Hardware-Software Inexactness in Noise-aware Design of Low-Power Body Sensor Nodes

Hossein Mamaghanian, Giovanni Ansaloni, Mohamed M. Sabry, David Atienza *and* Pierre Vandergheynst
Embedded Systems Laboratory (ESL) *and* Signal Processing Laboratory (LTS2)
École Polytechnique Fédérale de Lausanne, Switzerland
Email: {hossein.mamaghanian},{giovanni.ansaloni}, {mohamed.sabry}, {david.atienza}, {pierre.vandergheynst} @epfl.ch

I. Introduction and motivation

Wireless Body Sensor Nodes (WBSNs) are miniaturized and ultra-low-power devices, able to acquire and wirelessly transmit biosignals such as electrocardiograms (ECG) for extended periods of times and with little discomfort for subjects [1]. Energy efficiency is of paramount importance for WBSNs, because it allows a higher wearability (by requiring a smaller battery) and/or an increased mean time between charges.

In this paper, we investigate how noise-aware design choices can be made to minimize energy consumption in WBSNs. Noise is unavoidable in biosignals acquisitions, either due to external factors (in case of ECGs, muscle contractions and respiration of subjects [2]) or to the design of the front-end analog acquisition block. From this observation stems the opportunity to apply inexact strategies such as on-node lossy compression to minimize the bandwidth over the energy-hungry wireless link [3], as long as the output quality of the signal, when reconstructed on the receiver side, is not constrained by the performed compression.

To maximise gains, ultra-low-power platforms must be employed to perform the above-mentioned Digital Signal Processing (DSP) techniques. To this end, we propose an under-designed (but extremely efficient) architecture that only guarantees the correctness of operations performed on the most significant data (i.e., data most affecting the final results), while allowing sporadic errors for the less significant data [4].

In particular, the paper describes how the knowledge of the noise corrupting biosignals can be leveraged to minimize power consumption on WBSNs. Section II puts our work into context by acknowledging related efforts in the field, Section III describes an application-level solution based on lossy compression and low-power sensing, while Section IV proposes significance-based DSP as an architectural technique for inexact and ultra-low-power biosignals processing. Judicious implementation of these approaches results in substantial increases in WBSN energy efficiency, with minimal degradation of the quality of biosignal acquisitions.

This work has been partially supported by the EC FP7 FET Phidias project (Grant agreement no. 318013.), and the ObeSense (no. 20NA21_143081) and BodyPoweredSenSE (no. 20NA21_143069) RTD projects evaluated by the Swiss NSF and funded by Nano-Tera.ch with Swiss Confederation financing.

As illustrated in Figure 1, our contributions suggest that inexactness offers multiple opportunities for ultra-low-power WBSN designs, that we plan to synergistically exploit in our future works.

II. Related works

In recent years, advanced WBSN platforms have been proposed that, in addition to acquire and wirelessly transmit biosignals, also perform on-board processing. In this context, proposed strategies fall in two categories: in the first one an embedded bio-signal analysis is performed by the WBSN to derive relevant features [2] [5]. By transmitting the analysis results only, this approach minimizes the transmission bandwidth, but it also hides the acquired signals from the users. The second strategy, investigated in this paper, is to compress the signal before transmission. On one hand, compressed signals still require more bandwidth than analysis results. On the other hand, this scenario allows the retrieval of the acquisitions at the receiver, and is therefore preferable when detailed follow-up inspections of the recordings must be performed.

Most biosignals admit a *sparse* representation in a transformed domain. In the case of ECG for instance, only few coefficients carry most of the information in the Digital Wavelet Transform (DWT) domain. An effective compression strategy is therefore to discard less-significant DWT coefficients [6]. Alternatively, the sparse nature of the signals can be exploited to sub-sample the signal and reconstruct it at the receiver side by solving a convex optimization problem. The latter strategy is named Compressive Sensing (CS) [7] [3].

Sparsity also implies that some DWT coefficients are more significant than others, as they contribute more to the quality of the results. Consequently, guarantees on computations of

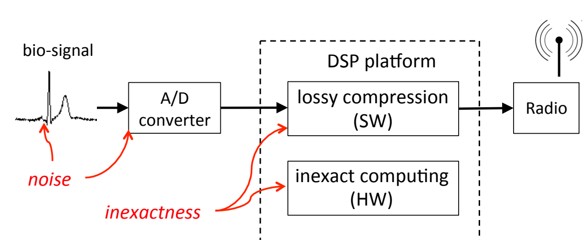


Fig. 1. Investigated WBSN platform. As inputs are partially corrupted by noise, hardware and/or software inexactness is not the quality bottleneck of the system.

