



HAL
open science

Analog-to-feature converter optimization through power-aware feature selection

Antoine Back, Paul Chollet, Olivier Fercoq, Patricia Desgreys

► **To cite this version:**

Antoine Back, Paul Chollet, Olivier Fercoq, Patricia Desgreys. Analog-to-feature converter optimization through power-aware feature selection. International Conference on Analog VLSI Circuits, Oct 2021, Bordeaux, France. hal-03329932v2

HAL Id: hal-03329932

<https://telecom-paris.hal.science/hal-03329932v2>

Submitted on 13 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analog-to-feature converter optimization through power-aware feature selection

Antoine Back, Paul Chollet, Olivier Fercoq and Patricia Desgreys

Laboratoire de Traitement et Communication de l'Information (LTCI)

Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France

antoine.back@telecom-paris.fr, pachollet@telecom-paris.fr, olivier.fercoq@telecom-paris.fr, patricia.desgreys@telecom-paris.fr

Abstract—Analog-to-feature (A2F) conversion is an acquisition method thought for IoT devices in order to increase wireless sensor's battery life. The operating principle of A2F is to perform classification tasks at sub-Nyquist rate, by extracting relevant features in the analog domain and then performing the classification step in the digital domain. We propose to use non-uniform wavelet sampling (NUWS) combined with feature selection to find and extract from the signal, a small set of relevant features for electrocardiogram (ECG) anomalies detection. A CMOS 0.18 μm mixed architecture for NUWS feature extraction is proposed, to obtain a power consumption model for A2F. This model can be taken into account in the feature selection step by evaluating the energy cost of each wavelet and then try to maximize classification accuracy while minimizing the energy needed for the extraction. We demonstrate the benefits of A2F showing that the energy needed can be divided by 15 compared to classical approach.

Index Terms—Analog-to-Feature converter, Bio-sensing acquisition, Feature selection, Low power, Non-Uniform Wavelet Sampling.

I. INTRODUCTION

Increasing the battery life of wireless sensors is a main challenge. As energy consumption of wireless communication systems represents the most significant part of the total consumption, new approaches, like compressed sensing (CS) [1], [2], have been proposed to decrease the amount of transmitted data, by sampling at sub-Nyquist rate. CS was proposed for full signal recovery, but it requires a complex reconstruction step [3], [4] and is limited in terms of achievable compression ratio [5]. Moreover, for certain applications, signal reconstruction is not always required. Analog-to-Feature (A2F) conversion is an acquisition method thought for IoT devices. The aim of this method is to decrease the amount of acquired samples by only extracting useful information directly on the analog signal. Transmitting only little information instead of the entire signal allows to considerably decrease the energy consumption due to wireless communication. Extracted information can be used in a classification task, i.e., to detect voice activity [6], [7] or electrocardiogram (ECG) anomalies [8].

A general architecture for A2F is proposed in [9] and depicted in Fig. 1. This architecture is composed of several configurable feature extractors which work in parallel. A context detector can activate, disable or reconfigure feature extractors depending on the application. It allows adapting accuracy by extracting more features, to decrease the number

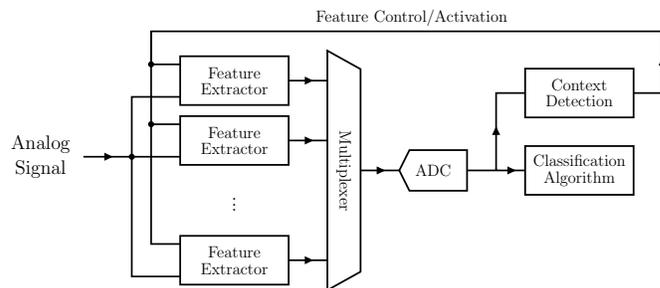


Figure 1. Proposed A2F architecture in [9]

of feature to adjust the power consumption to the minimum or to extract new features for another application. Features are finally used in a machine learning based classifier: this final step can be processed locally on the sensor to transmit only the classification result or later in the chain (data hub, base station, cloud, ...).

