A case study in low-complexity ECG signal encoding: How compressing is compressed sensing?

A Case Study in Low-Complexity ECG Signal Encoding: How Compressing is Compressed Sensing?

Valerio Cambareri, *Student Member, IEEE*, Mauro Mangia, *Member, IEEE*, Fabio Pareschi, *Member, IEEE*, Riccardo Rovatti, *Fellow, IEEE*, Gianluca Setti, *Fellow, IEEE*

Abstract

When transmission or storage costs are an issue, lossy data compression enters the processing chain of resource-constrained sensor nodes. However, their limited computational power imposes the use of encoding strategies based on a small number of digital computations. In this case study, we propose the use of an embodiment of compressed sensing as a lossy digital signal compression, whose encoding stage only requires a number of fixed-point accumulations that is linear in the dimension of the encoded signal. We support this design with some evidence that for the task of compressing ECG signals, the simplicity of this scheme is well-balanced by its achieved code rates when its performances are compared against those of conventional signal compression techniques.

Index Terms

Compressed Sensing, Lossy Compression, Low Complexity, Wireless Sensor Nodes

I. Introduction

Wireless sensor nodes operate on a tight resource budget, the most limiting constraint being low power consumption in data acquisition, encoding and transmission [1]. Since the power budget of a node is dominated by data transmission, minimising its bit-rate by suitable encoding stages is critical in saving the node's resources. In this context, we assume that data acquisition and compression are performed by low-power sensor nodes that transmit their encoded bitstreams to a central node, which is able to sustain a very large computational burden; such an extreme resource asymmetry limits the use of multimedia *Digital Signal Compression* (DSC) schemes designed on the opposite assumption that the encoding is performed only once (hence as computationally demanding as required),

V. Cambareri and R. Rovatti are with the Department of Electrical, Electronic and Information Engineering (DEI), University of Bologna, Italy (e-mail: valerio.cambareri@unibo.it, riccardo.rovatti@unibo.it). M. Mangia is with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, Italy (e-mail: mmangia@arces.unibo.it). F. Pareschi and G. Setti are with the Engineering Department in Ferrara (ENDIF), University of Ferrara, Italy (e-mail: fabio.pareschi@unife.it, gianluca.setti@unife.it).

whereas decoding is performed multiple times as users access the information content (hence as lightweight as possible).

Compressed Sensing (CS) [2] is a set of mathematical methods that enables the recovery of a Nyquist-rate signal representation from a set of undersampled measurements, as obtained by linear projection of the latter signal. This is made possible by leveraging a sparse signal model [3] that well matches the structure of many signals of interest. In this paper we study a lossy DSC scheme based on CS and targeted at ECG signal compression; its encoding stage projects the signal onto a Bernoulli random matrix stored at the sensor node, whereas the decoding stage entails sparsity-promoting optimisation algorithms that recover an approximation of the encoded signal, a task well-suited to central processing nodes.

Related investigations [4], [5] show that its rate-distortion performances are asymptotically sub-optimal w.r.t. traditional transform coding [6]. Although correct, these analyses do not account for the digital hardware complexity of transform coding, which generally requires floating-point multiplications for an exact transform implementation. Conversely, the encoder of the proposed CS-based DSC scheme is implemented by a lightweight, multiplierless fixed-point architecture. We here illustrate that, for the specific task of compressing ECG signals, the encoder-side complexity of CS is well-matched by its attained code rate. In particular, we show that when (i) a scalar quantiser is used to reduce the measurements' bit-rate, (ii) *Huffman Coding* (HC) is applied on the encoded bitstream, (iii) the random encoding matrix is adapted to the signal ensemble as in [7], [8], the attained code rates and recovery *Signal-to-Noise Ratio* (SNR) performances of a CS-based DSC are optimised. Finally, we compare this approach to some conventional DSC schemes as applied to single-lead ECG signals. The results highlight how the proposed DSC scheme is capable of attaining low code rates with a minimum amount of digital hardware.

