Zeroing for HW-efficient compressed sensing architectures targeting data compression in wireless sensor networks

Zeroing for HW-efficient Compressed Sensing Architectures Targeting Data Compression in Wireless Sensor Networks

Mauro Mangia[a,b], Daniele Bortolotti[a,*], Fabio Pareschi[c,b], Andrea Bartolini[a,d], Luca Benini[a,d], Riccardo Rovatti[a,b], Gianluca Setti[c,b]

[a] *DEI, University of Bologna, Italy*
[b] *ARCES, University of Bologna, Italy*
[c] *ENDIF, University of Ferrara, Italy*
[d] *ISL, ETH Zürich, Switzerland*

Abstract

The design of ultra-low cost wireless body sensor networks for wearable biomedical monitors has been made possible by today technology scaling. In these systems, a typically multi-channel biosignal sensor takes care of the operations of acquisition, data compression and final output transmission or storage. Furthermore, since these sensors are usually battery powered, the achievement of minimal energy operation is a fundamental issue. To this aim, several aspects must be considered, ranging from signal processing to architectural optimization. In this paper we consider the recently proposed rakeness-based compressed sensing (CS) paradigm along with its zeroing companion. With respect to a standard CS base sensor, the first approach allows us to further increase compression rate without sensible signal quality degradation by exploiting localization of input signal energy. The latter paradigm is here formalized and applied to further reduce the energy consumption of the sensing node. The application of both rakeness and zeroing allows for trading off energy from the compression stage to the transmission or storage one. Different cases are taken into account, by considering a realistic model of an ultra-low-power multicore DSP system.

Keywords: Compressed Sensing, Rakeness, Zeroing, Low-Power, Wireless Sensors Networks, Non Volatile Memories, Wearable Biomedical Monitors

1. Introduction and Related Work

Many diseases related to modern human behavior, such as cardiovascular ones, require precise and long-term medical supervision. This situation is, however, unsustainable for traditional healthcare systems [1], due to the increasing costs and medical management needs. A large-scale and cost-effective solution to the aforementioned issue may be given by the introduction and the large diffusion of ultra-low power (ULP) personal health monitoring systems. As a matter of fact, emerging and future healthcare policies are actually fueling a shift toward long term monitoring of biosignals by means of wearable embedded devices [2]. These healthcare applications need the emergence of new technologies to allow the development of a Wireless Body Sensor Networks (WBSNs), i.e, a network of extremely power-efficient, wearable, bio-sensing nodes with adequate performance, capable of enabling the continuous measurement of biomedical signals, and providing the ubiquitous, long-term and real-time monitoring required for fast coordination with medical personnel.

The availability of WBSNs paves the way to new medical strategies not limited to patient monitoring, but useful in many other application areas such as athletes' training improvement, elders continuous caring, or stress detection during safety critical tasks. Yet, such WBSNs are subject to conflicting strict requirements, in particular for the sensor nodes design. On the one hand they are constrained to an extremely low power budget in the effort to extend as much as possible battery life or to allow energy harvesting techniques [3]. On the other hand, there is an increasing demand for computation capabilities to locally process the acquired data and reduce to a minimum the data size while preserving the information.

Indeed, the final goal of minimal energy operation can be achieved only with an holistic approach, by considering and combining the several trade-offs and optimization aspects of the signal processing chain and of the different technological layers composing the ULP architecture.

The recently introduced Compressed Sensing (CS) signal acquisition and compression paradigm [4, 5] has proved to be effective in reducing energy consumption in biomedical monitors [6, 7], representing a good solution for the design of ULP sensor nodes. CS is a dimensionality reduction technique based on a linear transformation that is typically a projection on a set of random vectors (sensing vectors),

*Corresponding author
Email address: daniele.bortolotti@unibo.it (Daniele Bortolotti)

arranged as the rows of the so called sensing matrix. The vector of input signal samples is mapped into a *measurement* vector with a smaller size, from the knowledge of which (under assumptions that are usually satisfied) it is possible to reconstruct the original signal without significant loss of quality. This dimensionality reduction allows a potentially large saving in the resources needed at the sensing node (mainly power) since *i*- the amount of additional processing (a linear transformation) is intrinsically small and can be effectively reduced; *ii*- the transmitter (the most power hungry stage) significantly benefits from a reduced load [8]. For instance, the work presented in [9] shows a $\approx 40\%$ improvement in lifetime compared to state-of-the-art DWT-based compression techniques for an embedded ECG monitor.

Further system optimization is possible thanks to recently proposed improvements of the standard CS theory. An example is the introduction of the concept of *rakeness* [7, 10]. The basic idea behind this approach is to exploit statistical features of the signal (namely *localization*, i.e. the assumption that the information of the signal is not equally distributed in its domain) to improve performance. Roughly speaking, localization implies that some realizations of the input process have a higher probability with respect to other ones. Note that this is not a limitation, since the only class of signals where all possible realizations have the same probability is white noise.

In the effort of reducing the computational power required by a sensing node exploiting CS, we investigate in this paper a new rakeness-based approach called *zeroing*, recently introduced in [8] and developed in [11]. The main idea underlying this approach is to use a sparse sensing matrix or rather a matrix with fewer non-null elements compared to full sensing matrices, obtained by randomly setting to zero elements of rakeness-based sensing matrix. Such approach outperforms the rakeness one in terms of memory footprint and computation requirements: In addition the sparsity of its sensing matrix makes it suitable for a more efficient algorithmic implementation. The drawbacks is a distortion of the statistical characterization of the sensing matrix with a decrease in performance with respect to rakeness approach. Nevertheless, it opens up a new range of possible trade-offs.

In our reference WBSN architecture, previously introduced in [8] and [11], CS is applied into the digital domain and implemented within the sensing nodes. The output data is either transmitted to the WBSN-gateway or locally stored for later usage. The functionality can be logically divided into three separate phases : *i*- sensors data acquisition (*input collection*), *ii*- input data processing (*compression*) and finally *iii*- handling of the processed input (*transmission or storage*). In the first stage the input analog signal is sampled, the data is periodically made available to a digital signal processor (DSP) which is in charge of compressing the data by means of the CS algorithm. Next, the compressed output can either be transmitted (to the gateway of the WBSN, e.g., a smartphone) or stored for later off-line medical analysis.

If we consider the case of medical grade biomedical monitoring, where typically a multi-lead signal is deployed, the computation task exhibits high potential for parallel processing. Indeed, parallel processing enables a more aggressive voltage-frequency scaling than single-core solutions, though at ultra-low workload requirements the single-core solution pays off because of the lower leakage. A multi-core DSP architectural template showed its higher energy efficiency, with respect to single-core solutions [12, 13], within a workload range that comprises the requirements of biosignal processing applications. The authors in [8] adopted a multi-core architecture where each lead is processed on a separate core in a parallel fashion and each core has an area-efficient segregated instruction memory. Moreover, in [14] an hybrid memory extension is devised to reduce the consumption adapting the active memory portions to different CS workload requirements, while in [15] a robust reconstruction algorithm allows to exploit voltage scaling near to the threshold level. In this work the architecture proposed in [12] is further consolidated by optimizing different algorithmic aspects with the goal of increasing the energy efficiency at system-level.

Compared to our previous works [8, 11] this work discusses the trade-offs for both the compression and reconstruction part. Showing that the benefits of the zeroing approaches are maximized when also the reconstruction engine runs in a battery powered device.

The main contributions of this work are the following:

- we formalize the zeroing CS approach, which further extends the compressive sensing efficiency for bio-sensing applications.
- we show that the rakeness approach can lead to a higher compression ratio, for a fixed quality of service, if compared to the standard CS. Moreover, we show that the zeroing approach can reduce the cost of compression by trading off quality of service for computation in battery-limited bio-sensing nodes.
- according to the constraints on the personalized gateway different reconstruction algorithms can guarantee the desired target performance at different compression ratio values. Our results show that signal reconstruction performed by a greedy algorithm hides the degradation introduced by the zeroing approach, making it the best choice for the battery lifetime of both the bio-sensing node and the personalized gateway performing the decoding task.
- we consider different transmission technologies in our analysis, showing that the zeroing still performs the best in terms of battery life for a wide set of applications. Our results show that for all the technologies considered the proposed rakeness and zeroing approaches outperform the standard CS one.

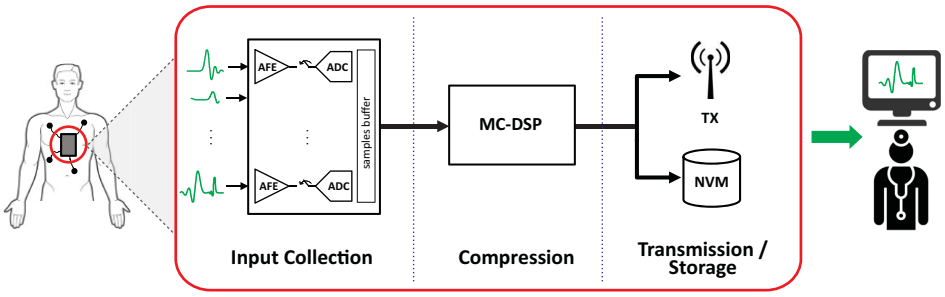

Figure 1: Block diagram of the considered biomedical system.