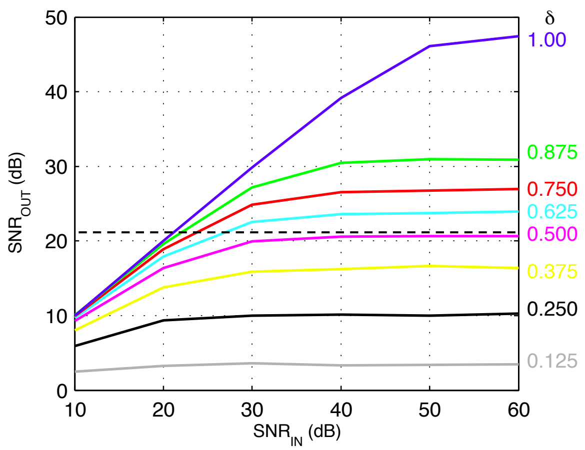


Fig. 2. Signal-to-noise ratio (SNR) of ECG signals after CS reconstruction when varying the SNR before CS, for different compression ratios δ.

less-significant coefficients can be relaxed, resulting in a more energy-efficient implementation of the DSP platform, with a minimal impact on the overall results [4]. Conversely to related works in the general-purpose field [8], approximate computing is here selectively applied, with the goal of preserving the quality of results even under high rates of errors in less significant data.

III. CS-BASED ON-NODE COMPRESSION

Compressed Sensing leverages the sparsity of many biosignal to perform inexact (lossy) compression. It effectively reduces the energy-hungry transmission bandwidth while requiring little computational effort for the encoder, and is therefore a well-suited strategy for maximizing the efficiency of WBSNs. CS decoding algorithms are instead more complex, but they are usually performed on less constrained platforms.

To briefly illustrate CS encoding and decoding, let's define $\mathbf{x} \in \mathbb{R}^N$ as the time series of the input signal, comprising a component of additive white gaussian noise: $\mathbf{x} = \mathbf{x}_{sig} + \mathbf{n}$. The signal admits an (approximate) sparse representation in a dictionary $\mathbf{\Psi}$ (here, DWT) if most coefficients of $\boldsymbol{\alpha} = \mathbf{\Psi}^{-1}\mathbf{x}$ are close to 0. When performing compressed sensing [7], $M < N$ linear combinations of the input are computed on the transmission side: $\mathbf{y} \in \mathbb{R}^M = \mathbf{\Phi}\mathbf{x} = \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\alpha}$, where $\mathbf{\Phi}$ represent the sensing matrix and is independent from the input signal. $\delta = \frac{M}{N}$ is defined as the Compression Ratio (CR), stating the decrease in bandwidth requirements and, ultimately, in the energy employed for wireless communication.

At the receiver side, the signal can be reconstructed by solving a convex optimization problem:

$$\min_{\hat{\alpha} \in \mathbb{R}^N} ||\hat{\boldsymbol{\alpha}}||_1 \qquad \text{subject to:} \qquad ||\mathbf{\Phi}\mathbf{\Psi}\ \hat{\boldsymbol{\alpha}} - \mathbf{y}||_2 \leq \sigma$$

Where $||\ ||_1$ represent the ℓ_1 norm which is proven that leads to sparse solutions and σ bounds the (non-sparse) noise corrupting the data. The recovered signal $\hat{\mathbf{x}}$ is then computed as $\hat{\mathbf{x}} = \mathbf{\Psi}\hat{\boldsymbol{\alpha}}$.

To investigate the performance of CS compression of biosignals under the presence of noise, we added different levels of zero-mean white gaussian noise to the ECG recordings provided in the MIT-Arrhythmia Database [9], performing their reconstruction varying the compression ratio δ. We defined the input Signal-to-Noise ratios before and after CS as $\text{SNR}_{\text{IN}} = 20 \times \log_{10} \frac{||\mathbf{x}_{sig}||_2}{||\mathbf{n}||_2}$ and $\text{SNR}_{\text{OUT}} = 20 \times \log_{10} \frac{||\hat{\mathbf{x}}||_2}{||\mathbf{x}_{sig} - \hat{\mathbf{x}}||_2}$, respectively.