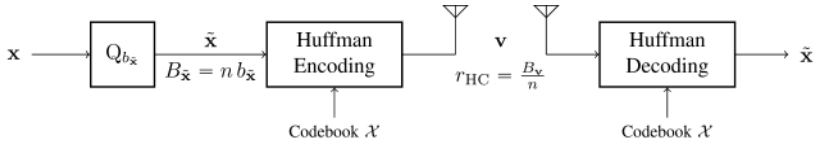
II. Compression Schemes for ECG Signals

We consider the reference case of ECG signals that typically comply with a sparse signal model [9], [10]; a segment of the analog ECG signal is acquired by *Analog-to-Digital* (A/D) conversion, mapping it to n Nyquist-rate samples $\mathbf{x} = \begin{bmatrix} x_0 & \cdots & x_{n-1} \end{bmatrix}^\top \in \mathbb{R}^n$ and finalised[1] by uniform scalar quantisation of each sample as $\tilde{\mathbf{x}} = \mathrm{Q}_{b_{\tilde{\mathbf{x}}}}(\mathbf{x})$, *i.e.*, with $b_{\tilde{\mathbf{x}}}\,\mathrm{bit}$ per sample and scaled to quantise the full signal range. This produces a *Pulse-Code Modulated* (PCM) bitstream of $B_{\tilde{\mathbf{x}}} = n b_{\tilde{\mathbf{x}}}\,\mathrm{bit}$. The task of encoding $\tilde{\mathbf{x}}$ prior to transmission can be divided in (i) a lossy stage that produces a reduced-rate bitstream $\tilde{\mathbf{y}}$ with some information loss w.r.t. $\tilde{\mathbf{x}}$, and (ii) a lossless stage that eliminates its remaining redundancy and outputs the encoded bitstream $\mathbf{v}$, such as an entropy coding scheme [11]. The two stages jointly achieve a code rate of $r = {}^{B_\mathbf{v}}\!/_n\,\mathrm{bps}$ (bits per sample). In particular, we here evaluate three DSC schemes as reported in Figure 1.

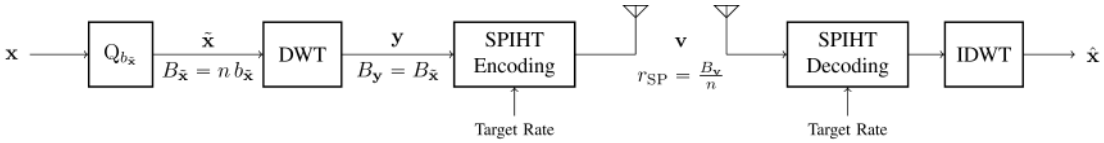
A. Huffman Coding

The lowest-complexity DSC scheme we consider is obtained by processing the PCM samples in $\tilde{\mathbf{x}}$ with standard, lossless HC [11] whose optimal codebook $\mathcal{X}$ is here assumed to be known *a priori* and practically trained on

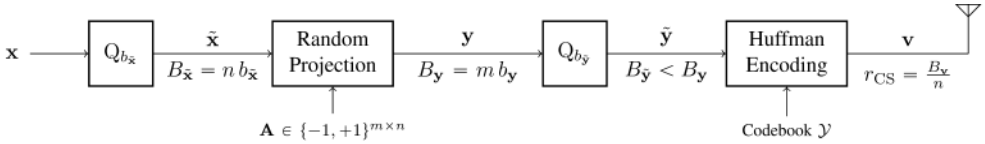
[1]The integration of non-uniform, minimum-distortion quantisers at the A/D converter or early digital processing stages is a technologically complex task; for this reason, we limit this study to uniform scalar quantisers.



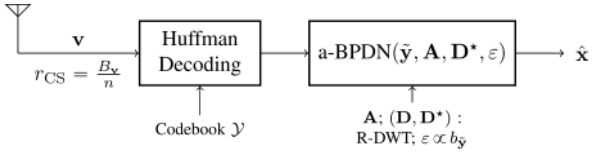
(a) Huffman coding: encoding and decoding stages



(b) Set partition coding: encoding and decoding stages



(c) CS-based DSC: encoding stage



(d) CS-based DSC: decoding stage

Fig. 1. Block diagram of the analysed DSC schemes; $B_{\cdot}$ denotes the lengths of the encoded bitstreams, $b_{\cdot}$ their word length per sample. Channel coding is regarded as part of the transmitter/receiver.

the empirical *Probability Mass Function* (PMF) of a very large set of $b_{\tilde{\mathbf{x}}}$ bit words. Since this training set might not contain all possible words, an *escape* codeword is added to $\mathcal{X}$ followed by $\lceil \log_2 q \rceil$ bit to represent any of the q symbols not appearing in the above set. This DSC is lossless w.r.t. $\tilde{\mathbf{x}}$ and consumes a minimum amount of computational resources: after the signal is quantised, $\tilde{\mathbf{x}}$ is encoded by a lookup table that maps its fixed-length words to variable-length codewords in the encoded bitstream $\mathbf{v}$. Thus, once $\mathcal{X}$ is stored at the sensor node, HC achieves a code rate r_{HC} with no fixed-point arithmetic operations involved, *i.e.*, requiring the computational complexity of $\mathcal{O}(n)$ table lookups.