When compared to the no compression case interesting results can be found, as the rakeness technique makes the compressing sensing approach beneficial for transmission and storage technology with medium-high energy cost.

The paper is organized as follows. Section 2 presents the WBSN system architecture and the multi-core DSP for signal processing. In Section 3 we propose an overview of both CS and rakeness-based CS and we introduce the zeroing approaches. Next, Section 4 discusses a case study of the different CS algorithms for biomedical monitors, highlighting several system-level trade-offs. Section 5 presents the experimental setup and the results of the evaluation of the proposed algorithms in terms of reconstruction quality, memory footprint and energy efficiency considering different technological solutions for output transmission and storage. Finally, the conclusion of this work is presented in Section 6.

2. System Architecture

2.1. Overview

In the present work we are considering a biomedical system where three phases can be identified: *Input Collection*, *Compression* and *Transmission/Storage*. A block diagram of the biomedical system is shown in Figure 1.

Input Collection. At first, the input multi-channel biosignal is sampled by the analog front-end (AFE) with a sampling frequency (f_s), according to the properties of the signal to analyze and the accuracy required. During the data collection phase the DSP waits for the number of samples (n) required to perform compression. Considering typical sampling frequencies for biomedical signals, this phase exceeds in time the phase of computation. For example, assuming $f_s = 256$Hz and $n = 512$, the input collection lasts for 2 seconds. During this phase, for most of the time the whole system is idle therefore we assume an ultra-low power state (almost zero power) for both the DSP and the transmission/storage back-end to avoid unnecessary power consumption. Once a set of n new samples is ready in the sampling buffer, the AFE triggers the DMA (via an interrupt mechanism) to move to the samples to the DSP data memory.

Compression. When all the samples from the different leads are collected, the compression algorithm is performed by the multi-core DSP (cf. Section 5.1). The considered system performs a burst of computation on the available data for future transmission or storage. In this phase the digital processor is subject to high workload requirements and a high memory footprint as described in [8]. As described in Section 4, the introduced zeroing paradigm allows for energy tradeoffs between the computational load and the size of the measurement vectors for subsequent transmission or storage.

Transmission/Storage. In the last stage, the compressed data are either transmitted (TX) in real-time or stored in a non-volatile memory (NVM) for later off-line usage, for instance for monitoring a long activity window of a patient to be analyzed by medical personnel. An overview of the technologies considered in this work is presented in Table 1, while the energy figures following for state-of-the-art TX protocol or NVM technology will be presented in Section 5.

Table 1: Summary of TX and NVM technologies considered.

	Technology	
TX	*Bluetooth Low Energy*	(BLE)
	Narrow Band	(NB)
	Human Body Channel	(HBC)
	Near-field Communication	(NFC)
NVM	*Resistive RAM*	(ReRAM)
	Spin-transfer torque magnetic RAM	(STT-MRAM)
	FG Flash Memory	(FLASH)
	Conductive Bridge RAM	(CBRAM)

During this phase, we further assume a power-efficient architecture in a deep low-power state with negligible consumption for the DSP, which is achievable thanks to the knobs offered by modern technological processes such as fully depleted silicon on insulator (FD-SOI) [16].

2.2. Multi-Core DSP

As introduced above, medical grade biomedical monitoring has an inherent parallel nature where multi-channel signal analysis is often embarrassingly parallel and can be perfectly mapped on a multi-core architecture. In this work we consider a DSP architecture representative of a class of modern low-power multi-core architectures, targeting digital biosignal processing [12, 13, 17]. Figure 2 shows our MC-DSP architecture, featuring 8 processing elements (PEs) with an Harvard architecture for the memories, i.e., separate instruction and data buses.

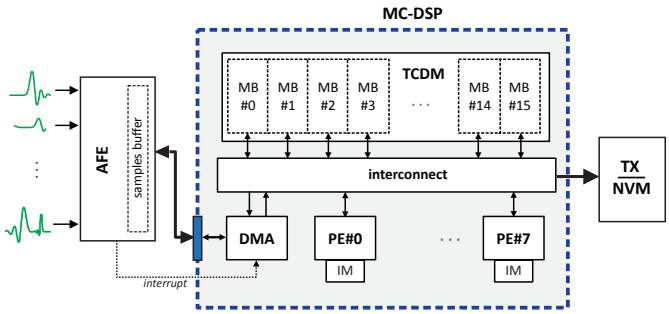

Figure 2: Multi-core DSP architecture for CS.

The instruction memory (IM) architecture is therefore simple and avoids long stalls for cache refill costs and coherency protocol overheads. Indeed, each PE has a private single-cycle IM, where the code is statically stored. On the data side, a multi-banked tightly coupled data memory (TCDM) is shared among all the cores acting as a scratch-pad memory. The number of memory ports of the TCDM is equal to the number of banks enabling concurrent access to different memory locations. Once a read or write requests is brought to the memory interface, the data is available on the negative edge of the same clock cycle, leading to a total latency of 2 clock cycles for TCDM access without conflicts[1]. The communication is based on a high-bandwidth low-latency interconnect, based on a mesh-of-trees logarithmic network able to support single-cycle communication between PEs and memory banks (MBs), as in [18]. In case of multiple conflicting requests, for fair access to memory banks, a round-robin scheduler arbitrates the accesses. To ease the negative impact of banking conflicts we consider a banking factor of 2. With reference to Figure 2, the digital samples corresponding to different channels are temporary stored in a buffer inside the analog front end during the input collection phase. To reduce the capacity requirements for such buffer, a DMA engine is periodically activated, via an interrupt from the AFE, to move the samples from the AFE buffer to the TCDM. Once the last series of samples is moved to the DSP memory, the multi-core compression algorithm starts, where each core is active and working on its own subset of the sampled data. We assume that the computation phase must be completed before the first sample of the next time window $(n+1)$ is available to avoid double buffering overhead.

[1] If conflicts occur there is no extra latency between pending requests, once a given bank is active, it responds with no wait cycles.

3. Basics of Compressed Sensing

The CS is a paradigm recently introduced in the area of signal processing with the purpose of merging signal acquisition and compression tasks. CS aims to overcome the limit of the Nyquist-Shannon sampling theorem by representing the information content of the input signal using fewer digital words with respect to Nyquist-rate. The only assumption is to have a to prior knowledge on the considered class of input signals, which must be *sparse*. This implies that the information content associated to the input signal is less than the amount of Nyquist-rate samples, thus enabling data compression possibilities [19].

From a mathematical point of view, the sparsity assumption can be formalized as follows. Let us refer to a given time window T, being n the number of Nyquist-rate samples (being them either analog samples or already quantized ones) necessary to represent a signal defined in T. Let us name $x = (x_0, \ldots, x_{n-1})^\top \in \mathbb{R}^n$ the vector containing the n Nyquist-rate input signal samples. The input signal is k-sparse if there exists a n-dimensional *sparsity basis* or *dictionary* $\Psi = \{\psi_0, \ldots, \psi_{N-1}\}$, $\Psi \in \mathbb{R}^{n \times N}$ and $N \geq n$, such that any realization $x = \Psi\alpha$ of the input signal is such that its coefficients vector $\alpha \in \mathbb{R}^N$ has at most $k \ll n$ non-zero components in its support.

In the following, a brief description of the standard CS approach is given, along with an overview of the *rakeness* and of the *zeroing* approaches.

3.1. Standard CS Framework

In the aforementioned signal model the actual number of degrees of freedom in x is considerably smaller than n. Leveraging this property, fundamental results [20] have showed that its salient information content can be captured in a set of $m < n$ linear *measurements*. These are gathered in the m-dimensional vector $y = (y_0, \ldots, y_{m-1})^\top \in \mathbb{R}^m$, as obtained by projecting the input signal x on a suitable set of n-dimensional sensing vectors ϕ_j, $j = 0, \ldots, m-1$ arranged row by row in the *sensing matrix* $\Phi \in \mathbb{R}^{m \times n}$. The CS *encoding* is represented as:

$$y = \Phi(x + \xi) = \Phi\Psi\alpha + \Phi\xi = \Phi\Psi\alpha + \nu \qquad (1)$$

where ξ is an additive term that can be used to take into account all non-idealities like the quantization error or the noise on the input signal and $\nu = \Phi\xi$ is a vector in $\mathbb{R}^m$ that counts the effect of noise sources on the measurements. The ability of the sensing stage to efficiently capture the signal information content is represented by the *compression ratio* (CR) of the CS system defined as the ratio n/m.

