

# An Architecture for Low-power Compressed Sensing and Estimation in Wireless Sensor Nodes

David Bellasi

Riccardo Rovatti

Luca Benini

Gianluca Setti

ETH Zürich, Switzerland DEI/ARCES - University of Bologna ETH Zürich, Switzerland ENDIF - University of Ferrara  
 Email: bellasid@iis.ee.ethz.ch Email: riccardo.rovatti@unibo.it Email: benini@iis.ee.ethz.ch Email: gianluca.setti@unife.it

**Abstract**—Radio communication is among the most energy consuming tasks in wireless sensor nodes. Reducing the amount of data to be transmitted holds a large power saving potential. The combination of compressed sensing (CS) and local signal parameter estimation can achieve a massive data rate reduction in applications where the primary interest is in the acquisition of a scalar feature of the signal rather than the reconstruction of the entire waveform. In this paper, We propose a compressed estimator, building upon an enhancement of the typical CS signal-modulation scheme via *punctured* sampling. Specifically, a subset of signal samples and associated weighting coefficients are chosen so as to minimize node power consumption while achieving a given estimation performance. We detail a corresponding *puncturing* algorithm and present the design of an integrated digital compressed estimation unit in 28 nm FDSOI CMOS. In a concrete case study, local estimation combined with subsampling is shown to result in a power reduction of up to an order of magnitude with respect to the standard solution of sampling and transmitting samples for off-board processing.

## I. INTRODUCTION

Compressed Sensing (CS) is a recently introduced signal acquisition paradigm applicable to signals whose intrinsic dimensionality is considerably smaller than what the Nyquist theorem suggests. In the discrete-time formulation the waveform to be acquired in a given time window is represented by a set of  $n$  samples collected in a vector  $x = (x_0, \dots, x_{n-1})^T$ . The key assumption of CS is *sparsity*, i.e., the existence of an  $n$ -dimensional basis  $D$  such that  $x = D\xi$  for every sample vector  $x$ , with  $\xi$  being a vector in which at most  $k \ll n$  components are non-zero. In this case, the actual number of degrees of freedom of the signal is considerably smaller than  $n$ , and its salient information content can be captured in a set of  $m < n$  properly designed linear, non-adaptive measurements. If the  $j$ -th measurement is given by the linear combination

$$y_j = \sum_{k=0}^{n-1} a_{j,k} x_k, \quad (1)$$

using suitable coefficients  $a_{j,k}$ , we obtain the measurement vector  $y = Ax = ADs$ , where we arranged the coefficients  $a_{j,k}$  in the  $m \times n$  matrix  $A$ . If  $A$  and  $D$  have certain mathematical properties [1] one can recover  $s$  (and thus  $x$ ) from  $y$  despite the fact that  $A$  (and thus  $AD$ ) incurs a dimensionality reduction. Recovery is guaranteed even in the presence of a certain amount of noise and by means of algorithms whose development has been a hot research field in recent years [1], [2], [3].

Thanks to the above properties, CS is a promising method to extend the lifetime of wireless sensor systems relying on *nodes* with limited energy resources. Specifically, the reduction of the data to be transmitted can result in massive power

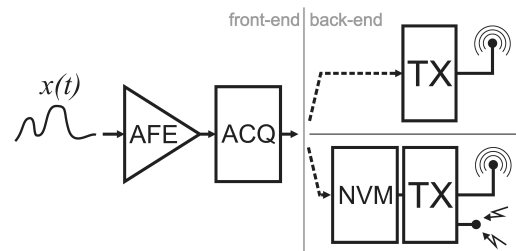


Fig. 1. Main system-level components of a wireless sensor node.

savings as the energy required for transmission is by far the largest contribution to the power budget of such a node [6].

A simplified overview of a typical sensor node is depicted in Fig. 1: the system architecture is schematically divided into a front-end that performs amplification (AFE) and signal acquisition including (ACQ) analog-to-digital conversion, and a back-end. The back-end either transmits (TX) the digital words from the front-end immediately or, in an alternative implementation, stores them in non-volatile memory (NVM), where the data resides until enough energy becomes available to trigger transmission (either by energy harvesting or because of intentional external power supply).