Results are plotted in Figure 2, with SNRs above 21 dB corresponding to a "good" quality of the reconstructed ECGs [10]. This figure highlights that SNR_{IN} lower than 20 db the quality of results is limited by the input noise for compression ratios above 0.625. Only when SNR_{IN} is above 21 dB further increasing δ leads to tangibly higher-quality outputs. From Figure 2 guidelines can also be derived for the design of the analog-to-digital front-ends performing the signal acquisition, whose precision (and, consequently, power consumption [11]) should be tailored to the performance of the overall system[1].

IV. INEXACT COMPUTING

Inexactness intrinsic in the signal acquisition (noise) and/or in the embedded DSP algorithms (lossy compression) can be accounted for at the microarchitectural level, with the aim of achieving further energy savings. This goal can be obtained by aggressively downscaling the operating voltage of DSP components, at the same time relaxing their computational reliability [12] [13]. While in conventional computing paradigms, the correctness of all operations is required, in inexact computing errors are selectively corrected based on their corresponding impact on the final results.

As a practical example, we considered an application where a WBSN nodes computes the DWT transform of an ECG signal ($\boldsymbol{\alpha} = \mathbf{\Psi}^{-1}\mathbf{x}$). We measured the quality of the output (in terms of percentage root mean-square difference, PRD) when a coefficient calculation is faulty, and hence set to zero. Results, shown in Figure 3, highlight that the correctness in the computation of certain coefficients must be ensured to avoid signal degradation (high PRD), while others are less critical (low PRD). A possible strategy is therefore to provide guaranteed error protection for only parts of the coefficients (the first 25% in Figure 3), employing only best-effort protection to the remaining ones.

[1]The A/D converter power consumption is, on a first approximation, proportional to the square of RMS of the input-referred noise.

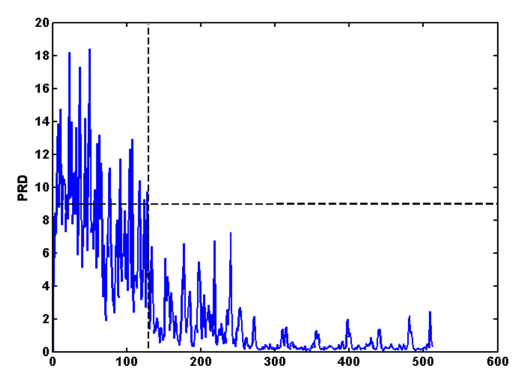
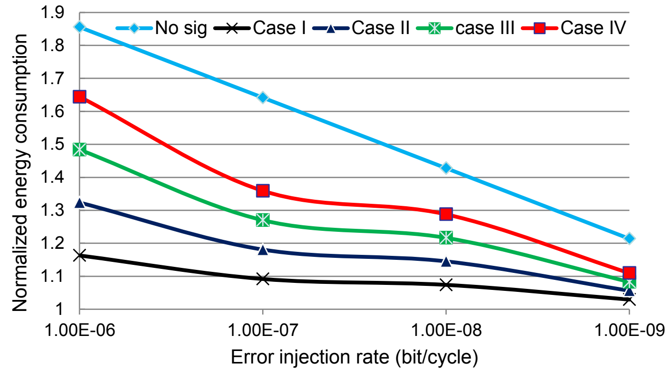


Fig. 3. Quality of reconstructed signals when a single DWT coefficient is incorrect. Higher PRD values are worse, the maximum PRD threshold that indicates the signal is of good quality for further processing is 9% [7].

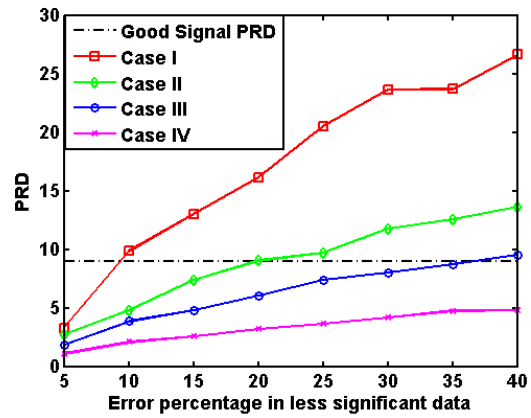
This paradigm (named *Significance-based computing* [4]) is suitable for many applications in bio-signal processing domain, as often biomedical applications exhibit a different level of significance (error induced on the output) for different operations. For instance, during frequency analysis of bio-signals data, some frequencies are of higher importance, while other frequencies instead play a less important role.