The decoder stage receives y and, thanks to the knowledge (shared with the encoder) of both Φ and Ψ, reconstructs an estimation $\hat{x}$ of the input signal x. Formal results [20, 21] guarantee that α (and thus x) can be recovered from y despite the fact that Φ (and thus $\Phi\Psi$) is a dimensionality reduction. In particular, correct reconstruction with probability equal to one is guaranteed by the so-called *Restricted Isometry Property* (RIP) of the matrix $\Phi\Psi$, ensuring that the sensing stage is able to conserve the input signal ℓ_2 norm also with $m < n$, i.e. with an ensuring CR¿1. The CS theory guarantees [4] that, for all possible sparsity matrices Ψ (i.e. for either any sparsity basis or dictionary), the RIP of $\Phi\Psi$ is always satisfied by adopting Φ composed of a collection of *independent and identically distributed* (i.i.d.) random variables with either Gaussian or sub-Gaussian distributions. Over this constraint the the minimum number of needed projections for a correct reconstruction is bounded as follow [4].

$$m \geq m_{\min} = Ck\log(n/k)$$

where C is a constant typically set around 4.

Note that the linearity of (1) requires only simple operations, that can be efficiently implemented in a low power sensor node. Yet, to further reduce the required power, it is possible to generate all elements of Φ as *binary antipodal* random variables, where the two possible values $+1$ and -1 occur with the same probability, i.e. $\Phi \in \{+1, -1\}^{m\times n}$. This still ensures RIP while at the same time reduces the sensing stage complexity by relaxing the product operations required to compute y into much simpler sign inversion operations, allowing a straightforward hardware implementation, since only $m \cdot n$ sums or subtractions are required. With this approach, no reduction in terms of CS performance with respect to the Gaussian case has been observed [19]. A graphical representation of the encoding process by using random antipodal sensing matrix is depicted in Figure 3.

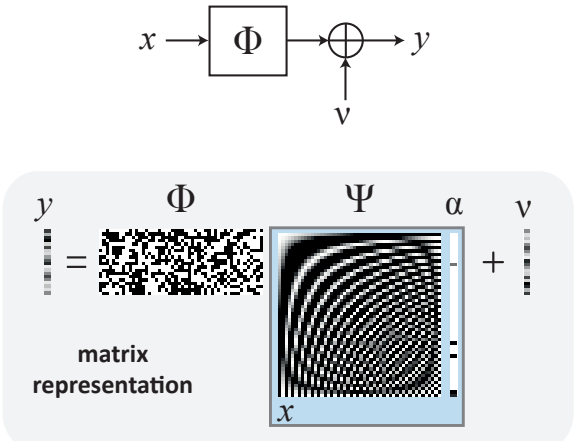

Figure 3: Matrix representation of the information extraction based on Compressed Sensing.

Assuming RIP is satisfied, signal reconstruction is achieved by solving dedicated optimization problems that look for the sparsest α subject to constraints forcing the corresponding measurements to be as close as possible to the observed y [20, 22]. Sparsity is generally promoted by the ℓ_1 norm instead of the computationally intractable count of non-zero components. Along this path at least three methods are commonly employed.

The first method is called Basis Pursuit (BP) and simply neglects disturbances to solve

$$\min_{\alpha} \|\alpha\|_1 \;\text{ s.t. }\; \Phi\Psi\alpha = y$$

where $\|\cdot\|_p$ indicates the usual ℓ_p norm. The main appeal of BP is that it can be recast into a fully linear optimization problem for which standard methods exists though, *ad hoc* techniques have been developed. The second method is called Basis Pursuit with Denoising (BPDn) and takes into account disturbances solving

$$\min_{\alpha} \|\alpha\|_1 \;\text{ s.t. }\; \|\Phi\Psi\alpha - y\|_2^2 \leq \epsilon^2 \tag{2}$$

where ϵ^2 is tuned on the characteristics of the disturbance term ν. A last method keeps the *denoising* formulation of BPDn but focuses directly on the target vector x instead of α. This can yield definite advantages in the quality of reconstruction when Ψ is a dictionary, i.e. $N > n$. In that case one solves

$$\min_{x} \|\Psi^* x\|_1 \;\text{ s.t. }\; \|\Phi x - y\|_2^2 \leq \epsilon^2 \tag{3}$$

where Ψ^* is a *analysis transform* operator [23] that for every x chooses one of the possible representations with respect to Ψ. This problem is called Analysis BPDn (ABPDn) and is of interest here since the change of point of view from α to x may imply a different computational burden even if Ψ is a basis and thus $\Psi^* = \Psi^{-1}$.

On top of these methods relying on sparsity promotion by means of the ℓ_1 norm, there are other greedy approaches that iteratively promote sparsity by observing intermediate and approximate solutions. Implementations of CS decoding on embedded, low-resources platforms usually look into this set of methods rather than feed a solver with a suitably defined optimization problem. In all cases, quality of reconstruction depends on m, i.e. on the amount of information that is passed from the encoder to the decoder. Remarkable, the m minimization is of paramount importance.

3.2. Rakeness Approach

To further increase the CS compression performance, in [10] authors introduced the *rakeness* extension, coupling the sparsity hypothesis with the additional assumption on the acquired class of signals to be *localized*, i.e. the information content in each instance x is non-uniformly distributed in the whole signal domain. In this scenario, CS compression capability is increased by introducing a new guideline in the sensing, aiming at collecting ("rake") as

much energy as possible when projecting the input signal onto the ϕ_j, by a statistical matching between x and ϕ_j. At the same time such approach preserves the randomness of Φ and so the RIP of the $\Phi\Psi$ operator. This allows either to increase the reconstruction quality or to reduce m (i.e. increase compression ratio) at a given reconstruction performance level.

More formally, let us model ϕ_j and x as realizations of two independent stochastic processes $\underline{\phi}$ and $\underline{x}$, whose $n \times n$ correlation matrices are given by C^{ϕ} and C^{x}. We define the *rakeness* ρ as

$$\rho(\underline{\phi}, \underline{x}) = \mathbf{E}_{\underline{\phi},\underline{x}} \left[|\langle \phi_j, x \rangle|^2 \right] = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} C^{\phi}{}_{i,j} C^{x}{}_{i,j} \qquad (4)$$

where $\mathbf{E}_{\underline{\phi},\underline{x}}$ stands for the expected value computed with respect to both $\underline{\phi}$ and $\underline{x}$, and $\langle \cdot, \cdot \rangle$ stands for the standard inner product such that $\langle \phi_j, x \rangle = \sum_{i=0}^{n-1} \phi_{j\,i} x_i$.

The idea to increase the collected signal energy by the generic ϕ_j, with the constraint that the sensing vectors are random enough to preserve the RIP, is translated into solving the following optimization problem

$$\begin{array}{ll} \max\limits_{\underline{\phi}} & \rho(\underline{\phi}, \underline{x}) \\ \text{s.t.} & \langle \phi_j, \phi_j \rangle = e \\ & \rho(\underline{\phi}, \underline{\phi}) \leq \tau e^2 \end{array} \qquad (5)$$

where the first constraint requires that the energy of each sampling vector is e, and the second one is an upper bound related to the randomness of the process generating all ϕ_j involved in the sensing. The optimization problem (5) has been analytically solved in [10]. Its solution, that is off-line computed, is given by a correlation matrix C^{ϕ}, that identifies the stochastic process to be used for generating sensing vectors. The tuning of τ on a proper range is not critical since it does not appreciably alter the overall system performance [24, 25].

Note that this approach is perfectly compatible with the generation of a binary antipodal sensing matrix[2]. Yet, while the generation of a (Gaussian) random vector with a prescribed correlation is quite easy, the generation of random binary antipodal vectors given C^{ϕ} is a substantially more complex task. To this aim, many different approaches have been proposed in the literature (see, e.g., [26, 27, 28]). The approach used in this paper is the so-called linear probability feedback generator [27].

3.3. Zeroing Approach

The zeroing approach [8, 11] aims to further reduce the complexity, and so the power consumption, of the encoding stage with respect to the adoption of a binary antipodal Φ. The basic idea is to alter a binary antipodal Φ by *zeroing* some entry and thus allowing $\Phi \in \{+1, 0, -1\}^{m \times n}$. More formally, let us introduce the zeroing parameter ζ such that $0 < \zeta \leq m$. Starting from an antipodal Φ, obtained either with the standard or the rakeness approach, we consider the generic j-th column of Φ, $j = 0 \ldots n-1$, and we set to zero $m - \zeta$ elements entries randomly chosen, thus leaving only ζ non-zero elements in each column.

By doing this the computational complexity of (1) is reduced from $m \cdot n$ sums to $\zeta \cdot n$ sums. The drawback is a perturbation of the statistical characterization of Φ. This has almost no effect when comparing with the standard CS approach. However, when starting from a Φ generated accordingly to the rakeness paradigm (we will always make this assumption from now on), by zeroing some entries of Φ we are altering the C^{ϕ} and so reducing the value of $\rho(\underline{\phi}, \underline{x})$. In conclusion we are expecting a reduction of the rakeness effect, with system performance degrading back to that of the standard CS approach.

Yet, we will consider in the following a case study in which we can show that

- the performance (in terms of signal reconstruction quality) of the zeroing approach is higher than the standard approach for almost all values of ζ;
- the signal reconstruction quality is decreasing with ζ, while energy requirements are increasing with ζ. This trade-off allows many advantages as shown in the following.