One of the challenges of A2F conversion is to identify what is useful information for a given application and how to extract it. The proposed A2F converter is based on the non-uniform wavelet sampling (NUWS) [10]. NUWS provides several degrees-of-freedom that enable flexibility. Therefore, the number of features possibly extracted is very important, solutions have to be found to select a small set of features that provides good classification performance while having a low power consumption. In this article, a power consumption model of an A2F converter for ECG anomaly detection is proposed. This model is based on state-of-the-art front-end power consumption data and power consumption of a proposed digital architecture for wavelet generation. We finally propose to use some well known features selection algorithms to maximize classification accuracy while minimizing energy consumption.

The outline for this paper is as follows. Section II presents the general architectures of the A2F and feature extractors, and introduce the problem brought by NUWS and its solution. In Section III, a power consumption model of our A2F converter is presented. In Section IV, a solution to optimize the power consumption by taking into account wavelet energy cost during selection step is proposed, performance obtained for this optimization step are then are presented. Finally, Section V

concludes the paper.

II. NON-UNIFORM WAVELET SAMPLING FOR FEATURE EXTRACTION

In order to implement an A2F converter, feature extraction is performed using non-uniform wavelet sampling (NUWS). The operating principle of NUWS is to perform a wavelet transform of the signal and then acquire a small set of wavelet samples. In other words, the incoming signal is mixed with several wavelets and then the result is integrated over the support of the wavelets. Each wavelet allows to obtain a feature. The general A2F architecture from Fig. 1 uses several features extractors which works in parallel. The architecture of these extractors is depicted in Fig. 2. It is composed of an analog mixer, an analog integrator, a wavelet generator and a control unit which defines, for a given mother wavelet, the integration time, wavelet's frequency, the sampling frequency and wavelet's time shift. After the features are extracted and acquired by the ADC, the classification can be performed.

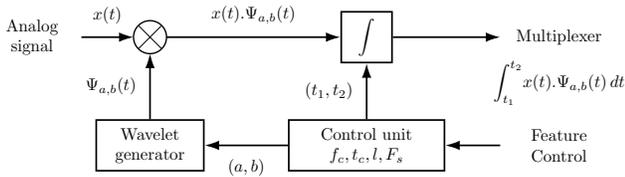


Figure 2. Proposed features extractor

The aim is to design an A2F converter for ECG arrhythmia detection. To evaluate the accuracy of our A2F, ECG signals from the MIT-BIH Arrhythmia database [11] are used. This database is composed of 48 ECG recording from 47 different patients, each recording last 30 minutes and are sampled at a frequency of 360 Hz. From these 48 recordings, 34 were used as training dataset and 14 were used as test dataset. The signal is divided in blocks of $N = 256$ samples centering every QRS complex on the same position. The resulting dataset contains 107,044 annotated beats. A binary classification is performed, using feedforward neural networks with one hidden layer of ten neurons, using *Softmax* activation function, to detect the presence of anomalies.

In [12], two types of wavelets are tested (Haar wavelets and Gabor wavelets) and concludes that Haar wavelets are good candidates for this application. Considering that these two types of wavelet have the same level of performance, Haar wavelets have benefits over Gabor wavelets because it is the simplest wavelet family. It is a family of square function which take values 0, -1 and 1 . Wavelet generation and mixing can thus be implemented in a very efficient way.

As described in [12], a Haar wavelets dictionary is built, including wavelets with different frequencies, time shift and length. Finally, a 502 wavelets dictionary is obtained. Given the number of wavelets, and so the number of features, used to describe the database, feature selection algorithms are required to select a small set of features which provides

good classification accuracy. Some selection algorithms are described in [12], where the benefits of Sequential Forward Search (SFS) are shown.

The exhaustive search, which consists in testing every possible solution quickly becomes impractical when the total number of features increases. The SFS [13], presented in Algorithm 1, is a well-known wrapper model, that consists of starting with an empty set S and then selecting the locally best feature according to the classification performance of the tested subset $J(S)$. In this study, the stopping condition of the algorithm is a maximal number of selected features (d), but it can also be stopped when none of the remaining features allow to improve the performance.

