demonstrate the flexibility of the proposed WUC, we report an additional experiment of pattern recog-

TABLE IV  
COMPARISON OF REAL-TIME WUC STATE-OF-THE-ART

|                           | JETCAS 11<br>Rumberg et al. [29]       | TBCAS 15<br>Izumi et al. [8] | JSSC 16<br>Badami et al. [9]                | JSSC 17<br>Jeong et al. [10] | <b>This<br/>Work</b>   |
|---------------------------|--|------------------------------|---|------------------------------|--|
| Architecture              | Filter Bank +<br>Power Level Detection | ADC + DSP                    | Filter Bank +<br>Decision Tree <sup>b</sup> | ADC + DSP                    | <b>LC-ADC + DPR</b>  |
| Process [ $\mu\text{m}$ ] | 0.5                                    | 0.13                         | 0.09  | 0.18                         | <b>0.13</b>  |
| Area [ $\text{mm}^2$ ]    | 2.25                                   | 6.8 <sup>a</sup>             | 2   | 0.75                         | <b>0.054</b>   |
| Supply [V]                | N.A.                                   | 1.2                          | N.A.  | 1.2                          | <b>1</b>   |
| Bandwidth [kHz]           | 10                                     | 0.7 (128 Sps)                | 5   | 0.47                         | <b>2.6</b>   |
| Power [ $\mu\text{W}$ ]   | 51                                     | 1.02                         | 6   | 0.0122                       | <b>2.2</b>   |
| Application               | Stationary Sounds                      | QRS Detection                | Voice / No-Voice                            | Stationary Sounds            | <b>Hand Gestures (Accelerometer)<br/>Pathologic ECG Patterns</b> |

<sup>a</sup> Estimated number.

<sup>b</sup> Assuming 7 off-chip tunable analog references to configure the circuit.

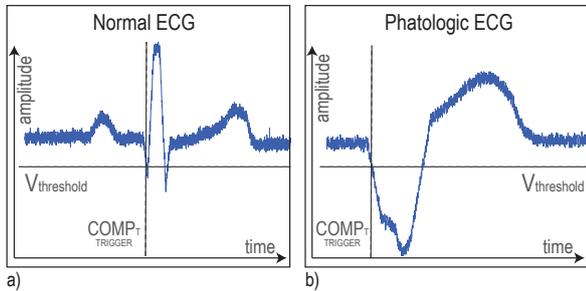


Fig. 16. Amplified ECG signal waveforms for (a) normal contractions (b) pathologic PVC contractions.

nition classification in the context of biomedical signals [14]. In this experiment, the WUC prototype input is connected to a Fluke prosim 8 vital signs patient simulator that models an ECG signal track exhibiting isolated premature ventricular contractions (PVCs). Amplified raw data are shown in Fig. 16. To accommodate the smaller amplitude range of the ECG signal, the gain of the preamplifier has been increased by 32 dB.

Even though single PVCs are often benign events, series (> 3) of PVCs lasting less than 30 s are called non-sustained ventricular tachycardia (NSVT), which may be asymptomatic but known to anticipate the often fatal ventricular fibrillation [30]. Hence, as only a series of events is dangerous, the WUC is configured to reduce unnecessary wake-up of the HPS by minimizing false positive (FP) events whereas relaxing constraints on false negative (FN) events. Experimental data shows that the WUC classification performance on single ECG beats is: TP = 37, TN = 49, FN = 29 and no FP events, yielding to 74.8 % of accuracy and 59.3 % of MMC. However, the probability to miss a series of 3 PVCs is only 8 % and quickly drops for longer-lasting (and thus more dangerous) NSVT episodes.