B. Set Partition Coding of Wavelet Coefficients

The highest-complexity, lossy DSC scheme evaluated here is *Set Partitioning In Hierarchical Trees* (SPIHT) [12]. The SPIHT encoder operates on the *Discrete Wavelet Transform* (DWT) coefficients of $\tilde{\mathbf{x}}$ (in particular the use of a 9/7 biorthogonal DWT [3], [13] is suggested in [12]) by constructing a map of their significance w.r.t. their magnitude and parent-offspring relationships. The critical arithmetic complexity of this encoding stage is in implementing the chosen DWT [14]; its cost is estimated as $\mathcal{O}(n \log_2 n)$ floating-point sums and multiplications (see, *e.g.*, [15]) for a high-precision DWT. While hardware-efficient DWT implementations exist [16], [17], their computation requires fixed-point multiplications with quantised filter coefficients, which cause some precision issues. Thus, we consider

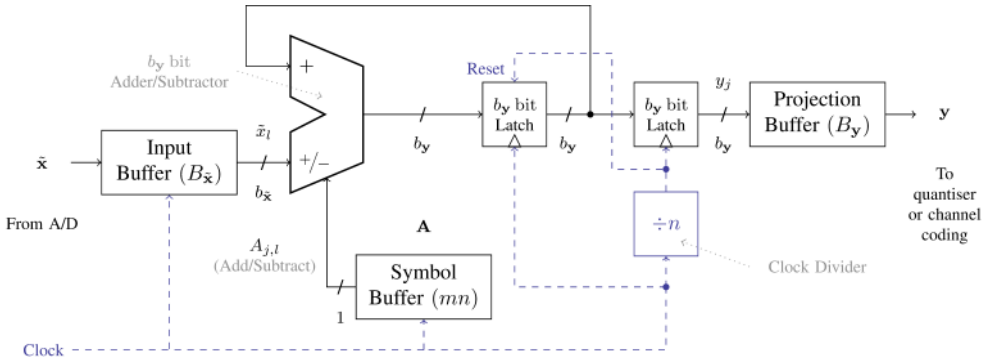


Fig. 2. A digital, multiplierless hardware implementation of the CS encoding stage with Bernoulli random matrices, using a single accumulator and fixed-point arithmetic. The buffers are local registers of size denoted by $(\cdot)$ bit; the input buffer retains $\tilde{\mathbf{x}}$ for mn clock cycles; the dashed lines denote synchronisation signals.

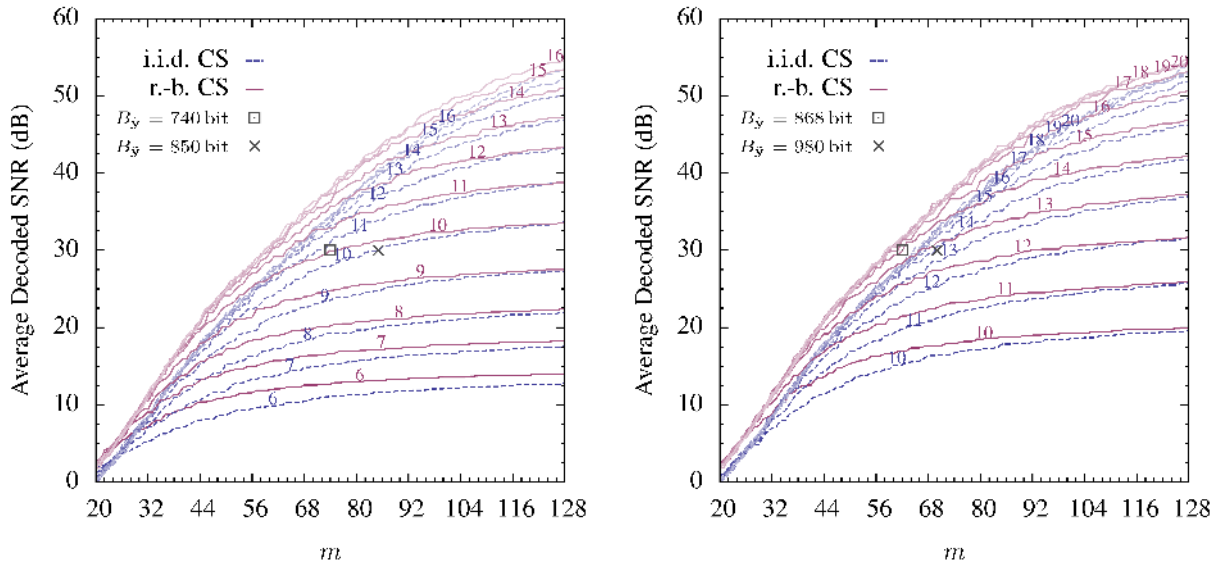
SPIHT to have relatively high-complexity for its integration into a resource-constrained sensor node. As a reference, we will report in Section III the attained, very low code rates r_{SP} of its floating-point implementation followed by an entropy coding stage [12].