Algorithm 1 Sequential Forward Search

```

( $S \leftarrow \emptyset$ )
repeat
  for all  $X_i \notin S$  do
    ( $J_i \leftarrow J(S \cup \{X_i\})$ )
  end for
  ( $i' \leftarrow \arg \max (J_i)$ )
  ( $S' \leftarrow S \cup \{X_{i'}\}$ )
  if  $J(S') > J(S)$  then
    ( $S \leftarrow S'$ )
    ( $J(S) \leftarrow J(S')$ )
  end if
until  $|S| == d$ 

```

In our previous study, it has been shown that, using feature selection algorithms, a classification accuracy of 98% can be reached by extracting 6 features. The problem is that only the classification accuracy is considered. The easiest way to decrease power consumption of the sensor is to decrease as much as possible the number of features. The second way is to reduce the number of feature extractor by extracting several features with the same extractor. Finally, to optimize the global power consumption, energy needed to extract the features can be taken into account during the selection process, to select features with high accuracy and low power consumption. Features can require more or less energy to be extracted, according to the time support of the related wavelet. Therefore, an estimation of power consumption of the different parts of the system is required.

III. POWER CONSUMPTION MODEL

Wavelet's extraction cost can be estimated by evaluating the power consumption of each component of the extractor. The bio-sensing front end circuit proposed in [14], and depicted in Fig. 3, contains a low noise amplifier (LNA), a programmable gain amplifier (PGA), a G_m -C low-pass filter (LPF) and a 10-bit SAR analog-to-digital converter. The amplifier stage power consumption is equal to $P_{LNA} + P_{PGA} = 5.04 \mu\text{W}$ (see Table I) and the ADC needs $E_{sampling} = 14.3 \text{ fJ/conversion step}$. The proposed filter has a programmable cut-off frequency between 150 and 10 kHz. NUWS requires a LPF with integrator behavior: the minimum integrated frequency is fixed by the time window (the minimum wavelet frequency). Given the minimum wavelet

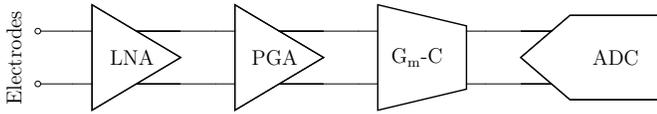


Figure 3. Bio-sensing front-end from [14]

frequency of 1.4 Hz, the filter does not achieve the needed characteristics, in terms of cut-off frequency. The problem in the context of the realization of this integrator is to design a first order low-pass filter with a sub-hertz cut-off frequency. Indeed, to reduce the cut-off frequency, it is necessary to use a larger capacitor, but its size is limited in an integrated circuit, or to reduce the value of the transconductance. The first order low-pass filter proposed in [15], and depicted in Fig. 4, has a programmable cut-off frequency between 220 mHz and 39.1 kHz, and has a power consumption of 1.08 μ W. It has two tuning voltages, V_{t2} and V_{gc} , allowing to modify transconductance value with a thick/thin relationship. Fig. 5 shows the frequency range, achieved by the LPF proposed in [15], according to the different tuning voltages and capacitor configurations. As Haar wavelets take values -1 and 1 , during their validity period, we propose to implement the mixing step as a switching system inverting positive and negative inputs of the integrator. As shown in Fig. 4, the LPF works in a single-ended mode, a buffer must be used between mixing step outputs and integrator input to convert signal from differential signal to single-ended signal. Table I summarizes the performance of analog component used for feature extraction.

As shown in [12], Haar wavelets are good candidates for NUWS feature extraction. These wavelets are square signals, so it means that it can be easily generated by a digital system. Fig. 6 is a simplified schematic of our wavelet generator's operation. Wavelets are described by three parameters: the frequency (divisor), the length and the time shift. The generator is a clock divider that divides the clock signal to obtain a 360 Hz clock signal (which corresponds to the precision of our database). This 360 Hz clock is divided a second time to obtain the wavelet signal. As Haar wavelet are "3 states" signals, our wavelet generator as two 1-bit outputs. The output *Enable*, represent the validity period of the wavelet, in other words, when the wavelet is null or not, and the output *Wavelet* is the waveform.