Furthermore, the zeroing approach ensures that the computational effort associated to each incoming sample is limited and fixed by $\zeta < m$. For example, considering the projection stage employed in a parallel structure, no more than ζ parallel paths must be considered, while for a programmable architecture, if at most ζ iterations are needed in each sample time. More details on this case are reported in a next section.

4. CS in biomedical monitors

The standard CS (SCS, where the sensing matrix Φ is made by binary antipodal randomly generated entries), the rakeness CS (RCS, with binary antipodal Φ generated with given C^{ϕ}) and the zeroing CS with ζ non-null entries each column (ZCS-ζ, obtained by randomly zeroing elements of a rakeness-based Φ) have been tested on real ECG signals available from the MIT-BIH arrhythmia database [29]. For the sake of illustration, here we present results from 71.1 s of the record 101. This signal is sampled at 360 Hz, and we estimate its signal-to-quantization noise ratio (SQNR) as 38.5 dB.

The signal has been partitioned in 50 non-overlapping time windows with $n = 512$ samples each, leading to $T \approx 1.42$ s. Different values of m are used; for each one, a unique sensing matrix Φ has been selected by means of preliminary tests on synthetic ECGs [30]. More specifically, we generated 100 synthetic ECGs and: (i) for the

[2] note that the energy of antipodal sampling sequences in (5) is always $e = n$.

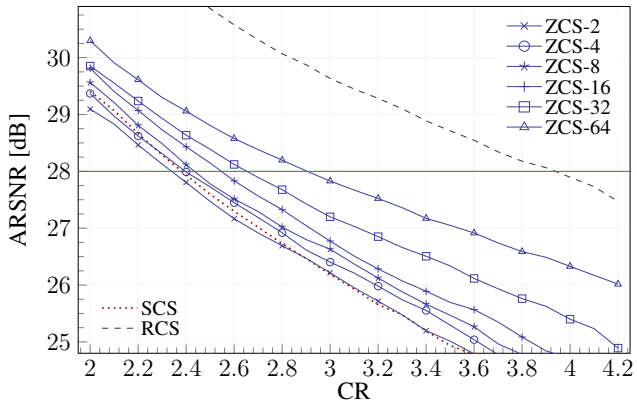

Figure 4: Average RSNR as a function of the compression ratio (CR) for SCS, RCS and ZCS-ζ for different encoding strategies.

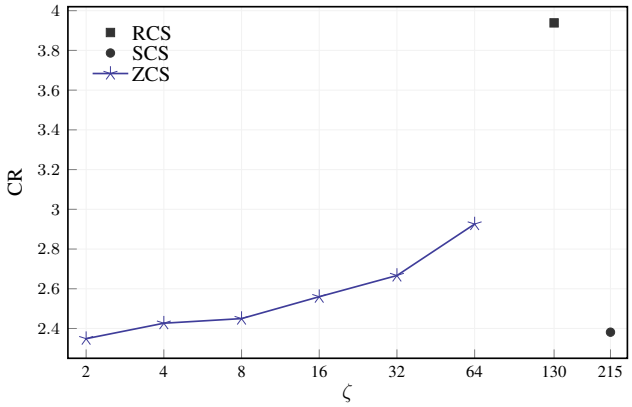

Figure 5: CS as a function ζ considering ABPDN for a fixed quality of service.

SCS case, we chose Φ as the random antipodal matrix ensuring the best average reconstruction performance; (*ii*) for the RCS case, using the same synthetic ECG generators to estimate the correlation profile C^x required by (5) we preliminary evaluate C^ϕ in order to generate a pool of sensing matrices and then we chose Φ again as the matrix guaranteeing best performance; (*iii*) for the ZCS-ζ case, we take the optimal matrix Φ in (*ii*) and randomly set to zero $m - \zeta$ entries in every column. This approach has been adopted since it ensures maximum fairness, being Φ not biased on any particular real signal.

Furthermore, three different algorithms have been used to investigate the effect of zeroing on different reconstruction strategies where as main figure of merit we consider the reconstruction signal-to-noise ratio, defined as

$$\text{RSNR} = \left[\frac{\|x\|_2}{\|x - \hat{x}\|_2}\right]_{\text{dB}}$$

and averaged over the 50 considered time windows to obtain the Average RSNR (ARSNR).

First, we consider a decoding procedure belongs to the class of algorithms solving (3) where reconstruction vectors $\hat{x}$ are computed by an overcomplete dictionary, that is a redundant discrete wavelet transform in the family of Symmlet-6 wavelet transformations with $J = 4$ sub-bands, i.e. for which $N = (J + 1)n$ [31, Chapter 5.2]. To enforce sparsity in this domain the solution of convex problem (3) is achieved by using Douglas-Rachford splitting as implemented in UNLocBoX [32, 33, 34]. Results in terms of ARSNR are shown in Figure 4 where the SCS, the RCS and the ZCS approach with different values of ζ are considered. As expected, the RCS outperforms all other approaches in terms of ARSNR for all considered CR values. Furthermore, performance for the ZCS is increasing with the ζ value, and drops to that of the SCS only for $\zeta = 2$.

As an additional figure of merit, CR assessed to a minimum target ARSNR is considered. For this decoding approach, we set this threshold at 28 dB where this quality of service represents, in our point of view, a good trade-off between the SQNR and a good visual representation of the signal. As a visual confirmation, Figure 6 shows short chunks of reconstructed ECG signals at the target ARSNR by RCS, SCS and ZCS-ζ. To highlight the effect of zeroing, in Figure 5 we also report results in terms of maximum CR as function of ζ by fixing ARSNR to 28 dB. Both SCS and RCS work with sensing matrices composed of antipodal symbols such that $\zeta = m$.

Note that the role of ZCS is not to outperform RCS in terms of either maximum CR for a fixed ARSNR or to increase ARSNR for a fixed CR. ZCS introduces another degree of freedom in the design of the entire sensor node relaxing the amount of elementary operations needed to compute y. Table 2 reports results which summarizes this novel trade-off. Here we report the maximum CR (and corresponding m) for all considered CS approaches, along with the number of sums required to compute y.

Table 2: CR, m and # of sums needed to target a SNR=28 dB for various CS cases.

CS		CR	m	# sums
SCS		2.38	215	1.3×10^5
RCS		3.94	130	6.7×10^4
ZCS	$\zeta = 64$	2.93	175	3.2×10^4
	$\zeta = 32$	2.67	192	1.6×10^4
	$\zeta = 16$	2.56	200	8192
	$\zeta = 8$	2.45	209	4096
	$\zeta = 2$	2.35	218	1024

With respect to traditional approaches such as SCS and RCS, the zeroing technique can greatly reduce the computational cost required to compress a signal instance, at the expense of a lower CR (with respect to RCS), i.e. of increasing the minimum m. It is important to recall that a reduced m is reflected into a reduced amount of bits that must be transmitted/stored for each window. Yet, in scenarios where the computational cost is not negligible with respect to the transmission/storing one, ZCS can be used to minimize the power requirements of the entire sensor node. This will be shown in a few examples in the fol-

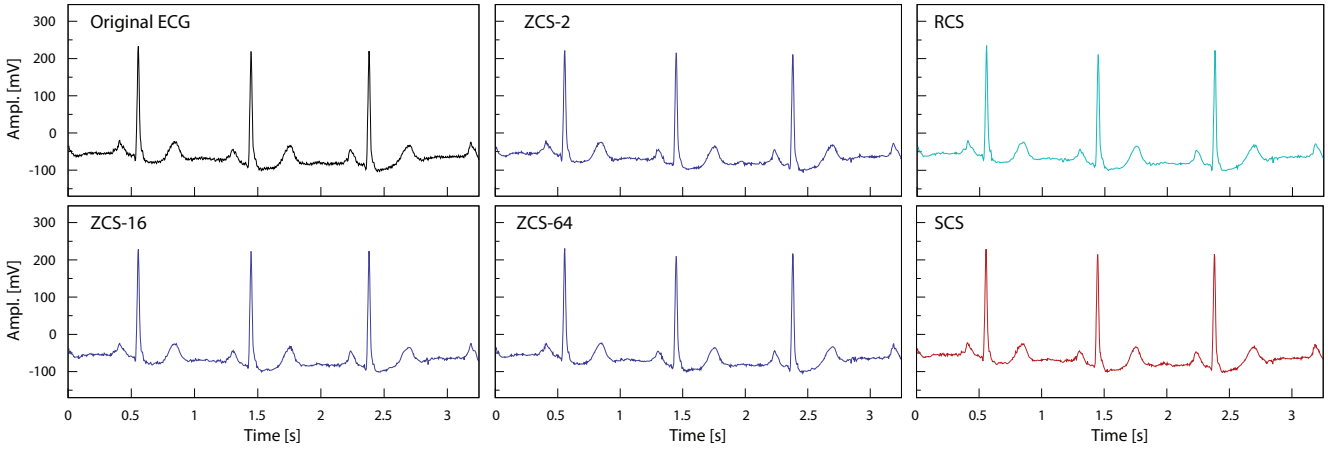

Figure 6: Chunks of reconstructed signal by ABPDN at target RSNR for all CS approaches.

lowing section. In this case, the choice of ζ is a trade-off, depending on the power consumption of the DSP performing CS and on the energy required to store/transmit the m measurements composing the vector y.