C. Lossy Compression by Compressed Sensing

1) Encoding Stage: the CS encoding is carried out as $\mathbf{y} = \mathbf{A}\mathbf{x}$, with $\mathbf{y} = \begin{bmatrix} y_0 & \cdots & y_{m-1} \end{bmatrix}^\top \in \mathbb{R}^m$ the *measurements* and $\mathbf{A}$ a random *encoding matrix* [18] that we assume $\mathbf{A} \in \{-1, +1\}^{m \times n}, m < n$ to minimise its implementation complexity. By applying this encoding on $\tilde{\mathbf{x}}$, we actually collect $\mathbf{y} = \mathbf{A}\tilde{\mathbf{x}}$ represented by m words of[2] $b_{\mathbf{y}} = b_{\tilde{\mathbf{x}}} + \lceil \log_2 n \rceil$ bit. Thus, each y_j is obtained by accumulation of the PCM samples in $\tilde{\mathbf{x}}$ modulated by a sequence of sign changes; the cost of this operation is $\mathcal{O}(mn)$ fixed-point sums. Thus, the CS encoder is conveniently mapped on mn clock cycles of a single accumulator, as in the straightforward digital architecture of Figure 2. To reduce the rate of the encoded bitstream, we further process $\mathbf{y}$ by a second uniform scalar quantiser as $\tilde{\mathbf{y}} = \mathrm{Q}_{b_{\tilde{\mathbf{y}}}}(\mathbf{A}\tilde{\mathbf{x}})$ yielding m words of $b_{\tilde{\mathbf{y}}}$ bit, as obtained by keeping only $b_{\tilde{\mathbf{y}}}$ *most significant bits* (MSBs) from each y_j. After this, we apply lossless HC with an optimal codebook $\mathcal{Y}$ constructed on the empirical PMF of each element of $\tilde{\mathbf{y}}$. Thus, the encoded bitstream $\mathbf{v}$ attains a code rate r_{CS} that depends on $(m, b_{\tilde{\mathbf{x}}}, b_{\tilde{\mathbf{y}}})$, the presence or absence of HC, and a suitable choice of the encoding matrix. These degrees of freedom are numerically compared in Section III.

2) Encoding Matrix Design: although assuming $\mathbf{A} \in \{-1, +1\}^{m \times n}$ as a Bernoulli random encoding matrix with *independent and identically distributed* (i.i.d.) entries and equal-probability symbols is a *universal* choice for any signal ensemble [18], it may be sub-optimal when additional priors on $\mathbf{x}$ are verified besides sparsity. In particular, by letting $\mathbf{x} \in \mathbb{R}^n$ be a *Random Vector* (RV) that models a signal ensemble, we say that it is *localised* if its correlation matrix $\mathbf{C}_{\mathbf{x}}$ is non-white [8]. With this hypothesis, recent contributions [7], [8], [19] have shown how

[2]This choice perfectly quantises the full range of the j-th measurement $y_j \in [-Y, Y]$, where $Y = 2^{b_{\tilde{\mathbf{x}}}-1} n = \max_{\tilde{\mathbf{x}}} \sum_{l=0}^{n-1} \tilde{x}_l$ with $\|\tilde{\mathbf{x}}\|_\infty \leqslant 2^{b_{\tilde{\mathbf{x}}}-1}$.



(a) Measurements $\tilde{\mathbf{y}}$ quantised with $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}}$ bit (values of $b_{\tilde{\mathbf{y}}}$ reported on curves)

(b) Measurements $\tilde{\mathbf{y}}$ quantised with $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}} + \frac{1}{2}\lceil \log_2 n \rceil$ bit (values of $b_{\tilde{\mathbf{y}}}$ reported on curves)

Fig. 3. Average Decoded SNR for i.i.d. (dashed) and rakeness-based (solid) CS with different $\mathrm{Q}_{b_{\tilde{\mathbf{y}}}}$, $b_{\tilde{\mathbf{x}}} = \{6, \ldots, 16\}$, as m varies w.r.t. $n = 256$. For $b_{\tilde{\mathbf{x}}} = 10$, the points corresponding to bit budgets that allow an ADSNR $\approx 30\,\mathrm{dB}$ are highlighted with $\times$ (i.i.d. CS) and $\square$ (rakeness-based CS).