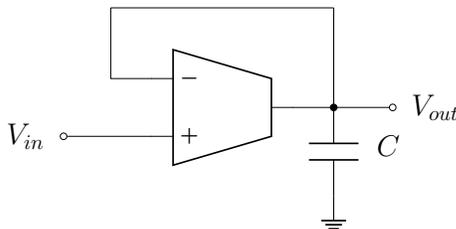


Figure 4. G_m -C integrator

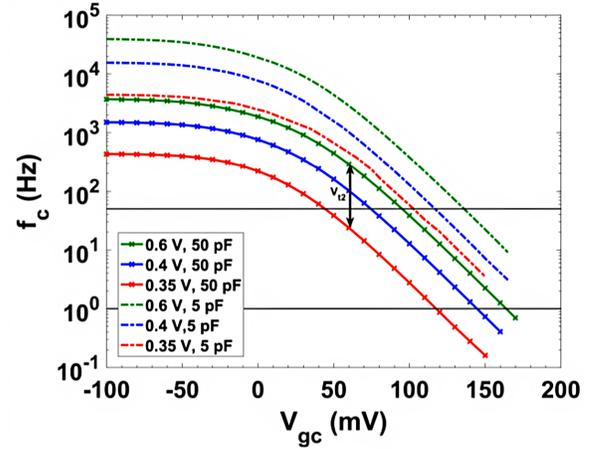


Figure 5. Frequency range achieved by LPF proposed in [15] for different voltages [V_{t2} ; V_{gc}] and capacitor configurations

This configurable wavelet generator is described in the hardware description language Verilog and its power consumption is estimated for the same technology as previously (XFAB, 0.18 μ m) with an operating frequency of 2 MHz. After place and route simulation, the estimation of the power consumption of a wavelet generator is about 22.4 μ W during generation phase. The generator occupies 0.035 mm² of silicon area. The result of place and route step is depicted in Fig. 7.

In order to decrease power consumption of this part of the A2F, a very well known technique, used in synchronous circuits, called Clock Gating [16], is used. This technique consists in disabling the clock signal in unused parts of the circuit. Using clock gating, the power consumed by the clock tree can be decreased: disabling the clock in unused parts of the circuit make possible to reduce the load on the clock tree and so to reduce the number of buffer used in the tree. Flip-flops power consumption is also decreased because, without clock signal, these are not triggered and so the dynamic power is saved. The compensation of it, is that more logic must be added to the circuit. Thanks to clock gating, power consumption of the wavelet generator is reduced from 22.4 μ W to $P_{Generator} = 9 \mu$ W.

Table I
PERFORMANCE SUMMARY

LNA and PGA [14]		LPF [15]		ADC [14]	
Supply voltage (V)	1.4	Supply voltage (V)	1.8	Supply voltage (V)	1
Current (A)	3.6 μ	Current (A)	0.6 μ	Current (A)	0.3 μ
Gain (dB)	38 – 72	C (pF)	50	SNDR(dB)	56.1
Cut-off frequencies (Hz)	0.5 – 300	Cut-off frequency (Hz)	0.22 - 39.1k	ENOB (bits)	9.02