A CS-based acquisition stage must implement (1), which can be done either in the analog-domain prior to A/D conversion or in the digital-domain after digitization. However, it was suggested that early digitization may be more energy-efficient in CS-based nodes [6]. While a CS front-end alone achieves a data reduction in the order of  $m/n$ , the combination of CS-acquisition and parameter estimation can result in an even larger reduction of the power consumption. This applies to the case in which we are interested in quantifying a single signal parameter rather than in the knowledge of the exact waveform. We therefore propose a novel early-digitizing CS front-end based on an enhanced multiplier-accumulator multiply-and-accumulate unit C-MAC that, in addition to the reconstruction-oriented acquisition, enables parameter estimation in the classical linear setting [7]. The C-MAC can be reconfigured to implement an effective estimator whose result is the only measurement transmitted by the node. As  $m = 1$  in this case, the consumption of the back-end in Fig. 1 is reduced by a factor of  $m$  accordingly. Additionally, we optimize the estimation such that it requires a minimum number of samples, resulting in *punctured* sampling which allows to switch-off one or more circuit parts when no data is processed.

In the following, we first describe the architecture and derive a corresponding power model with reference to a monolithic implementation in a state-of-the-art 28 nm FDSOI CMOS technology in Section II. In Section III, we recall the

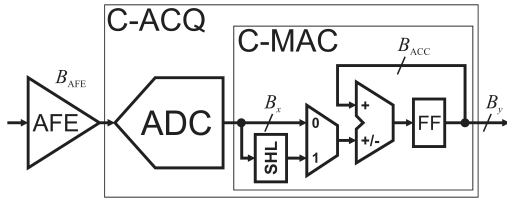


Fig. 2. Proposed compressed sensing front-end C-ACQ based on early-digitization and a specialized MAC unit C-MAC.

classical linear setting for parameter estimation and derive the *puncturing* algorithm that yields the optimal sampling positions for a fixed number of samples. This allows to set up a design loop in which, given a specific estimation performance target, one is able to identify the sampling strategy that minimizes the associated power consumption. This design flow is exemplified in a case study in Section IV, where we show the enhanced energy-efficiency of our proposed architecture compared to the conventional node implementation. Finally, Section V concludes the paper.

## II. A MIXED PROJECTION-ESTIMATION ARCHITECTURE

The stable reconstruction of  $x$  from  $y$  by CS algorithms is guaranteed with high probability if  $A$  consists of symbols randomly chosen from a limited set of different values [4]. Typically, the hardware-friendly coefficients  $a_{j,k} \in \{-1, +1\}$  are chosen in practical CS-based signal acquisition systems [5]. Yet, our aim is to reduce the number of samples used in the computation, which can be achieved by considering an extended set of coefficients. Since we have early digitization, samples are fed into the projection block already as binary words, hence, their multiplication by 2 is equivalent to a shift operation and can be realized almost for free using a simple multiplexer.

We propose the specialized acquisition stage C-ACQ, shown in Fig. 2. The analog signal, after amplification by the AFE stage, is sampled and quantized by ADC in  $B_x$  bits. The projection in (1) is performed by an enhanced MAC unit C-MAC that is able to apply the coefficients  $\pm 1$  and  $\pm 2$  explicitly, and 0 implicitly (simply by not taking any sample), while featuring virtually the same hardware complexity as a standard MAC. The accumulation is performed by the C-MAC unit using an adder/subtractor with a wordwidth of  $B_{ACC}$  bits, of which  $B_y$  bits are delivered to the back-end for immediate or delayed transmission.