To assess the impact of significance-based computing on achieving significant energy savings, we applied this paradigm on a low-power simulated processing platform [14], executing the DWT algorithm. We randomly injected memory errors at varying rates to simulate several inexact working conditions. We consider four cases of significance, namely: *Case I*: 12.5% of the DWT coefficients are significant, *Case II*: 25%, *Case III*, and *Case IV*: 50% of the DWT coefficients are significant (and hence their correctness is guaranteed). For the significant coefficients, we use a HW-assisted 4-bit Error Correcting Code (ECC) protection, while for the least significant we only detect the errors using two-dimensional parity codes and set the faulty word to zero.

Figure 4 shows the normalized energy consumption values, to the case with no protection mechanism, of the mentioned significance cases. Compared to the case with no significance, we can achieve up to 35% energy savings with significance-based computing, particularly *Case I*. However, we need to carefully select the significance case based on the operating condition characterized by the expected error injection rate. In fact, not knowing the expected error rate can lead to either over-protection or significant signal degradation cases, as shown in Figure 4(b). For instance, if the target platform is expected to operate at all the error injection cases, then *Case III* guarantees that the results signal is "good" for further analysis ($PRD < 9\%$ for all error cases [7]), while also guaranteeing 20% energy savings with respect to a case where all data is considered significant and protected by 4-bit ECC.



(a) Energy



(b) Quality

Fig. 4. (a) Normalized energy consumption and (b) reconstructed signal PRD with different significance cases.

References

[1] Y. Hao and R. Foster, "Wireless body sensor networks for health-monitoring applications," *Physiological Measurement*, vol. 29, no. 11, p. R27, Oct. 2008.

[2] R. Braojos, G. Ansaloni, D. Atienza, and F. Rincon, "Embedded real-time ecg delineation methods: A comparative evaluation," in *Bioinformatics Bioengineering (BIBE), 2012 IEEE 12th International Conference on*, Nov. 2012, pp. 99–104.

[3] F. Chen, F. Lim, O. Abari, A. Chandrakasan, and V. Stojanovic, "Energy-aware design of compressed sensing systems for wireless sensors under performance and reliability constraints," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 60, no. 3, pp. 650–661, Mar. 2013.

[4] M. M. Sabry, G. Karakonstantis, D. Atienza, and A. Burg, "Design of energy efficient and dependable health monitoring systems under unreliable nanometer technologies," in *Proceedings of the 7th International Conference on Body Area Networks*, ser. BodyNets '12, 2012, pp. 52–58.

[5] F. Rincon, J. Recas, N. Khaled, and D. Atienza, "Development and evaluation of multilead wavelet-based ECG delineation algorithms for embedded wireless sensor nodes," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, no. 6, pp. 854–863, Nov. 2011.

[6] B. Rajoub, "An efficient coding algorithm for the compression of ecg signals using the wavelet transform," *Biomedical Engineering, IEEE Transactions on*, vol. 49, no. 4, pp. 355–362, Apr. 2002.

[7] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 9, pp. 2456–2466, May 2011.

[8] H. Esmaeilzadeh, A. Sampson, L. Ceze, and D. Burger, "Architecture support for disciplined approximate programming," *SIGARCH Comput. Archit. News*, vol. 40, no. 1, pp. 301–312, Mar. 2012.

[9] PhysioBank. [Online]. Available: www.physionet.org/physiobank/

[10] Y. Zigel, A. Cohen, and A. Katz, "The weighted diagnostic distortion (wdd) measure for ecg signal compression," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 11, pp. 1422–1430, Nov. 2000.

[11] F. Chen, A. Chandrakasan, and V. Stojanovic, "Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors," *Solid-State Circuits, IEEE Journal of*, vol. 47, no. 3, pp. 744–756, Mar. 2012.

[12] A. Lingamneni, K. K. Muntimadugu, C. Enz, R. M. Karp, K. V. Palem, and C. Piguet, "Algorithmic methodologies for ultra-efficient inexact architectures for sustaining technology scaling," in *Proceedings of the 9th Conference on Computing Frontiers*, ser. CF '12, 2012, pp. 3–12.

[13] T. Gemmeke, M. M. Sabry, J. Stuijt, P. Raghavan, F. Catthoor, and D. Atienza, "Resolving the memory bottleneck for single supply near-threshold computing," in *Proceedings of the Design, Automation and Test in Europe Conference and Exhibition*, 2014.

[14] A. Dogan, J. Constantin, M. Ruggiero, A. Burg, and D. Atienza, "Multi-core architecture design for ultra-low-power wearable health monitoring systems," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2012*, Mar. 2012, pp. 988–993.