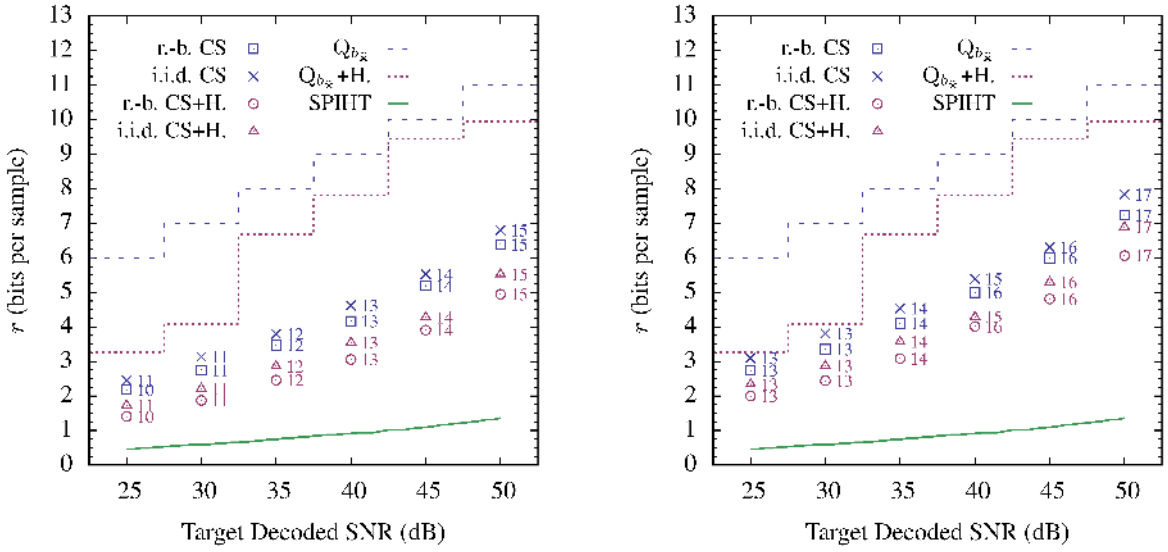
$\mathbf{A}$ can be adapted to the RV $\mathbf{x}$, yielding substantial performance gains; to summarise this so-called *rakeness*-based approach, we let the rows of $\mathbf{A}$ be m independent copies of a RV $\mathbf{a}$ defined on $\{-1, +1\}^n$ whose correlation matrix

$$\mathbf{C_a} = \tau \tfrac{n}{\mathrm{tr}(\mathbf{C_x})} \mathbf{C_x} + (1 - \tau)\mathbf{I}_n,\ \tau \in (0, 1) \tag{1}$$

where $\mathbf{I}_n$ is the n-dimensional identity and τ only needs to be chosen so that $\mathbf{C_a}$ is positive-definite (*e.g.*, $\tau = 1/2$). Moreover, since $\mathbf{a}$ defined on $\{-1, +1\}^n$ imposes[3] $\mathrm{diag}\,(\mathbf{C_a}) = \mathbf{I}_n$ we scale the correlation matrix as $\tilde{\mathbf{C}}_\mathbf{a} = \mathbf{\Gamma} \mathbf{C_a} \mathbf{\Gamma}$ where $\mathbf{\Gamma} = \mathrm{diag}\,(\mathbf{C_a})^{-\frac{1}{2}}$. Then, we synthesise the m rows of $\mathbf{A}$ as in [20, (13)] by taking $\mathbf{A} = \mathrm{sign}\,(\mathbf{T})$, where $\mathbf{T} \in \mathbb{R}^{m \times n}$ collects m instances of a RV $\mathbf{t} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{C_t})$ with $\mathbf{C_t} = \sin(\frac{\pi}{2} \tilde{\mathbf{C}}_\mathbf{a})$. If $\mathbf{C_t}$ is positive-definite (as in most cases with $\tau = 1/2$) then $\mathbf{a}$ with the desired $\tilde{\mathbf{C}}_\mathbf{a}$ can be generated [20] and m of its instances can be stored in the symbol buffer of Figure 2. This encoding matrix design was shown to lower the requirements on m when the correlation matrix $\mathbf{C_x}$ is a stationary property of the RV $\mathbf{x}$. Thus, we use it as an asset to further reduce r_{CS}.

3) Decoding Stage: since $\mathbf{A}$ is a dimensionality reduction, the recovery of $\mathbf{x} \in \mathbb{R}^n$ from its m-dimensional measurements $\tilde{\mathbf{y}}$ hinges on a sparse signal model by which $\mathbf{x}$ has a *k-sparse* representation $\mathbf{s} \in \mathbb{R}^p$ w.r.t. a *synthesis transform* $\mathbf{D} \in \mathbb{R}^{n \times p}$, $p \geqslant n$, *i.e.*, $\mathbf{x} = \mathbf{D}\,\mathbf{s}$, $k = |\mathrm{supp}(\mathbf{s})|$. In fact, recovery error bounds exist [21, Theorem 1.4] when $p > n$, relating k and the minimum number of measurements $m = \mathcal{O}(k \log p/k)$ that ensures the approximate recovery of $\mathbf{x}$ from $\tilde{\mathbf{y}}$ in the presence of noise. Motivated by established results [21], [22] we here choose as a

[3]This can be easily verified to hold for any joint PMF in $\{-1, +1\}^n$. Note that $\mathrm{diag}\,(\cdot)$ here extracts a diagonal matrix from a full matrix.