This first set of results partially covers the characterization of a ZCS-based system. To have a complete overview, considerations on the role of the decoder stage are needed. This topic involves a huge amount of works in literature (e.g., [22, 35, 36]) where different approaches cope with many possible scenarios. Among them, a first corner is given by high complexity decoding procedures aiming to increase as much as possible the affinity of the reconstructed signal to the original one, as for the first decoder considered above. At the opposite corner there are greedy approaches aiming to reduce as possible the computational cost of the decoder stage with a limited lossy in the reconstruction performances. Obviously, many other cases can be ideally placed between these two corners.

All these solutions are intrinsically non-linear, and the effect of the zeroing on the sensing matrix entries can be pursued only by simulations. Additionally to results describing the first proposed decoding algorithm (as for Figure 4), in both plots in Figure 7 the results of the effects of the zeroing considering two other decoders are presented. In particular, Figure 7-(a) refers to the solution of (2) by the large-scale solver named SPGL1 [36] while Figure 7-(b) shows performances obtained by a popular greedy approach, the orthogonal matching pursuit (OMP) [35]. The comparison of these results clearly shows that the performance degradation expected by ZCS with respect to RCS is strongly depending on the considered decoding strategy. This can be either negligible, as for ZSC-64 where OMP is used for the signal reconstruction, or significant as shown in in the first set of results presented in this section.

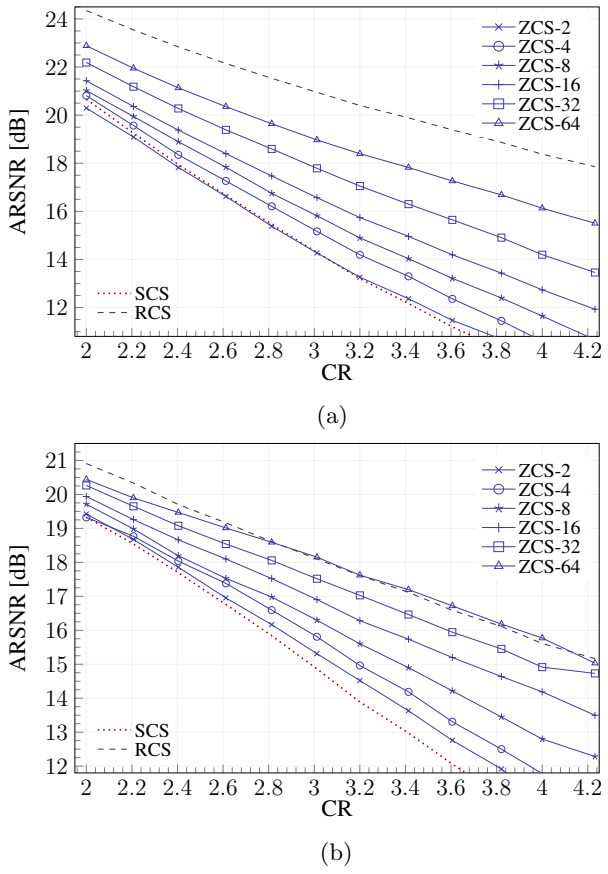

Figure 7: Average RSNR as a function of the compression ratio (CR) for SCS, RCS and ZCS-ζ for different encoding strategies: signal reconstruction by SPGL1 in (a) and by OMP in (b).

5. Evaluation

To accurately profile the execution of the CS algorithms on the multi-core digital processor, we utilized cycle-accurate simulator [37], written in SystemC. The virtual platform contains a model of the multi-core DSP architecture described in Section 2, which can be enriched with power consumption numbers for the memory subsystem and the rest of the logic (interconnect, cores) extracted from a RTL-equivalent architecture [17]. The robustness

of the results hereby presented is based on the very low misalignment[3] of the SystemC platform and the RTL platform (silicon-equivalent), which is below 7%.

Based on the analysis of the CS code (detailed in the next section), we chose as a baseline configuration the architectural parameters reported in Table 3.

Table 3: MC-DSP architecture baseline configuration.

parameter	*value*
number of PEs	8
IM capacity (per-core)	1KB
TCDM number of banks	16
TCDM capacity	192KB
stack size	512B

For what concerns the CS data structures (samples, sensing matrix, etc.), static data allocation is performed by means of cross-compiler attributes and linker script sections. Finally, the application support in [37] allows for an easy bare-metal parallelization of the CS algorithm by means of explicit pointers to code sections. Such mechanism allows to statically bind an input data channel to a given PE, avoiding the need of a full runtime system which would add extra overhead and penalize the energy efficiency of the system.

5.1. CS on MC-DSP

The three different CS paradigms introduced in Section 4 (SCS, RCS and ZCS-ζ) are suitble for two different algorithmic implementations due to their sensing matrix (Φ) structure. In both cases the multi-core DSP is operating in a parallel fashion, where each processing element is compressing the input data related to the associated input channel. The SCS and the RCS compression algorithms rely on the same computation, however they deploy two different sensing matrix designs.

A straightforward implementation for such algorithms is a matrix multiplication, where the matrix is composed by antipodal symbols, i.e. the elements are either $+1$ or -1. In this case the matrix is named full (sparsity = 0%) and the projection of the input vector over the sensing matrix rows can be implemented with a standard Full Matrix (FM) multiplication, shown in Listing 1, and the Φ matrix can be represented with `signed char` datatype.

[3]The alignment is intended as the difference in terms of execution cycles for a matrix multiplication benchmark (256x256), highly similar to the straightforward CS implementation, on a 4 PEs architecture running on both the SystemC and RTL platforms.

```
void cs_fm(short *x, short *y)
{

  unsigned short i, j;

  for (i = 0; i < m; i++)
    for (j = 0; j < n; j++)
      y[i] += phi[i][j] * x[j];

}
```

Listing 1: Full matrix version of CS.

If we consider the ZCS-ζ case, Φ is a ternary sparse sensing matrix ($+1, 0$ or -1 elements) and this motivated a more efficient algorithm implementation based on Look-Up Tables (LUTs), shown in Listing 2. Obviously, the sparsity of such matrix is strongly related to ζ.

Such algorithm implementation encodes the information related to the position and the sign of the non-zero elements of the sensing matrix in two different LUTs, LUT Index and LUT Sign, respectively accessed via the `*lip` and `*lsp` pointers.

```
void cs_lut(short *x, short *y)
{

  unsigned short i, j;
  unsigned char *lip = &lut[0];
  signed char *lsp = &lut_s[0];

  for (i = 0; i < n; i++)
    for (j = 0; j < k; j++)
      y[*(lip++)] += (*(lsp++)) * x[i];

}
```

Listing 2: Look-up table version of CS.

In Listing 2 a case with input size $n = 512$ and a CR > 2 is considered, allowing for respectively `unsigned char` and `signed char` data types, while a `short` representation applies for x and y due to the 12-bit resolution ADC considered.

To compare all mentioned CS approached, in the rest of the paper we refer to the operative points reported in Table 2 where the decoding algorithm and the target reconstruction quality are fixed. To achieve the same target reconstruction quality, due to the different algorithmic implementation of the matrix multiplication characterizing CS encoders and to the different number of measurements required to achieve the desired ARSNR, the SCS, RCS and ZCS-ζ cases lead to different memory footprints in the DSP data memory. For all cases the TCDM memory has to allocate the input samples s_i, considering $n = 512$, 8 channels and 12-bit ADC resolution it leads to 8KB using `short` datatypes. For the measurement vectors y associated to each PE and the sensing matrix Φ (or the LUT structures), the requirements are different and this information is reported in Table 4.

5.2. Experimental Results

We present here the energy consumption analysis on the sensor node, considering the three different stages.

Table 4: MC-DSP data memory footprint requirements for SCS, RCS and ZCS-ζ, targeting RSNR=28dB as in Section 4.

algorithm	*type*	*m*	*output* (B)	*sensing* (KB)
SCS	FM	256	4096	128
RCS	FM	130	2080	65
ZCS $\zeta = 64$	LUT	175	2800	64
ZCS $\zeta = 32$	LUT	192	3072	32
ZCS $\zeta = 16$	LUT	200	3200	16
ZCS $\zeta = 8$	LUT	209	3344	8
ZCS $\zeta = 2$	LUT	218	3488	2

Compression. To evaluate the energy consumption of the SCS, RCS, ZCS-ζ algorithms running on the MC-DSP, we instrumented our architectural simulator [37] in order to collect the activity of every component of the modeled DSP architecture. RTL simulations were run to profile the power consumption of the architectural elements to be later back-annotated inside the power models of the SystemC simulator. As a design corner we considered the operating point $T = 20^o$C, $V_{dd} = 0.5$V at 5.5MHz in a 28nm FDSOI technology [16], choosing the RVT library. Such operating point is indeed within the range of maximum efficiency of a real silicon implementation of an almost identical[4] architecture, as presented in [38]. Moreover, considering the memory-bound nature of the CS compression task, the bandwidth requirements for PE-TCDM communication imply a higher supply voltage (set at 0.8V) for the memory sub-system to sustain the throughput. To increase the energy efficiency the DSP logic is clock-gated during the idle phases, eliminating the dynamic power, while to reduce the leakage we consider a reverse body bias voltage $V_{RBB} = 0.5$V, leading to a $\approx 65\%$ leakage power reduction [38].