### A. Signal model

A multi-channel CS acquisition system may include multiple parallel C-ACQ units. Focusing on a single path of C-ACQ (and thus dropping the index  $j$  in (1)), the system computes the linear combination  $y = \sum_{k=0}^{n-1} a_k x_k$  with the coefficients  $a_k \in \{-2, -1, 0, +1, +2\}$  indicating whether the  $k$ -th sample is discarded ( $a_k = 0$ ), or acquired and added or subtracted from the overall sum ( $|a_j| = 1$  or  $|a_j| = 2$ ). The number of non-zero coefficients  $s$  is the number of samples actually processed. By defining the projection vector  $a = (a_0, \dots, a_{n-1})^T$  we have  $y = a^T x$ .

### B. Power model

Here, we detail simple power models for each of the circuit blocks in Fig. 2. The power model for the analog

front-end is adapted from [6] assuming a gain of 40 dB, a noise efficiency  $NEF=3$ , and an input bandwidth of  $f_{ADC}^{\max}/2$  where  $f_{ADC}^{\max}$  is the maximum sampling rate of the subsequent ADC<sup>1</sup>. Assuming that the total disturbance at the output allows an equivalent resolution of  $B_{AFE}$  bits (i.e. that the noise variance is  $\propto 2^{-2B_{AFE}}$ ), the power consumed by the AFE is  $2^{2B_{AFE}} \times 36 \text{ fJ} \times f_{ADC}^{\max}$ . Yet, whenever the ADC is not acquiring any sample (i.e.,  $a_{j,k} = 0$ ) the AFE can be deactivated and cut-off from the supply by a power-gating transistor. This ensures that the power consumption scales with the sampling frequency below the idle power limit [8]. Hence, the above AFE power can be scaled first by the fraction  $f_s/f_{ADC}^{\max}$  of the sample time  $1/f_s$  required for digitization and then by the fraction of samples  $s/n$  actually acquired in the estimation time window. This yields

$$\mathcal{P}_{AFE} = \frac{s}{n} \times 2^{2B_{AFE}} \times 36 \text{ fJ} \times f_s,$$

The power-gating scheme, likewise applied to the ADC, effectively removes virtually any static component from the power budget<sup>2</sup>. Hence, the power consumption of the ADC can be modeled via the commonly accepted Walden figure-of-merit (FOM) [9]. Assuming a  $FOM=22.4 \text{ fJ/conversion-step}$  [8] we have

$$\mathcal{P}_{ADC} = \frac{s}{n} \times 2^{B_x} \times 22.4 \text{ fJ} \times f_s.$$

The consumption of C-MAC was modeled by designing dedicated units with different input  $B_x$  and output  $B_y$  bit numbers in a low-leakage 28 nm FDSOI CMOS technology and by performing Monte-Carlo simulations<sup>3</sup>. By fitting simulation data we derived the following linear model

$$\mathcal{P}_{C-MAC} = \frac{s}{n} \times (B_x + B_{ACC} - 1) \times 1 \text{ pJ} \times f_s + \mathcal{P}_{C-MAC,0}.$$

Here,  $\mathcal{P}_{C-MAC,0}$  is the leakage power of the C-MAC, unavoidable due to the fact that the idle accumulator cannot be power-gated without losing its state.  $\mathcal{P}_{C-MAC,0}$  is proportional to the circuit area and was also derived by fitting synthesis data as

$$\mathcal{P}_{C-MAC,0} = (3B_x + 4B_{ACC} - 5) \text{ nW}.$$

The transmitter's power consumption is obtained from the energy per transmitted data bit and the total number of transmitted bits. Depending on the distance covered by the wireless link either narrow band or ultra-wideband (UWB) systems can be used. They differ considerably in transmission energy, with values in the range of a few nJ/bit for narrow band links [10] and tens of pJ/bit for UWB links [12]. Assuming a narrowband TX we have

$$\mathcal{P}_{TX} = \frac{m}{n} \times B_y \times 3 \text{ nJ} \times f_s,$$

where  $B_y$  is the number of bits used to encode each measurement. The absence of any static contribution is due to the fact that TX can be power-gated.