(a) Measurements $\tilde{\mathbf{y}}$ quantised with $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}}$

(b) Measurements $\tilde{\mathbf{y}}$ quantised with $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}} + \frac{1}{2}\lceil \log_2 n \rceil$

Fig. 4. Achieved code rates of the evaluated DSC schemes and their variants for the chosen ADSNR target specifications; "+H." denotes the use of HC as in Figure 1. For CS-based DSC, the value of $b_{\tilde{\mathbf{y}}}$ that allows a given rate is reported to the right of each marker.

decoding stage the *analysis* form of *Basis Pursuit with De-Noising* (BPDN), *i.e.*,

$$\hat{\mathbf{x}} = \underset{\boldsymbol{\xi} \in \mathbb{R}^n}{\arg\min} \ \|\mathbf{D}^\star \boldsymbol{\xi}\|_1 \quad \text{s.t.} \ \|\tilde{\mathbf{y}} - \mathbf{A}\boldsymbol{\xi}\|_2 \leqslant \varepsilon \tag{a-BPDN}$$

where $\mathbf{D}^\star$ is an *analysis transform* mapping $\boldsymbol{\xi}$ to its transform-domain representation and $\varepsilon \geqslant 0$ controls the data fidelity with which the measurements are matched in the presence of noise sources, in our case limited to $\mathrm{Q}_{b_{\tilde{\mathbf{y}}}}$ (*i.e.*, ε will depend on $b_{\tilde{\mathbf{y}}}$). While this choice of decoding by (a-BPDN) is not as computationally efficient as using greedy algorithms (*e.g.*, [23], [24]), we here adopt it as a reference to provide high-accuracy signal recovery in the presence of quantisation noise and with minimum *a priori* information. We here assume $(\mathbf{D}, \mathbf{D}^\star)$ of a redundant DWT (R-DWT, [3], [25]) that constitutes a tight frame. Since promoting the sparsity of $\mathbf{D}^\star\boldsymbol{\xi}$ while verifying the constraint of (a-BPDN) allows for improved recovery performances in the presence of noise [22] we leverage this property to mitigate the impact of quantisation on the quality of $\hat{\mathbf{x}}$. Specifically, the evaluated decoding stage assumes $(\mathbf{D}, \mathbf{D}^\star)$ of a Symmlet-6 R-DWT with $J = 4$ sub-bands (*i.e.*, $p = (J+1)n$) [3, Chapter 5.2], while the solution of (a-BPDN) is provided by Douglas-Rachford splitting [26] (as implemented in UNLocBox [27]). Finally, we assume a noise norm $\varepsilon = \|\tilde{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2$ and ensure that the solver converges up to a variation of 10^{-7} in the objective function.

III. Performance Evaluation on ECG Signals

In this Section we compare the performances after decoding of the DSC schemes in Figure 1, with an emphasis on evaluating the proposed CS-based DSC and its configurations. In the following we will be concerned with evaluating the average SNR of the decoded signal, $\mathrm{ADSNR} = 10\log_{10}\mathbb{E}\left[\frac{\|\mathbf{x}\|_2^2}{\|\hat{\mathbf{x}}-\mathbf{x}\|_2^2}\right]\mathrm{dB}$ as a performance index, where $\hat{\mathbf{x}}$ is the decoded output of the considered techniques.

A. Signal Generation and PCM Quantisation

We use a synthetic ECG generator [28] to produce 10^4 training instances of $\mathbf{x}$ with $n = 256$ (sampled at $256\,\mathrm{Hz}$). The parameters of the generator are drawn in the same ranges of [7] to obtain a training set oscillating at natural heart rates. Each instance is then quantised to $\tilde{\mathbf{x}}$ by $\mathrm{Q}_{b_{\tilde{\mathbf{x}}}}$. Since the ECG PCM samples' empirical PMF is not uniformly distributed, the intrinsic SNR w.r.t. uniform white quantisation noise can be estimated as $\mathrm{SNR}_{\tilde{\mathbf{x}},\mathbf{x}} = 10\log_{10}\mathbb{E}\left[\frac{\|\mathbf{x}\|_2^2}{\|\tilde{\mathbf{x}}-\mathbf{x}\|_2^2}\right] = 6.02 b_{\tilde{\mathbf{x}}} - 20\log_{10}\frac{\mathrm{CF}_{\mathbf{x}}}{\sqrt{3}} \approx 6.02\, b_{\tilde{\mathbf{x}}} - 11\,\mathrm{dB}$ (as will be reported in Figure 4), obtained by computing the ECG signals' crest factor $\mathrm{CF}_{\mathbf{x}} = \frac{\|\mathbf{x}\|_\infty}{\sqrt{n}\|\mathbf{x}\|_2}$ on the training set.