Figure 8 shows on the x-axis the different CS algorithms while on the y-axis the energetic cost for multi-channel input compression. First we can notice that RCS and ZCS improves the energy efficiency when compared to SCS. Indeed the RCS approach consumes $\approx 37\%$ less energy than SCS, while achieving a larger compression. ZCS leads to an additional energy gain case up to $\approx 91\%$ with respect to the SCS case, thanks to the efficient LUT-based algorithmic implementation and the increased sparsity of the sensing matrix.

Transmission/Storage Phase. To complete the analysis we evaluated the energy consumption of the last stage of our biomedical monitor. We derived a simple model to cope with some of the most widespread or forthcoming technologies for either transmission or storage.

Different protocols were considered for the transmission option ranging from the power-hungry Near Field Communication (NFC) and the Bluetooth Low Energy (BLE) protocols, to the efficient Narrow Band (NB) solution. In the storage scenario, i.e. memorizing the compressed data in a Non Volatile Memory (NVM) for a long-term monitoring scenario, we considered different promising technologies ranging from the Resistive RAM (ReRAM) to the Conductive Bridging RAM (CBRAM). Table 5 reports the energy per bit for both transmission and storage of such state-of-the art solutions for biomedical monitors, as carefully detailed in [39].

[4]Since the chip fabricated in [38] has 4 cores and 8 TCDM banks, the power consumption numbers were scaled up for the architecture considered in this work.

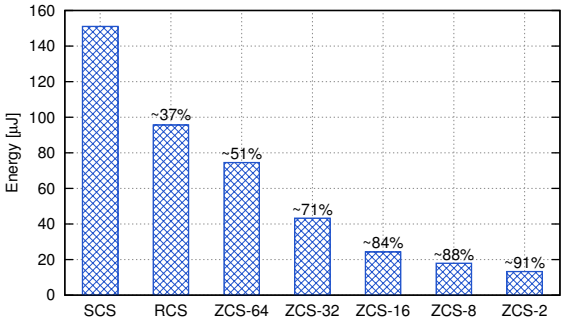

Figure 8: Energy spent in the compression stage for SCS, RCS and ZCS-ζ. Energy savings with respect to the SCS case are reported on top of each bar.

Table 5: Energy per transmitted/stored bit assuming different transmission (TX) and storage (NVM) technologies.

		Energy [nJ/bit]	Reference
TX	BLE	1	[40]
	NB	0.1	[41]
	HBC	0.24	[42]
	NFC	10	[43]
NVM	ReRAM	2	[44]
	STT-MRAM	0.1	[45]
	FLASH	0.01	[46]
	CBRAM	0.001	[47]

Data Transmission. The first set of results compares the different compression schemes in terms of the total energy consumed by the bio-sensing node to acquire a window of samples at a fixed reconstruction quality (28dB), configured with the different transmission technologies. Each plot in Figure 9 refers to a given transmission technology and reports with different bars (x-axis) the different compression schemes, different colors are used for the different contribution to the total energy: DSP (bottom part of each bar), transmission (top part of each bar). From the figure we can notice that the RCS achieves for all the cases the lowest transmission energy while ZCS can trade-off this gain with lower DSP energy. For this reason the proposed RCS and ZCS always outperform standard CS in terms of energy efficiency. When comparing to the no compression case (Shannon sampling), interesting results can be observed:

- low transmission cost (NB, HBC): for these technologies the DSP power dominates the transmission mak-

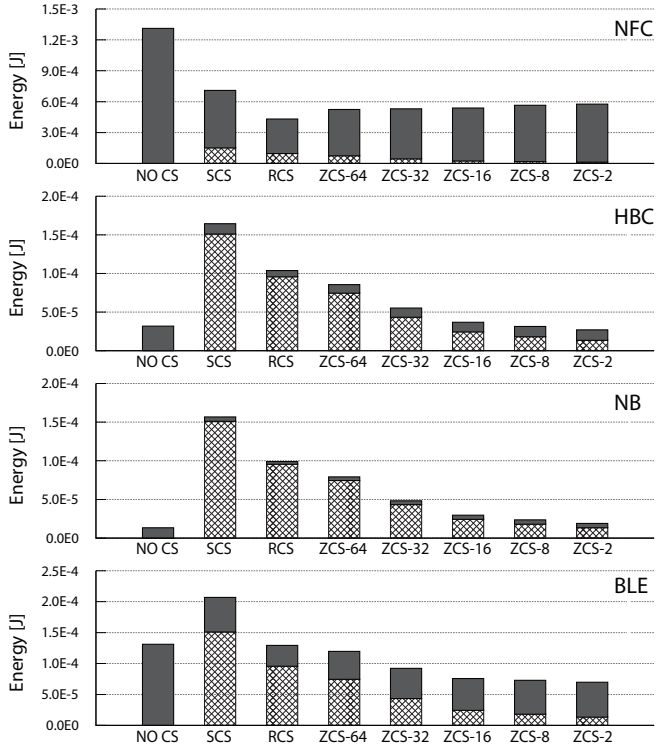

Figure 9: Energy per window, considering different TX technologies. From top to bottom: Near Field Communication, Human Body Channel, Narrow Band and Bluetooth Low Energy.

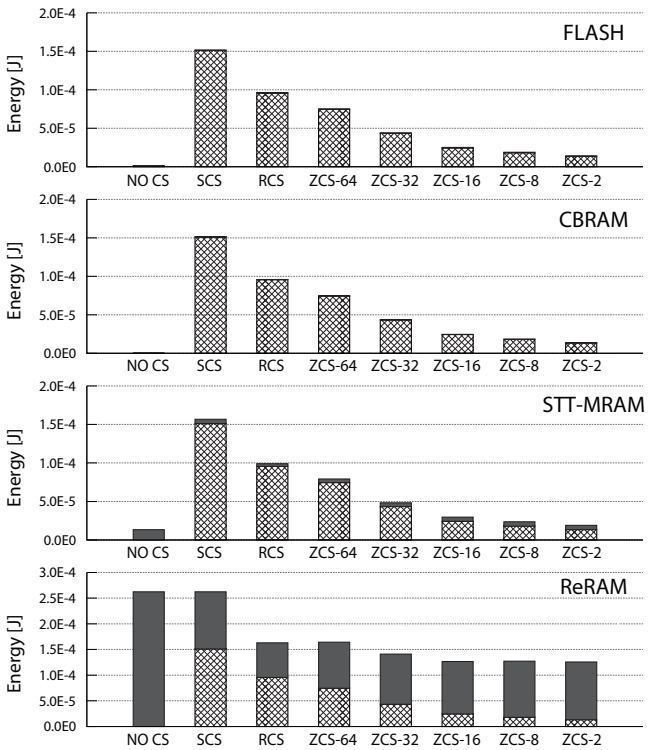

Figure 10: Energy per window for different NVM technologies. From top to bottom: Flash technology, Conductive Bridge RAM, Spin-transfer torque magnetic RAM and Resistive RAM.

ing the NO CS case the most energy efficient one, as it does not require DSP computation. In this condition however the proposed zeroing strategy shows its effectiveness as ZCS-2 is capable of hiding the DSP cost with lower transmission rate reducing the gap against NO CS.

- high transmission cost (NFC): for this technology the transmission power dominates the DSP. As a result RCS achieves the best energy-efficiency. With a reduction of 67% w.r.t. NO CS and 37% w.r.t. the SCS case. Even if the proposed zeroing approach is still more energy-efficient than the NO CS and SCS it is less convenient than the RCS. ZCS-64 achieves the 60% of energy saving w.r.t. NO CS which decreases to the 56% for ZCS-2.
- intermediate transmission cost (BLE): for this technology the ZCS-2 approach shows the best energy-efficiency performance, leading to a gain of 47% w.r.t. the NO CS case, while RCS saves only the 2% and SCS increases the overall energy of 58%.

We can conclude that rakeness and zeroing always outperform the standard CS and are preferable to Shannon sampling when transmission cost is relevant, which is the case of today's mature technologies (BLE, NFC).