As far as NVM is concerned, data storage comes at a per-bit energy cost on the order of 1 pJ/bit [11]. As NVMs retain their state even if the power supply is removed, no leakage power is considered. Hence, we have

$$\mathcal{P}_{NVM} = \frac{m}{n} \times B_y \times 1 \text{ pJ} \times f_s.$$

<sup>1</sup>Further, the analog supply voltage  $V_{dd,A}=1 \text{ V}$  and the absolute temperature  $\Theta=300 \text{ K}$  have been assumed.

<sup>2</sup>The static leakage current of the power-gate itself is neglected in our model.

<sup>3</sup>Synopsys Design Compiler and Mentor Graphics Modelsim were used for logic synthesis and netlist simulation.

---

**Algorithm 1** Puncturing algorithm
 

---

```

1:  $J \leftarrow \text{permute } \{0, \dots, n-1\} \text{ s.t. } |z_{J(j)}| \geq |z_{J(j+1)}|$ 
   for  $j = 0, \dots, n-2$ 
2: Initialize  $a \leftarrow (0, \dots, 0)^\top$  and  $\bar{\omega} \leftarrow \infty$ 
3: for  $j = 0, \dots, s-1$  do
4:    $\bar{A} \leftarrow 0$ 
5:   for  $A = 2, \dots, 1$  do
6:      $\omega \leftarrow \frac{a^\top a + A^2}{(a^\top z + A|z_{J(j)}|)^2}$ 
7:     if  $\omega < \bar{\omega}$  then
8:        $\bar{\omega} \leftarrow \omega$  and  $\bar{A} \leftarrow \text{sign}(z_{J(j)})A$ 
9:     end if
10:  end for
11:   $a_{J(j)} \leftarrow \bar{A}$ 
12: end for
13:  $a \leftarrow \frac{a}{\min\{|a_j| \text{ s.t. } a_j \neq 0, j = 0, \dots, n-1\}}$ 

```

---

### III. ESTIMATION

The proposed architecture enables both compressed acquisition as well as compressed estimation. While CS acquisition requires multiple projection paths, estimation only requires a single one. As it aims at the extraction of a single parameter instead of waveform reconstruction, estimation is a much simpler task.

We adopt the classical linear model [7], in which  $x(t) = \alpha z(t) + \eta(t)$  with  $z(t)$  being a known waveform represented by the samples in  $z = (z_0, \dots, z_{n-1})^\top$ ,  $\eta(t)$  being a realization of a disturbance process producing the samples in  $\eta = (\eta_0, \dots, \eta_{n-1})^\top$ , and  $\alpha$  an unknown parameter modulating the amplitude of  $z(t)$ . We assume that  $\eta$  is made of independent samples with zero mean and variance  $\sigma_\eta^2$ . Starting from an instance of  $x$  we want to compute an estimate  $\hat{\alpha}$  of  $\alpha$ . Given our architecture it is most natural to set  $\hat{\alpha} = y/a^\top z$  and look for the coefficients  $a$  that maximize the estimation performance. The estimator is unbiased since  $\mathbf{E}[\hat{\alpha}] = \frac{\alpha a^\top z}{a^\top z} + \frac{a^\top \mathbf{E}[\eta]}{a^\top z} = \alpha$  and the disturbance is assumed to be zero mean. The estimator variance is given by

$$\sigma_{\hat{\alpha}}^2 = \mathbf{E}[(\hat{\alpha} - \alpha)^2] = \sigma_\eta^2 \frac{a^\top a}{(a^\top z)^2}, \quad (2)$$

which implies that drawing  $a$  at random, as in the case of full acquisition, is not a good choice. In fact, in the unfavorable case of  $a$  and  $z$  being almost orthogonal, their scalar product  $a^\top z$  tends to vanish thereby increasing the variance without bound.

Since  $z$  is known, we may choose  $a$  so that  $\sigma_{\hat{\alpha}}^2$  is minimized. Without any constraint on  $a$ , the best set of coefficients would be  $a = z$  so that  $\sigma_{\hat{\alpha}}^2 = \sigma_\eta^2 / z^\top z$ . In our case,  $a \in \{-2, -1, 0, +1, +2\}^n$  and, in general, we want to proceed with  $s < n$  samples, so that  $n - s$  of the entries of  $a$  must be zero.