A comparison with state-of-the-art ultra-low-power prototyped solutions is reported in Table IV. However, profoundly different architectures and optimization goals make a fair comparison a challenging task. Nevertheless, the power consumption of our WUC is lower than other solutions with the exception of [8], [10]. The paper [10] reports better power

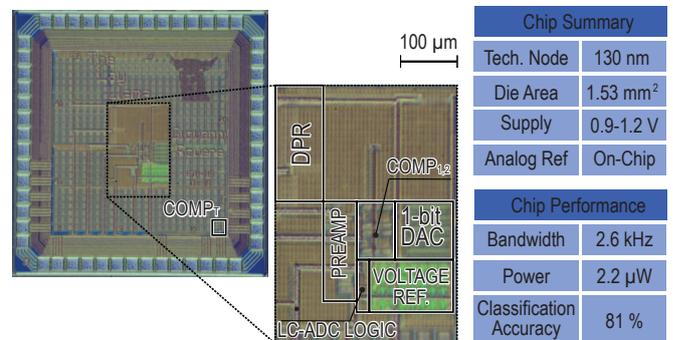


Fig. 17. Chip micrograph of the fabricated prototype with summary of the features.

and classification performances, however their work is limited to stationary signals discrimination up to 470 Hz. Similarly, the [8] implementation allows for low-frequency ECG heart-beat detection only. At least an order of magnitude less in terms of required silicon area, together with the demonstrated flexibility of our architecture, elect our design as potential candidate for duty-cycling HPS in IoT sensor nodes.

## VII. CONCLUSIONS

In this paper we reported the design of an ultra-low-power event-driven WUC in 130 nm CMOS technology. The chip micrograph is depicted in Fig. 17. The proposed WUC can be used to dynamically trade energy for quality of a HPS by duty-cycling it to a power saving mode, waking it up when high-precision acquisition is demanded.

To the best of our knowledge, for the first time a LC-ADC has been combined with a programmable DPR asynchronous classifier to implement an event-driven mixed-signal WUC with real-time pattern recognition capabilities. The DPR classifier parameters are optimally configured on the considered dataset with an ad hoc data-driven off-line training algorithm. The algorithm captures the relevant traits of a target signal while discarding irrelevant information for the classification task. The WUC high power efficiency is achieved by non-uniform sampling, on-chip voltage references and approximate

computing allowing relaxed analog design, since noise and distortion non-idealities are compensated for by the training algorithm.

In comparison with other state-of-the-art WUCs, the flexibility and generality of this design allows using it with many different sensors and signals as well as a wide variety of wake-up trigger patterns. To show the effectiveness of the proposed WUC, it has been demonstrated in an accelerometer based hand-gesture recognition application, reaching 81% of classification accuracy; and in pathologic ECG classification, reaching 74.8% of classification accuracy. The fabricated prototype has been tested and consumes only 2.2  $\mu\text{W}$  of power at 1 V supply and requires few non-critical off-chip components, thus making it suitable for extended-lifetime IoT sensing node applications.

#### ACKNOWLEDGMENT

The authors gratefully acknowledge the support of P. Schönle, T. Burger, J. Bösser, M. Lanz, T. Kleier; as well as ACP AG, Zollikon, CH for chip fabrication and the ETH Microelectronic Design Center for EDA support.