NVM Storage. The second set of results analyses the case in which the bio-sensing node stores internally the sampled values in different non-volatile memory technologies. Similarly to the previous figure, each plot in Figure 10 refers to a different NVM technology.We can notice that STT-MRAM, CBRAM and FLASH configurations are characterized by a negligible storage energy cost which makes the NO CS case more energy-efficient. However when ReRAM technology is used for the internal samples storage conclusions changes and the proposed rakeness and zeroing approaches pay-off. Indeed in this configuration the ZCS-2 leads to best energy efficiency with 52% of energy-efficiency gain w.r.t. the NO CS case while the RCS achieves the 38% of energy saving. In this case NO CS and SCS have similar energy consumption. Differently from the transmission case, where data is stored in the bio-sensing node, also the storage footprint is important due to the limited NVM memory which can be embedded in the monitoring node. Table 6 shows the maximum

Table 6: Monitoring time as a function of different NVM storage capacities (technology independent).

	32KB (s)	1MB (m)	128MB (h)	1GB (d)
NO CS	2.82	1.51	3.21	1.07
SCS	6.66	3.55	7.58	2.53
RCS	11.04	5.89	12.56	4.19
ZCS $\zeta = 64$	8.26	4.41	9.40	3.13
ZCS $\zeta = 32$	7.61	4.06	8.66	2.89
ZCS $\zeta = 16$	7.23	3.86	8.23	2.74
ZCS $\zeta = 8$	6.79	3.62	7.72	2.57
ZCS $\zeta = 2$	6.60	3.52	7.51	2.50

Table 7: Energy consumption and storage requirements for one monitoring window (1.412 s) for different NVM technologies.

	ReRAM Energy [μJ]	STT-MRAM Energy [μJ]	CBRAM Energy [μJ]	FLASH Energy [μJ]	Storage [KB]
NO CS	262.14	13.11	0.13	1.31	2048
SCS	262.32	156.77	151.27	138.99	868
RCS	162.67	98.95	95.63	90.33	524
ZCS $\zeta = 64$	164.14	79.02	74.58	72.10	700
ZCS $\zeta = 32$	140.56	48.15	43.33	44.92	760
ZCS $\zeta = 16$	126.69	29.41	24.34	28.39	800
ZCS $\zeta = 8$	127.01	23.40	18.01	22.88	852
ZCS $\zeta = 2$	125.42	18.89	13.34	18.83	876

monitoring time for different embedded storage size.

If we consider a NVM storage size of 128MB, RCS allows to store more than 12 hours of samples while the SCS allows only to store 7 hours of samples and without compressive sensing (NO CS) only 3 hours can be stored. The zeroing approach is in between the SCS and RCS values.

Finally, Table 7 summarizes the results showing that for all the configurations the proposed RCS and ZCS schemes outperforms standard CS in both energy-efficiency and storage effectiveness. Compressive sensing approach is always effective in reducing the memory storage area with a 4x improvement in storage footprint w.r.t NO CS for the rakeness case. When highly energy-efficient storage technology are used (CBRAM and FLASH) the proposed ZCS can reduce the related computational overhead of the 88% w.r.t. standard CS but still consuming one (two for FLASH) order of magnitude more energy than the no compression case. Last, for the ReRAM technology the proposed techniques (RCS and ZCS) are the optimal one for both storage efficiency and energy efficiency.

6. Conclusion

To achieve minimal energy consumption in low-cost WBSN-based biosignal monitors both architectural and signal processing aspects must be considered. The rakeness and the zeroing approaches for CS has been taken into account, enabling trading off the computation workload with the number of measurements for later transmission or storage. In this paper such tradeoffs have been evaluated by considering a multi-core DSP and different technologies for storage or transmission. Experimental results showed that the zeroing approach proves to be more energy efficient with respect both to the standard CS and the rakeness CS when energy requirement for storage or transmission is significant. Moreover, when compression energy and storage energy are comparable, such approaches allow the flexibility of several design choices for what concerns energy consumption and monitoring time.

Acknowledgments

This work was partially supported by the ICYSoC RTD project (no. 20NA21 150939), evaluated by the Swiss NSF and funded by Nano-Tera.ch with Swiss Confederation financing.

References

[1] World Health Organization.
URL http://www.who.int/mediacentre/factsheets/fs317

[2] A. Munir, A. Gordon-Ross, S. Ranka, Multi-core embedded wireless sensor networks: Architecture and applications, Parallel and Distributed Systems, IEEE Transactions on 25 (6) (2014) 1553–1562. doi:10.1109/TPDS.2013.219.

[3] X. Zhang, H. Jiang, L. Zhang, C. Zhang, Z. Wang, X. Chen, An energy-efficient asic for wireless body sensor networks in medical applications, Biomedical Circuits and Systems, IEEE Transactions on 4 (1) (2010) 11–18. doi:10.1109/TBCAS.2009.2031627.

[4] E. J. Candes, J. K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Communications on Pure and Applied Mathematics 59 (8) (2006) 1207–1223. doi:10.1002/cpa.20124.

[5] D. L. Donoho, Compressed Sensing, IEEE Transactions on Information Theory 52 (4) (2006) 1289–1306. doi:10.1109/TIT.2006.871582.

[6] D. Gangopadhyay, E. Allstot, A. Dixon, K. Natarajan, S. Gupta, D. Allstot, Compressed Sensing Analog Front-End for Bio-Sensor Applications, IEEE Journal of Solid-State Circuits 49 (2) (2014) 426–438. doi:10.1109/JSSC.2013.2284673.

[7] F. Pareschi, P. Albertini, G. Frattini, M. Mangia, R. Rovatti, G. Setti, Hardware-algorithms co-design and implementation of an analog-to-information converter for biosignals based on compressed sensing, Biomedical Circuits and Systems, IEEE Transactions ondoi:10.1109/TBCAS.2015.2444276.

[8] D. Bortolotti, M. Mangia, A. Bartolini, R. Rovatti, G. Setti, L. Benini, Rakeness-based compressed sensing on ultra-low power multi-core biomedicai processors, in: Design and Architectures for Signal and Image Processing (DASIP), 2014 Conference on, 2014, pp. 1–8. doi:10.1109/DASIP.2014.7115599.

[9] H. Mamaghanian, N. Khaled, D. Atienza, P. Vandergheynst, Compressed Sensing for Real-Time Energy-Efficient ECG Compression on Wireless Body Sensor Nodes, IEEE Transactions on Biomedical Engineering 58 (9) (2011) 2456–2466. doi:10.1109/TBME.2011.2156795.

[10] M. Mangia, R. Rovatti, G. Setti, Rakeness in the Design of Analog-to-Information Conversion of Sparse and Localized Signals, IEEE Transactions on Circuits and Systems I: Regular Papers 59 (5) (2012) 1001–1014. doi:10.1109/TCSI.2012.2191312.

[11] M. Mangia, D. Bortolotti, A. Bartolini, F. Pareschi, L. Benini, R. Rovatti, G. Setti, Long-term ecg monitoring with zeroing compressed sensing approach, in: Nordic Circuits and Systems Conference (NORCAS): NORCHIP International Symposium on System-on-Chip (SoC), 2015, 2015, pp. 1–4. doi:10.1109/NORCHIP.2015.7364394.

[12] A. Dogan, J. Constantin, M. Ruggiero, A. Burg, D. Atienza, Multi-core architecture design for ultra-low-power wearable

health monitoring systems, in: Design, Automation Test in Europe Conference Exhibition (DATE), 2012, 2012, pp. 988–993. doi:10.1109/DATE.2012.6176640.

[13] R. Dreslinski, B. Zhai, T. Mudge, D. Blaauw, D. Sylvester, An energy efficient parallel architecture using near threshold operation, in: Parallel Architecture and Compilation Techniques, 2007. PACT 2007. 16th International Conference on, 2007, pp. 175–188. doi:10.1109/PACT.2007.4336210.

[14] D. Bortolotti, A. Bartolini, C. Weis, D. Rossi, L. Beninio, Hybrid memory architecture for voltage scaling in ultra-low power multi-core biomedical processors, in: Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014, 2014, pp. 1–6. doi:10.7873/DATE.2014.182.

[15] D. Bortolotti, H. Mamaghanian, A. Bartolini, M. Ashouei, J. Stuijt, D. Atienza, P. Vandergheynst, L. Benini, Approximate compressed sensing: Ultra-low power biosignal processing via aggressive voltage scaling on a hybrid memory multi-core processor, in: Proceedings of the 2014 International Symposium on Low Power Electronics and Design, ISLPED '14, ACM, New York, NY, USA, 2014, pp. 45–50. doi:10.1145/2627369.2627629.

[16] D. Jacquet, F. Hasbani, P. Flatresse, R. Wilson, F. Arnaud, G. Cesana, T. D. Gilio, C. Lecocq, T. Roy, A. Chhabra, C. Grover, O. Minez, J. Uginet, G. Durieu, C. Adobati, D. Casalotto, F. Nyer, P. Menut, A. Cathelin, I. Vongsavady, P. Magarshack, A 3 ghz dual core processor arm cortex tm -a9 in 28 nm utbb fd-soi cmos with ultra-wide voltage range and energy efficiency optimization, IEEE Journal of Solid-State Circuits 49 (4) (2014) 812–826. doi:10.1109/JSSC.2013.2295977.