To devise an effective solution of the resulting discrete optimization problem, note that the balance between the numerator and the denominator of (2) is biased in favor of the denominator when a large-modulus component of  $z$  is considered (i.e., multiplied by a non-zero coefficient). Hence, it is sensible to sort the components of  $z$  in non-increasing order of their modulus and associate the first  $p$  of them to the coefficients in  $a$  minimizing  $\sigma_{\hat{\alpha}}^2$ . The resulting *puncturing* algorithm is summarized in Alg. 1 and selects the  $s$  best samples to use out of the  $n$  uniform ones along with the optimal corresponding coefficients from the set  $\{-2, -1, +1, +2\}$ . The final scaling step at line 13 ensures that the minimum non-zero value in  $a$

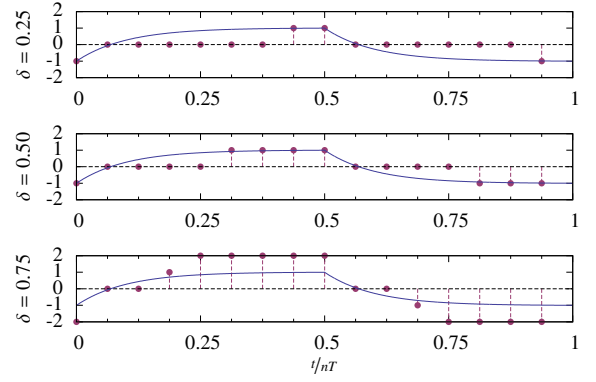


Fig. 4. Optimal sampling strategies for estimating the amplitude of an RC pulse when (from top to bottom) 25%, 50%, and 75% of the available samples are used.

is always 1, which minimizes the dynamic range requirement of the accumulator

### IV. CASE STUDY AND PERFORMANCE EVALUATION

We next study the specific case where  $z(t)$  is a sequence of RC-type charge and discharge transients of known time constant  $\tau$  but unknown amplitude  $\alpha$ . If  $z(t)$  is normalized in  $t \in [0, 1]$  we have

$$z(t) = -1 + 2 \begin{cases} \frac{1 - e^{-t/\tau}}{1 - e^{-1/2\tau}} & \text{if } t \leq 1/2 \\ 1 - \frac{1 - e^{-(t-1/2)/\tau}}{1 - e^{-1/2\tau}} & \text{if } t > 1/2 \end{cases}$$

We assume  $n = 16$  and  $s \in \{4, 8, 12\}$  and define the sampling density  $\delta = s/n$ , which evaluates to  $\delta \in \{0.25, 0.5, 0.75\}$  for the given  $n$  and  $s$ . For each of these densities, the procedure in Alg. 1 is applied and yields the sampling positions that maximize the rejection of the measurement noise, as well as the corresponding optimal coefficients  $a$  (excluding zero). The results are illustrated in Fig. 4 and show that the optimal sampling positions tend to be concentrated where the magnitude of the basic waveform is the largest. The availability of four instead of only two coefficient values is exploited to improve performance when the density increases and values are sampled that are significantly smaller than the maximum.

We are ultimately interested in obtaining amplitude estimations with an accuracy equivalent to  $B_\alpha$  bits, and observe a double RC transient with the total duration of  $T = 0.1$  s. For this, we set the noise performance of the AFE to  $B_{\text{AFE}} < B_\alpha$  equivalent bits and define the resolution of the ADC as  $B_x = B_\alpha$  bits at a maximum conversion rate of  $f_{\text{ADC}}^{\text{max}}$ . Our requirement translates into  $\sigma_\alpha^2 / \sigma_\eta^2 \leq 2^{2(B_{\text{AFE}} - B_\alpha)}$  so that  $\Delta B = B_\alpha - B_{\text{AFE}}$  assumes the significance of a logarithmic processing gain. We sweep  $f_{\text{ADC}}^{\text{max}}$  exponentially from 100 Hz (for  $\Delta B = 1$ ) to 1.1 MHz (for  $\Delta B = 8$ ).