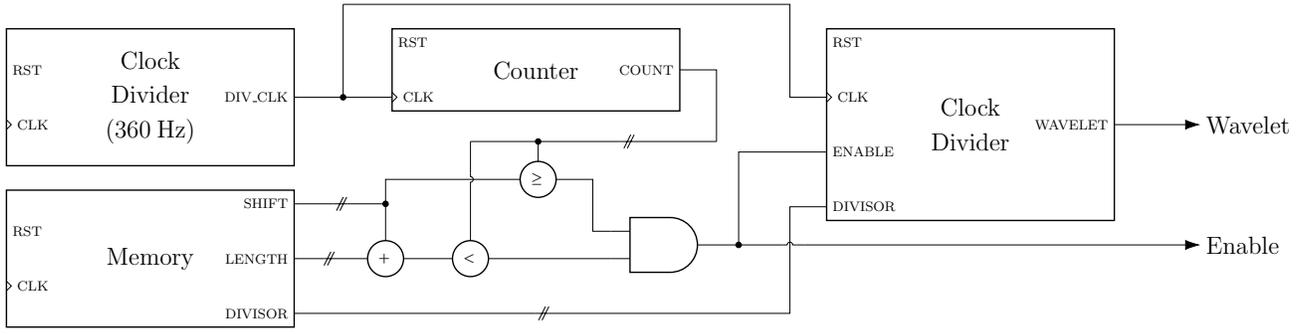


Figure 6. Schematic presenting wavelet generator's operation

IV. POWER CONSUMPTION OPTIMIZATION

The SFS algorithm, presented in Section II, is modified to take into account two new ideas: the maximum number of feature extractor used in parallel and the power consumption of each wavelet. The main way to reduce the global power consumption is to minimize the number of branches and acquire several features from the same extractor. To ensure that, two features must use wavelets which are not overlapping in time domain. After evaluating each combination, features are ranked according their performance. Knowing the number of extractors and the previously selected features, SFS selects the first feature of the rank, which can be extracted. If there is no solution, SFS ends. Fig. 8 shows the classification accuracy obtained from this new version of SFS by limiting to 3 and 5 the number of branches in our A2F, where the different marks indicate the branch chosen to extract the selected feature.

To minimize the power consumption, while maximizing classification accuracy, a new evaluation function in SFS is proposed:

$$J_i = \frac{A(S \cup \{X_i\}) - A(S)}{E_{feature}} \quad (1)$$

where S is the set of selected features, X_i is the tested feature, $A(S)$ is the classification accuracy of the set S and $E_{feature}$ is the energy needed for the feature extraction. The initial accuracy $A(0)$ is set to 85% which corresponds to the accuracy of a zero rule classifier, used as reference. A ZR classifier assigns to each data item the most frequently occurring class

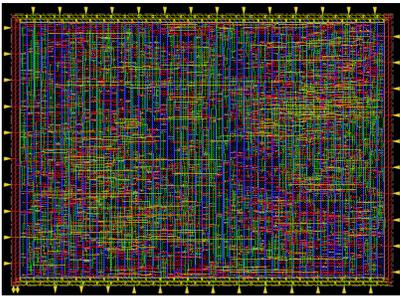


Figure 7. Wavelet generator XFAB 180 nm place and route schematic

in a data set. In our test dataset, about 85% of data was normal, so ZR is right 85% of the time.

The estimation of the extraction cost of each wavelet ($E_{feature}$) is calculated according to the following formula:

$$E_{feature} = E_{sampling} + E_{analog} + E_{digital} \quad (2)$$

$$E_{sampling} = \frac{P}{2^{ENOB} \cdot f_s} \quad (3)$$

$$E_{analog} = \Delta t \cdot (P_{LNA} + P_{PGA} + P_{GMC}) \quad (4)$$

$$E_{digital} = \Delta t \cdot P_{Generator} \quad (5)$$

where Δt is the wavelet time support length, $E_{sampling}$ is the energy needed per analog-to-digital conversion, this value is constant regardless of the wavelet, E_{analog} is the energy needed for the analog component (LNA and PGA from [14] and integrator presented in Section III), it is proportional to the wavelet length, and $E_{digital}$ is the energy needed by the digital system to generate the wavelet. Fig. 9 shows the classification accuracy and the total energy cost of selected features according to this new criteria. We can see that with the previous criteria, an accuracy of 98.4% can be reached while requiring 10.9 μ J (6 features) and with this new criteria, an accuracy of 98% can be reached while requiring 3.3 μ J (for 10 features).