[17] M. Gautschi, D. Rossi, L. Benini, Customizing an open source processor to fit in an ultra-low power cluster with a shared l1 memory, in: Proceedings of the 24th Edition of the Great Lakes Symposium on VLSI, GLSVLSI '14, ACM, New York, NY, USA, 2014, pp. 87–88. doi:10.1145/2591513.2591569.

[18] A. Rahimi, I. Loi, M. Kakoee, L. Benini, A fully-synthesizable single-cycle interconnection network for shared-l1 processor clusters, in: Design, Automation Test in Europe Conference Exhibition (DATE), 2011, 2011, pp. 1–6. doi:10.1109/DATE.2011.5763085.

[19] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, A pragmatic look at some compressive sensing architectures with saturation and quantization, IEEE Journal on Emerging and Selected Topics in Circuits and Systems 2 (3) (2012) 443–459. doi:10.1109/JETCAS.2012.2220392.

[20] E. J. Candes, T. Tao, Decoding by linear programming, IEEE Transactions on Information Theory 51 (12) (2005) 4203–4215. doi:10.1109/TIT.2005.858979.

[21] D. Donoho, Compressed sensing, Information Theory, IEEE Transactions on 52 (4) (2006) 1289–1306. doi:10.1109/TIT.2006.871582.

[22] D. Needell, J. Tropp, Cosamp: Iterative signal recovery from incomplete and inaccurate samples, Applied and Computational Harmonic Analysis 26 (3) (2009) 301 – 321. doi:10.1016/j.acha.2008.07.002.

[23] E. J. Cands, Y. C. Eldar, D. Needell, P. Randall, Compressed sensing with coherent and redundant dictionaries, Applied and Computational Harmonic Analysis 31 (1) (2011) 59 – 73. doi:10.1016/j.acha.2010.10.002.

[24] N. Bertoni, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, Correlation tuning in compressive sensing based on rakeness: A case study, in: Electronics, Circuits, and Systems (ICECS), 2013 IEEE 20th International Conference on, 2013, pp. 257–260. doi:10.1109/ICECS.2013.6815403.

[25] M. Mangia, J. Haboba, R. Rovatti, G. Setti, Rakeness-based approach to compressed sensing of ecgs, in: Biomedical Circuits and Systems Conference (BioCAS), 2011 IEEE, 2011, pp. 424–427. doi:10.1109/BioCAS.2011.6107818.

[26] G. Jacovitti, A. Neri, G. Scarano, Texture synthesis-by-analysis with hard-limited gaussian processes, Image Processing, IEEE Transactions on 7 (11) (1998) 1615–1621. doi:10.1109/83.725369.

[27] R. Rovatti, G. Mazzini, G. Setti, Memory-m Antipodal Processes: Spectral Analysis and Synthesis, IEEE Transactions on Circuits and Systems I: Regular Papers 56 (1) (2009) 156–167. doi:10.1109/TCSI.2008.920986.

[28] A. Caprara, F. Furini, A. Lodi, M. Mangia, R. Rovatti, G. Setti, Generation of antipodal random vectors with prescribed non-stationary 2-nd order statistics, Signal Processing, IEEE Transactions on 62 (6) (2014) 1603–1612. doi:10.1109/TSP.2014.2302737.

[29] A.L. Goldberger *et al.*, PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) 215–220. doi:10.1161/01.CIR.101.23.e215.

[30] P. McSharry, G. Clifford, L. Tarassenko, L. Smith, A dynamical model for generating synthetic electrocardiogram signals, IEEE Transactions on Biomedical Engineering 50 (3) (2003) 289–294. doi:10.1109/TBME.2003.808805.

[31] S. Mallat, A Wavelet Tour of Signal Processing: The Sparse Way, 3rd Edition, Academic Press, 2008.

[32] E. J. Candes, Y. C. Eldar, D. Needell, P. Randall, Compressed sensing with coherent and redundant dictionaries, Applied and Computational Harmonic Analysis 31 (1) (2011) 59 – 73. doi:10.1016/j.acha.2010.10.002.

[33] P. Combettes, J. Pesquet, A douglas-rachford splitting approach to nonsmooth convex variational signal recovery, Selected Topics in Signal Processing, IEEE Journal of 1 (4) (2007) 564–574. doi:10.1109/JSTSP.2007.910264.

[34] N. Perraudin, D. Shuman, G. Puy, P. Vandergheynst, UNLocBoX A matlab convex optimization toolbox using proximal splitting methods, ArXiv e-printsarXiv:1402.0779.

[35] J. Tropp, A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, Information Theory, IEEE Transactions on 53 (12) (2007) 4655–4666. doi:10.1109/TIT.2007.909108.

[36] E. van den Berg, M. P. Friedlander, Probing the pareto frontier for basis pursuit solutions, SIAM Journal on Scientific Computing 31 (2) (2008) 890–912. doi:10.1137/080714488.

[37] D. Bortolotti, C. Pinto, A. Marongiu, M. Ruggiero, L. Benini, Virtualsoc: A full-system simulation environment for massively parallel heterogeneous system-on-chip, in: Parallel and Distributed Processing Symposium Workshops PhD Forum (IPDPSW), 2013 IEEE 27th International, 2013, pp. 2182–2187. doi:10.1109/IPDPSW.2013.177.

[38] D. Rossi, A. Pullini, I. Loi, M. Gautschi, F. K. Grkaynak, A. Bartolini, P. Flatresse, L. Benini, A 60 gops/w, 1.8 v to 0.9 v body bias {ULP} cluster in 28 nm {UTBB} fd-soi technology, Solid-State Electronics 117 (2016) 170 – 184. doi:http://dx.doi.org/10.1016/j.sse.2015.11.015.

[39] D. E. Bellasi, L. Benini, Energy-efficiency analysis of analog and digital compressive sensing in wireless sensors, Circuits and Systems I: Regular Papers, IEEE Transactions on 62 (11) (2015) 2718–2729. doi:10.1109/TCSI.2015.2477579.

[40] Y.-H. Liu, X. Huang, M. Vidojkovic, A. Ba, P. Harpe, G. Dolmans, H. de Groot, A 1.9 nj/b 2.4 ghz multistandard (bluetooth low energy/zigbee/ieee802. 15.6) transceiver for personal/body-area networks, in: Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International, IEEE, 2013, pp. 446–447. doi:10.1109/ISSCC.2013.6487808.

[41] G. Papotto, F. Carrara, A. Finocchiaro, G. Palmisano, A 90nm cmos 5mb/s crystal-less rf transceiver for rf-powered wsn nodes, in: Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International, IEEE, 2012, pp. 452–454. doi:10.1109/ISSCC.2012.6177087.

[42] J. Bae, K. Song, H. Lee, H. Cho, L. Yan, H. J. Yoo, A 0.24nj/b wireless body-area-network transceiver with scalable double-fsk modulation, in: 2011 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011, pp. 34–36. doi:10.1109/ISSCC.2011.5746207.

[43] W. L. Lien, T. Y. Choke, Y. C. Tan, M. Kong, E. C. Low, D. P. Li, L. Jin, H. Zhang, C. H. Leow, S. L. Chew, et al., 9.1 a self-calibrating nfc soc with a triple-mode reconfigurable pll and a

single-path picc-pcd receiver in 0.11μm cmos, in: Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International, IEEE, 2014, pp. 158–159. doi:10.1109/ISSCC.2014.6757380.

[44] M.-F. Chang, C.-W. Wu, C.-C. Kuo, S.-J. Shen, K.-F. Lin, S.-M. Yang, Y.-C. King, C.-J. Lin, Y.-D. Chih, A 0.5v 4mb logic-process compatible embedded resistive ram (reram) in 65nm cmos using low-voltage current-mode sensing scheme with 45ns random read time, in: Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International, 2012, pp. 434–436. doi:10.1109/ISSCC.2012.6177079.

[45] D. Halupka, S. Huda, W. Song, A. Sheikholeslami, K. Tsunoda, C. Yoshida, M. Aoki, Negative-resistance read and write schemes for stt-mram in 0.13μm cmos, in: Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International, IEEE, 2010, pp. 256–257. doi:10.1109/ISSCC.2010.5433943.

[46] D. Shum, J. Power, R. Ullmann, E. Suryaputra, K. Ho, J. Hsiao, C. Tan, W. Langheinrich, C. Bukethal, V. Pissors, et al., Highly reliable flash memory with self-aligned split-gate cell embedded into high performance 65nm cmos for automotive & smartcard applications, in: Memory Workshop (IMW), 2012 4th IEEE International, IEEE, 2012, pp. 1–4. doi:10.1109/IMW.2012.6213670.

[47] N. Gilbert, Y. Zhang, J. Dinh, B. Calhoun, S. Hollmer, A 0.6v 8 pj/write non-volatile cbram macro embedded in a body sensor node for ultra low energy applications, in: VLSI Circuits (VLSIC), 2013 Symposium on, 2013, pp. C204–C205.