B. Measurements' Quantisation and Signal-to-Noise Ratio

The main noise sources in Figure 1 are the quantisers $\mathrm{Q}_{b_{\tilde{\mathbf{x}}}}, \mathrm{Q}_{b_{\tilde{\mathbf{y}}}}$. While $\mathrm{Q}_{b_{\tilde{\mathbf{x}}}}$ is common to all evaluated DSC schemes, the latter is only used in the CS-based DSC to reduce each element of $\mathbf{y}$ to $b_{\tilde{\mathbf{y}}} < b_{\mathbf{y}}\,\mathrm{bit}$. Since these elements are approximately Gaussian-distributed, letting $b_{\mathbf{y}} = b_{\tilde{\mathbf{x}}} + \lceil \log_2 n \rceil$ might exceed the precision actually needed to represent $\mathbf{y}$ with negligible distortion. Thus, to explore the effect of $\mathrm{Q}_{b_{\tilde{\mathbf{y}}}}$ we encode the ECG training set by letting $\mathbf{A} \in \{-1,+1\}^{m\times n}$ be a Bernoulli random matrix and evaluate two quantisation policies, *i.e.*, $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}}$ or $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}} + \frac{1}{2}\lceil \log_2 n \rceil$ (this second option serving as a mid-range choice for $b_{\tilde{\mathbf{y}}} \in [b_{\tilde{\mathbf{x}}}, b_{\tilde{\mathbf{x}}} + \lceil \log_2 n \rceil]$), where the range of $\mathrm{Q}_{b_{\tilde{\mathbf{y}}}}$ is rescaled to the extreme values of $\mathbf{y}$. Then we apply the same encoding on 64 test instances, solve (a-BPDN) and compute the ADSNR while varying m up to $n/2$, $b_{\tilde{\mathbf{x}}} = \{6,\ldots,16\}$.

The same procedure is repeated when $\mathbf{A}$ follows a rakeness-based design, with $\mathbf{C_a}$ obtained by plugging the sample correlation matrix $\hat{\mathbf{C}}_{\mathbf{x}}$ of the training set in (1); the range of $\mathrm{Q}_{b_{\tilde{\mathbf{y}}}}$ is rescaled according to the extreme values of $\mathbf{y}$, whose variance is increased due to this design of $\mathbf{A}$.

The results reported in Figure 3 allow us to observe that (i) rakeness-based CS outperforms standard, i.i.d. CS in all the examined cases; (ii) the quality gain obtained by using more bits for both $(b_{\tilde{\mathbf{x}}}, b_{\tilde{\mathbf{y}}})$ progressively saturates at an ADSNR limit imposed by the sparsity of ECG signals w.r.t. the chosen R-DWT; (iii) for a fixed value of $b_{\tilde{\mathbf{x}}}$, the total *bit budget* $B_{\tilde{\mathbf{y}}} = m b_{\tilde{\mathbf{y}}}$ required to reach an ADSNR target indicates the efficiency of the chosen quantisation policy. This quantity is highlighted in both Figures 3a and 3b, and shows how $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}}$ allows for lower code rates; thus, choosing a more accurate quantiser $\mathrm{Q}_{b_{\tilde{\mathbf{y}}}}$ for $\tilde{\mathbf{y}}$ must be matched with a smaller m, whose impact is more critical in achieving high ADSNR levels.

C. Rate Performances for ECG Signal Compression

Given the observed quantisation effects, we now compare the rate performances of two conventional DSC schemes of Figure 1 with different embodiments of CS in search for the lowest attained rate r_{CS} at some fixed target decoding performances, *i.e.*, for $\mathrm{ADSNR} = \{25, 30, 35, 40, 45, 50\}\,\mathrm{dB}$.

For a fair comparison, SPIHT [12] is run from the authors' code by fitting four test instances of $\tilde{\mathbf{x}}$ into frames of 1024 PCM samples quantised at different $b_{\tilde{\mathbf{x}}}$. The SPIHT encoder takes r_{SP} as an input, which we vary in $[0.05, 2]$; the minimum r_{SP} that guarantees the target ADSNR after decoding is then reported in Figure 4. As a further reference, we report the rates attained by the scheme of Figure 1a, *i.e.*, by uniform PCM quantisation, achieving a rate r_{HC} with optimal HC; since it is lossless, the achievement of an ADSNR target w.r.t. $\mathbf{x}$ actually depends on

$b_{\tilde{\mathbf{x}}}$. While r_{HC} could be estimated by the entropy of PCM samples, to account for the presence of escape symbols we run HC on the test set to find the true r_{HC}.