This sets the scene for the evaluation of six different sensor node typologies we perform next. They are obtained by combining the available sampling, processing and transmission options from Fig. 1, i.e., on-board or off-board estimation (**ON\_** or **OFF\_**), full sampling taking  $n$  samples or sub-sampling acquiring  $s$  out of  $n$  samples (**\_FS** or **\_SS**), and immediate transmission or storage in the NVM (**\_TX** or **\_NV**). In general, we consider a sampling rate  $f_s$  producing  $n = f_s T$  potential samples, of which  $s \leq n$  are used

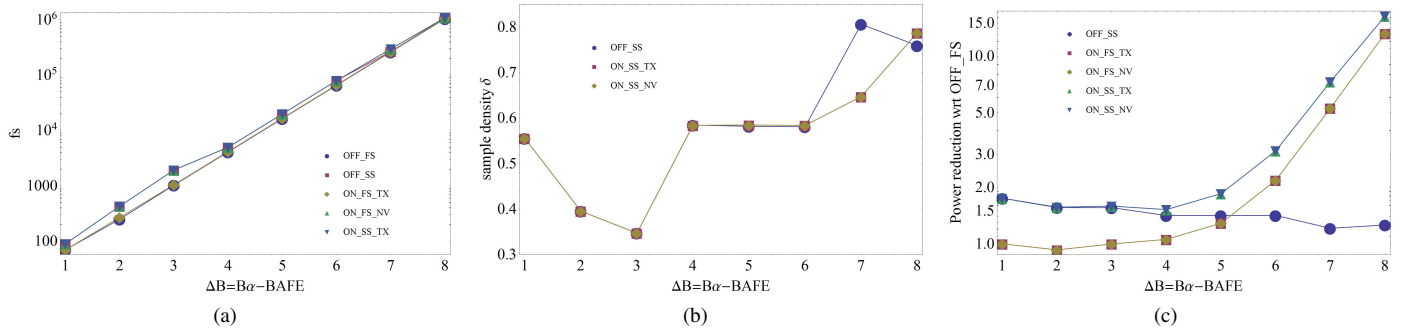


Fig. 3. (a) Sampling rates needed by each option to meet estimation requirements and minimize power; (b) The sample density  $\delta$  minimizing power consumption for subsampling options; (c) Power reduction with respect to **OFF\_FS**. On-board, subsampling options reduce power consumption by an order of magnitude.

for estimation. The (**OFF\_**) options entail the immediate transmission of all the acquired samples (thus  $m = s$  and  $B_y = B_\alpha$ ), and result in the total power consumption of  $\mathcal{P}_{AFE} + \mathcal{P}_{ADC} + \mathcal{P}_{TX}$ . Estimation is performed off-board with non-quantized coefficients and using standard optimal linear estimation techniques [7] that straightforwardly extends to subsampling. The on-board options (**ON\_**) rely on local estimation (thus  $m = 1$  and  $B_y = B_\alpha$ ) with quantized coefficients and selected samples, both optimally chosen using our puncturing algorithm (Alg. 1). The power consumption of (**ON\_**) amounts to either  $\mathcal{P}_{AFE} + \mathcal{P}_{ADC} + \mathcal{P}_{C-MAC} + \mathcal{P}_{TX}$  or  $\mathcal{P}_{AFE} + \mathcal{P}_{ADC} + \mathcal{P}_{C-MAC} + \mathcal{P}_{NVM}$ , depending on whether the estimates are immediately transmitted (**\_TX**) or stored in the NVM (**\_NV**). In the **\_NV** option the transmit power disappears from the power budget as no data is transmitted while the node is in the self-sustained operation mode. In all cases, the degrees of freedom ( $n$  and possibly  $s$ ) were set to yield the needed processing gain  $\Delta B$  and, when multiple configurations are possible, the one with minimum power was chosen.