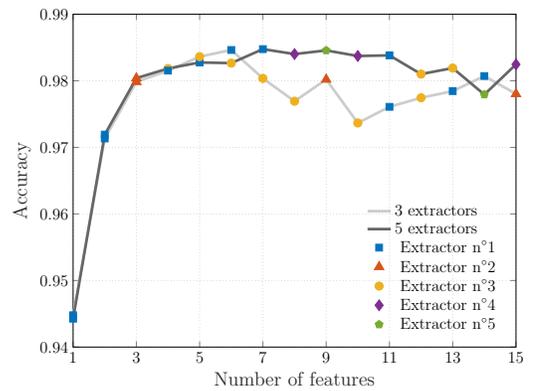


Figure 8. Results of SFS with fixed number of extractors

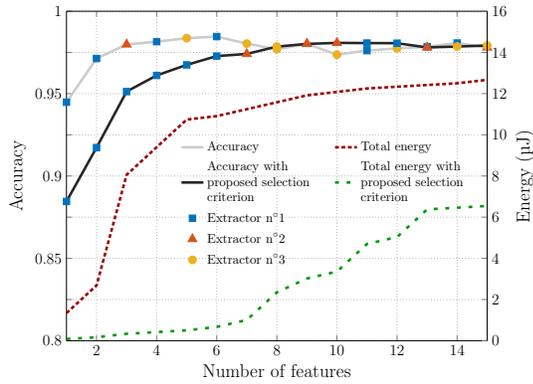


Figure 9. Comparison of classification accuracy and total energy cost with/without power optimization

Now, a comparison of the three different approaches of wireless sensors can be made to show the benefits of A2F over classical approaches using Nyquist acquisition and analog to information (A2I) sensors based on compressed sampling. The bio-sensing front-end from [14] is considered as the classical approach. It has a power consumption of $6.04 \mu\text{W}$ and work at a sampling frequency $F_s = 2 \text{ kHz}$. The second approach is a random modulator pre-integrator A2I converter described in [5]. This A2I works at a frequency of 2 kHz and is configured with 32 branches and an analysis time window of $N = 128$ samples. Each branch of referred A2I has a power consumption of 28 nW . In this study, the energy for recovery step of A2I is not considered. We consider here, the wireless transmission system from [17], based on Bluetooth low energy (BLE) protocol which requires 37 nJ to transmit a 10-bit sample. The Fig. 10 presents the energy needed by these three types of sensor to make the acquisition of 10 seconds of signal. The vertical axis, for energy, has a logarithmic scale. The energy needed by the classical approach is about $800.4 \mu\text{J}$, while the energy consumed by the A2I sensor is about $194 \mu\text{J}$. The proposed solution only consumes $52.3 \mu\text{J}$, so the energy needed is divided by about 15.3 compared to classical ap-

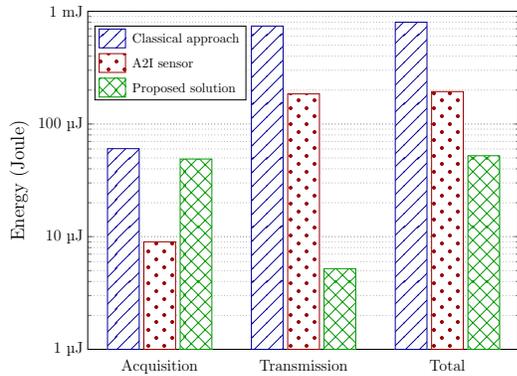


Figure 10. Comparison of the energy needed to acquire 10 seconds of signal by the 3 types of sensors

proach and by approximately 3.7 compared to A2I sensor. It shows the benefits of A2F over the two other approaches: by adding three analog feature extractors, the number of samples is wildly reduced compared to other approaches and therefore global power consumption of wireless sensors decreases.

V. CONCLUSION

In this paper, a $0.18 \mu\text{m}$ architecture is proposed to implement and make an estimation of the power consumption of an Analog-to-feature converter, based on non-uniform wavelet sampling, for ECG anomalies detection. The amount of possible wavelets provided by NUWS is very high, so well known feature selection algorithms were used to decrease the size of the wavelet dictionary while maximizing the classification accuracy. In order to optimize the power consumption of this A2F, we proposed to take into account two new criteria in the selection process, using the Sequential Forward Search, besides accuracy: the number of extractors used in parallel and the energy needed to extract each feature. The proposed SFS algorithm using this criterion allows to both decrease the number of extracted features, and therefore the amount of data the sensor had to transmit, and the number of extractors by sharing them to extract several features during a unique analysis time window. This selection step allows to know the circuit's complexity by determining the minimal number of extractors needed to reach the desired accuracy, and therefore to make a physical implementation of the entire analog-to-feature converter. Finally, benefits of A2F over the two others classical approaches and A2I sensors, were shown. Another future work could also focus on the generalization of the A2F to different low frequency signals, such as electroencephalogram (EEG) or voice activity, to take advantage of the A2F's reconfigurability.