These two reference methods are compared with various CS configurations in Figure 4, which reports the cases in Figure 3 that match the desired ADSNR with minimum r_{CS}. There, we observe that the rates attained in Figure 4a are generally lower than those in Figure 4b, thus confirming the advantage of assuming $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}}$. In addition, (i) the use of HC on the measurements significantly reduces the code rate of CS; (ii) by considering r_{CS} of rakeness-based CS with HC, Figure 4a shows that an $\mathrm{ADSNR} \approx 25\,\mathrm{dB}$ is achieved at $b_{\tilde{\mathbf{y}}} = b_{\tilde{\mathbf{x}}} = 10\,\mathrm{bit}$ by $r_{\mathrm{CS}} \approx 1.41\,\mathrm{bps}$, while $r_{\mathrm{HC}} = 3.27\,\mathrm{bps}$. Moreover, while outperformed by floating-point SPIHT, under low ADSNR requirements the CS encoder in Figure 2 is a viable alternative to provide DSC with a critical digital hardware simplification that should be matched with sensor node constraints.

IV. Conclusion

As a lossy DSC, CS was shown to be capable of achieving low code rates with extremely low computational complexity at the encoder; these rates were optimised by some additional considerations on the CS encoder. Given these low digital hardware requirements, CS lends itself as an agile scheme for DSC tasks under tight resource constraints.

References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[2] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[3] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Access Online via Elsevier, 2008.

[4] V. K. Goyal, A. K. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 48–56, 2008.

[5] A. Fletcher, S. Rangan, and V. Goyal, "On the rate-distortion performance of compressed sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. ICASSP 2007*, vol. 3, Apr. 2007, pp. III–885–III–888.

[6] V. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, Sep. 2001.

[7] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 59, no. 5, pp. 1001–1014, 2012.

[8] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A rakeness-based design flow for analog-to-information conversion by compressive sensing," in *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 1360–1363.

[9] Z. Zhang, T.-P. Jung, S. Makeig, and B. Rao, "Compressed Sensing for Energy-Efficient Wireless Telemonitoring of Noninvasive Fetal ECG Via Block Sparse Bayesian Learning," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 300–309, Feb. 2013.

[10] L. Polania, R. Carrillo, M. Blanco-Velasco, and K. Barner, "Compressed sensing based method for ECG compression," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 761–764.

[11] W. A. Pearlman and A. Said, *Digital Signal Compression: principles and practice*. Cambridge University Press, 2011.

[12] Z. Lu, D. Y. Kim, and W. A. Pearlman, "Wavelet compression of ecg signals by the set partitioning in hierarchical trees algorithm," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 7, pp. 849–856, 2000.

[13] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 45, no. 5, pp. 485–560, 1992. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/cpa.3160450502/abstract

[14] T. Fry and S. Hauck, " image compression on FPGAs," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 9, pp. 1138–1147, Sep. 2005.

[15] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, Jul 1992.

[16] S. Rein and M. Reisslein, "Low-Memory Wavelet Transforms for Wireless Sensor Networks: A Tutorial," *IEEE Communications Surveys Tutorials*, vol. 13, no. 2, pp. 291–307, 2011.

[17] K. Kotteri, S. Barua, A. Bell, and J. Carletta, "A comparison of hardware implementations of the biorthogonal 9/7 DWT: convolution versus lifting," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 52, no. 5, pp. 256–260, 2005.

[18] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5406–5425, 2006.

[19] A. Caprara, F. Furini, A. Lodi, M. Mangia, R. Rovatti, and G. Setti, "Generation of Antipodal Random Vectors With Prescribed Non-Stationary 2-nd Order Statistics," *IEEE Transactions on Signal Processing*, vol. 62, no. 6, pp. 1603–1612, Mar. 2014.

[20] G. Jacovitti, A. Neri, and G. Scarano, "Texture synthesis-by-analysis with hard-limited Gaussian processes," *IEEE Transactions on Image Processing*, vol. 7, no. 11, pp. 1615–1621, Nov. 1998.

[21] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

[22] I. W. Selesnick and M. A. Figueiredo, "Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors," in *SPIE Optical Engineering + Applications*. International Society for Optics and Photonics, 2009, pp. 74 460D–74 460D.

[23] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[24] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[25] J. Fowler, "The Redundant Discrete Wavelet Transform and Additive Noise," *IEEE Signal Processing Letters*, vol. 12, no. 9, pp. 629–632, Sep. 2005.

[26] P. L. Combettes and J.-C. Pesquet, "A douglas–rachford splitting approach to nonsmooth convex variational signal recovery," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 564–574, 2007.

[27] N. Perraudin, D. Shuman, G. Puy, and P. Vandergheynst, "UNLocBoX: A matlab convex optimization toolbox using proximal splitting methods," *arXiv preprint arXiv:1402.0779*, 2014.

[28] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 3, pp. 289–294, 2003.