#### REFERENCES

- [1] M. Alioto, "Enabling the Internet of Things: From Integrated Circuits to Integrated Systems," Springer, 2017.
- [2] C. Perera, A. Zaslavsky, P. Christen and D. Georgakopoulos, "Context Aware Computing for the Internet of Things: A Survey," *IEEE Commun. Surveys Tut.*, vol. 16, no. 1, pp. 414–454, January 2014.
- [3] P. Schönle et al., "A Multi-Sensor and Parallel Processing SoC for Wearable and Implantable Telemetry Systems," *Proc. IEEE European Solid-State Circuits Conf. (ESSCIRC)*, pp. 215–218, September 2017.
- [4] Analog Devices, "Micropower, 3-Axis, 2g /4g /8g Digital Output MEMS Accelerometer," *ADXL362 datasheet*, 2017.
- [5] NXP Semiconductors, "MMA8453Q, 3-Axis, 10-bit/8-bit Digital Accelerometer," *MMA8453Q datasheet*, 2017.
- [6] STMicroelectronics, "MEMS Digital Output Motion Sensor: Ultra-Low-Power High-Performance 3-Axis "Nano" Accelerometer," *LIS3DH datasheet*, 2017.
- [7] PUI Audio - Vesper, "Low-Noise Bottom Port Piezoelectric MEMS Microphone with Wake on Sound Feature," *PMM-3738-VM1010-R datasheet*, 2017.
- [8] S. Izumi et al., "Normally Off ECG SoC With Non-Volatile MCU and Noise Tolerant Heartbeat Detector," *IEEE Trans. Biomed. Circuits and Syst.*, vol. 9, no. 5, pp. 641–651, October 2015.
- [9] K. M. Badami, S. Lauwereins, W. Meert, and M. Verhelst, "A 90 nm CMOS, 6  $\mu\text{W}$  Power-Proportional Acoustic Sensing Frontend for Voice Activity Detection," *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 291–302, January 2016.
- [10] S. Jeong, Y. Chen, T. Jang, J.M. Tsai, D. Blaauw, H. Kim, and D. Sylvester, "Always-On 12-nW Acoustic Sensing and Object Recognition Microsystem for Unattended Ground Sensor Nodes," *IEEE J. Solid-State Circuits*, 2017.
- [11] X. Zhang and Y. Lian, "A 300-mV 220-nW Event-Driven ADC with Real-Time QRS Detection for Wearable ECG Sensors," *IEEE Trans. Biomed. Circuits and Syst.*, vol. 8, no. 6, pp. 834–843, December 2014.
- [12] W. Tang et al., "Continuous Time Level Crossing Sampling ADC for Bio-Potential Recording Systems," *IEEE Trans. Circuits and Syst. I, Reg. Papers*, vol. 60, no. 6, pp. 1407–1418, June 2013.
- [13] N. Ravanshad, H. Rezaee-Dehsorkh, R. Lotfi and Y. Lian, "A Level-Crossing Based QRS-Detection Algorithm for Wearable ECG Sensors," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 1, pp. 183–192, January 2014.
- [14] G. Rovere, S. Fateh, and L. Benini, "A 2.1  $\mu\text{W}$  Event-Driven Wake-Up Circuit Based on a Level-Crossing ADC for Pattern Recognition in Healthcare," *IEEE Biomed. Circuits and Syst. Conf. (BioCAS)*, pp. 332–335, October 2017.
- [15] Y. Li, D. Zhao, and W. A. Serdijn, "A Sub-Microwatt Asynchronous Level-Crossing ADC for Biomedical Applications," *IEEE Trans. on Biomed. Circuits and Syst.*, vol. 7, no. 2, pp. 149–157, April 2013.
- [16] F. Marvasti, "Nonuniform Sampling: Theory and Practice," *Springer Science & Business Media*, 2012.
- [17] K. Kozmin, J. Johansson and J. Delsing, "Level-Crossing ADC Performance Evaluation Toward Ultrasound Application," *IEEE Trans. Circuits and Syst. I: Reg. Papers*, vol. 56, no. 8, pp. 1708–1719, August 2009.
- [18] T. Marisa et al., "Pseudo Asynchronous Level Crossing ADC for ECG Signal Acquisition," *IEEE Trans. Biomed. Circuits and Syst.*, vol. 11, no. 2, pp. 267–278, April 2017.
- [19] P. Harpe, E. Cantatore and A. van Roermund, "An Oversampled 12/14b SAR ADC with Noise Reduction and Linearity Enhancements Achieving up to 79.1 dB SNDR," in *IEEE Int. Solid-State Circuits Conf. (ISSCC)*, February 2014.
- [20] Q. Ning, H. Mostafa, F. Corradi, M. Osswald, F. Stefanini, D. Sumislawska and G. Indiveri, "A Reconfigurable On-Line Learning Spiking Neuromorphic Processor Comprising 256 Neurons and 128K Synapses," *Frontiers in neuroscience*, vol. 9, no. 141, April 2015.
- [21] F. Akopyan et al., "Truenorth: Design and Tool Flow of a 65mW 1 Million Neuron Programmable Neurosynaptic Chip," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 34, no. 10, pp. 1537–1557, October 2015.
- [22] R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation*. John Wiley & Sons, 2010, vol. 3rd Edition.
- [23] S. Boughorbel, F. Jarray and M. El-Anbari, "Optimal Classifier for Imbalanced Data Using Matthews Correlation Coefficient Metric," *PLoS one*, vol. 12, no. 6, June 2017.
- [24] Linear Technology, "Precision Dual/Quad CMOS Rail-to-Rail Input/Output Amplifiers," *LTC6081/LTC6082 datasheet*, 2017.
- [25] L. Magnelli, F. Crupi, P. Corsonello, C. Pace and G. Iannaccone, "A 2.6nW, 0.45V Temperature-Compensated Subthreshold CMOS Voltage Reference," *IEEE J. Solid-State Circuits*, vol. 46, no. 2, pp. 465–474, February 2011.
- [26] Analog Devices, "Small, Low Power, 3-Axis 16g Accelerometer," *ADXL326 datasheet*, 2017.
- [27] TE Connectivity, "Model 832 Accelerometer," *832-0025 datasheet*, 2017.
- [28] S. Benatti et al., "A Sub-10mW Real-Time Implementation for EMG Hand Gesture Recognition based on a Multi-Core Biomedical SoC," *IEEE Int. Workshop on Advances in Sensors and Interf. (IWASI)*, 2017.
- [29] B. Rumberg, D.W. Graham, V. Kulathumani, and R. Fernandez, "Hibernets: Energy-Efficient Sensor Networks Using Analog Signal Processing," *IEEE J. Sel. Topics in Circuits and Syst.*, vol. 1, no. 3, pp. 321–334, September 2011.
- [30] B. A. Koplán and W. G. Stevenson, "Ventricular Tachycardia and Sudden Cardiac Death," *Elsevier Mayo Clinic Proceedings*, vol. 84, no. 3, pp. 289–297, March 2009.