REFERENCES

- [1] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [2] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, May 2008.
- [3] E. Crespo Marques, N. Maciel, L. Naviner, H. Cai, and J. Yang, "A Review of Sparse Recovery Algorithms," *IEEE Access*, vol. 7, pp. 1300–1322, 2019.
- [4] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: From theory to applications, a survey," *Journal of Communications and Networks*, vol. 15, no. 5, pp. 443–456, Oct. 2013.
- [5] D. Gangopadhyay, E. G. Allstot, A. M. R. Dixon, K. Natarajan, S. Gupta, and D. J. Allstot, "Compressed Sensing Analog Front-End for Bio-Sensor Applications," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 2, pp. 426–438, Feb. 2014.
- [6] K. M. H. 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 Journal of Solid-State Circuits*, vol. 51, no. 1, pp. 291–302, Jan. 2016.
- [7] X. Liu, E. Gönültaş, and C. Studer, "Analog-to-Feature (A2F) Conversion for Audio-Event Classification," in *2018 26th European Signal Processing Conference (EUSIPCO)*, Sep. 2018, pp. 2275–2279.
- [8] X. Tang, Q. Hu, and W. Tang, "Analog to Digital Feature Converter based on Oversampling Modulators for ECG Delineation," in *2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug. 2019, pp. 121–124.

- [9] M. Verhelst and A. Bahai, "Where Analog Meets Digital: Analog to Information Conversion and Beyond," *IEEE Solid-State Circuits Magazine*, vol. 7, no. 3, pp. 67–80, 2015.
- [10] M. Pelissier and C. Studer, "Non-Uniform Wavelet Sampling for RF Analog-to-Information Conversion," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 2, pp. 471–484, Feb. 2018.
- [11] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, May 2001.
- [12] A. Back, P. Chollet, O. Fercoq, and P. Desgreys, "Feature selection algorithms for flexible analog-to-feature converter," in *2020 18th IEEE International New Circuits and Systems Conference (NEWCAS)*, Jun. 2020, pp. 186–189.
- [13] F. J. Ferri, P. Pudil, M. Hatef, and J. Kittler, "Comparative study of techniques for large-scale feature selection," in *Machine Intelligence and Pattern Recognition*, ser. Pattern Recognition in Practice IV, E. S. Gelsema and L. S. Kanal, Eds. North-Holland, Jan. 1994, vol. 16, pp. 403–413.
- [14] H. Bhamra, J. Lynch, M. Ward, and P. Irazoqui, "A Noise-Power-Area Optimized Biosensing Front End for Wireless Body Sensor Nodes and Medical Implantable Devices," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 10, pp. 2917–2928, Oct. 2017.
- [15] J. Pérez-Bailón, B. Calvo, N. Medrano, and J. Ramírez-Angulo, "A 1.8 V Gm-C Highly Tunable Low Pass Filter for Sensing Applications," pp. 1–5, May 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9128692/>
- [16] Q. Wu, M. Pedram, and X. Wu, "Clock-gating and its application to low power design of sequential circuits," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, no. 3, pp. 415–420, Mar. 2000.
- [17] F.-W. Kuo, S. Binsfeld Ferreira, H.-N. R. Chen, L.-C. Cho, C.-P. Jou, F.-L. Hsueh, I. Madadi, M. Tohidian, M. Shahmohammadi, M. Babaie, and R. B. Staszewski, "A Bluetooth Low-Energy Transceiver With 3.7-mW All-Digital Transmitter, 2.75-mW High-IF Discrete-Time Receiver, and TX/RX Switchable On-Chip Matching Network," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 4, pp. 1144–1162, Apr. 2017.