The results of evaluating all six options are reported in Figs. 3(a)-(c) as a function of the gain  $\Delta B$ . In Fig. 3(a) we plot the actual  $f_s$  needed by each configuration to meet the requirements and minimize power. We observe that the sampling rates increase with the same exponential trend for all the six options. The evaluation of the required  $\delta$ , shown in Fig. 3(b), only applies to options with  $\delta < 1$ , i.e., subsampling options. The results suggest that higher densities must be employed for higher gains  $\Delta B$ , thus, by comparison with Fig. 4, triggering the use of all coefficient values that our architecture allows. Finally, Fig. 3(c) highlights the reduction in power consumption achieved by the various typologies with respect to **OFF\_FS**. It is clear that, although partially effective, simple subsampling (**OFF\_SS**) has a much smaller impact than local estimation (**ON\_FS\_TX** and **ON\_FS\_NV**). As anticipated, pairing local estimation with subsampling (**ON\_SS\_TX** and **ON\_SS\_NV**) yields the largest power reduction up to one order of magnitude for large process gains.

## V. CONCLUSION

We have presented a compressed estimation hardware architecture for wireless sensor nodes, which leverages the combined data reduction effect of subsampling and parameter estimation to achieve low-power operation. The introduction of *punctured* sampling effectively allows us to minimize power consumption of the node while achieving a specified detection performance. A corresponding specialized multiplier-accumulator unit was designed in a 28 nm FDSOI CMOS technology and characterized in terms of power consumption. In a case

study we found that the advantage of the compressed sensing and estimation approach over other sensor node architectures can be as high as an order of magnitude in reduction of the total power consumption.

## VI. ACKNOWLEDGMENTS

The research was partially funded by the project *UltrasoundToGo*, financed with a grant from the Swiss Nano-Tera.ch initiative, and evaluated by the Swiss National Science Foundation.

## REFERENCES

- [1] E.J. Candes, T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005
- [2] D. Needell, 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
- [3] P. Maechler, C. Studer, D.E. Bellasi, A. Maleki, A. Burg, N. Felber, H. Kaeslin, R.G. Baraniuk, "VLSI Design of Approximate Message Passing for Signal Restoration and Compressive Sensing," *IEEE J. Emerging Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 579–590, Sept. 2012
- [4] E.J. Candes, T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006
- [5] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, "A Pragmatic Look at Some Compressive Sensing Architectures With Saturation and Quantization," *IEEE J. Emerging Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 443–459, Sept. 2012
- [6] F. Chen, A.P. Chandrakasan, V.M. Sojanovic, "Design and analysis of Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *IEEE J. Solid State Circuits*, vol. 47, no. 3, pp. 744–756, 2012
- [7] S.M. Kay, *Fundamental of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, 1993
- [8] M. Yip, A.P. Chandrakasan, "A resolution-reconfigurable 5-to-10b 0.4-to-1V power scalable SAR ADC," in *Dig. Tech. Papers IEEE Int. Solid-State Circuits Conf.*, pp. 190–192, Feb. 2011
- [9] R.H. Walden, "Analog-to-Digital Converter Survey and Analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, 1999
- [10] J.L. Bohorquez, J.L. Dawson, A.P. Chandrakasan, "A 350uW CMOS MSK transmitter and 400 uW OOK super-regenerative receiver for medical implant communications," in *Dig. Tech. Papers IEEE Sym. VLSI Circuits*, pp. 32–33, 2008
- [11] N. Gilbert, Z. Yanqing, J. Dinh, B. Calhoun, "A 0.6V 8 pJ/write non-volatile CBRAM macro embedded in a body sensor node for ultra low energy applications," in *Dig. Tech. Papers IEEE Sym. VLSI Circuits*, pp. 16–2, 2013
- [12] Y. Park, D.D. Wentzloff, "An All-Digital 12 pJ/Pulse IR-UWB Transmitter Synthesized From a Standard Cell Library," *IEEE J. Solid-State Circuits*, vol. 46, no. 5, pp. 1147–1157, May. 2011