**Giovanni Rovere** (S'17) received the M.S. degree in electrical engineering from the University of Padua, Padua, Italy, in 2013. In 2014 he was an asynchronous digital designer at the Italian Institute of Technology, Italy, designing a ser-des circuit for event-driven vision sensor for robotic applications. He was a visiting student at the Institute of Neuroinformatics, ETH, Zurich in 2013 and at Cornell University, Ithaca, NY, USA in 2014. He is currently working towards the Ph.D. degree at the Integrated Systems Laboratory, ETH, Zurich. His

current research focuses on analog/mixed-signal efficient computation for pattern recognition and on system-on-chip for biomedical applications.



**Schekeb Fateh** (M'17) received the M.S. degree in electrical engineering from the Eidgenössische Technische Hochschule (ETH) Zurich, Switzerland, in 2009. In the same year, he joined the Integrated Systems Laboratory (IIS) of ETH Zurich, where he obtained the Ph.D. degree. His research interests include the design of high-speed analog and mixed-signal circuits and systems for wireless communications with emphasis on analog-to-digital converters as well as the implementation of multi-purpose low-power and miniaturized biomedical devices for IoT

healthcare applications. Currently, he is senior design engineer at Miromico. He received (jointly with Dr. Studer and Dr. Seethaler) the Swisscom and ICTnet Innovations Award 2010 for the work on "VLSI Implementation of Soft-Input Soft-Output Minimum Mean-Square Error Parallel Interference Cancellation".



**Luca Benini** (F'07) holds the chair of digital Circuits and systems at ETHZ and is Full Professor at the Universit di Bologna. His research interests are in energy-efficient system design for embedded and high-performance computing. He is also active in the area of energy-efficient smart sensors and ultra-low power VLSI design. He has published more than 800 papers and five books. He is a Fellow of the ACM and a member of the Academia Europaea. He is the recipient of the 2016 IEEE CAS Mac Van Valkenburg award.