Projected-gradient-descent in rakeness-based compressed sensing with disturbance rejection

Projected-Gradient-Descent in Rakeness-based Compressed Sensing with Disturbance Rejection

Mauro Mangia*, Letizia Magenta†, Alex Marchioni*,†, Fabio Pareschi§,*, Riccardo Rovatti*,†, Gianluca Setti‡,*
*ARCES - University of Bologna - via Toffano 2/2 - Bologna - ITALY
†DEI - University of Bologna - viale Risorgimeto 2 - Bologna - ITALY
‡DET - Politecnico di Torino - corso Duca degli Abruzzi 24 - Torino - ITALY
§DE - University of Ferrara - via Saragat 1 - Ferrara - ITALY

***Abstract*—Compressed Sensing (CS) has recently emerged as an effective tool to simultaneously acquire and compress analog waveforms in low-resource sensing devices. Its mechanisms have been also extended by both adapting the sensing stage to the actual class of input signals, and granting it the ability to reject disturbances. Regrettably, the resulting design flow entails the solution of two optimization problems with a potentially huge number of variables. This work overcomes this impasse by proposing a Project-Gradient-Descend method algorithm that drastically reduces the required CPU time to obtain a solution.**

I. Introduction

In many modern applications, sensors are the natural interface between data processing and the physical world. The produced data comes from analog waveforms that are filtered, converted in digital words and eventually compressed/encrypted before they are wirelessly dispatched. Furthermore, in many distributed sending paradigms, sensors must be as *parsimonious* as possible, i.e., signal acquisition, digitalization and any other signal manipulation before transmission must be done at an almost negligible cost in order to preserve battery lifetime.

In this scenario, Compressed Sensing (CS) [1] is as an innovative solution capable to merge data compression and low-power requirements. Compression can be performed directly in the analog domain [2]–[4]; alternatively, CS is to be considered as a low-power digital compression algorithm [5]–[7]. In addition, it also provides some level of privacy [8] without paying the cost of implementing a typical encryption system [10].

An encoder block based on CS take chunks of input signals and projects them over a limited set of predefined *sensing sequences*. The corresponding output, the measurements, represent the compressed data to be transmitted to the decoder stage that is able to recover the original waveform. By arranging the sensing sequences in a matrix, one highlights the main CS feature: data compression is performed by a simple matrix multiplication. This simple approach is still able to guarantee compression ratios that are close to values obtained with other more computational hungry algorithms [8].

The classical CS theory has been extend by [11] with the introduction of the *rakeness* concept that adapts the sensing sequences to the actual class of input signal[1] This rakeness-based approach has been further expanded in [14] considering also the impact of disturbances affecting the acquired signal, i.e., the sensing sequences statistics depend also on the disturbances features. A block scheme of this approach is reported in Figure 1, where x and $\hat{x}$ are the input and the reconstructed signal while ν represents disturbances. In [14], disturbance rejection is achieved by solving an optimization problem with a number of variables that is quadratic with respect to the signal dimensionality. The authors used a general purpose solver that works only if the signal dimensionality is limited. Yet, several CS applications involve highly dimensional signals and a method able to solve the same optimization problem for hundreds or thousands of variables is required. Here we propose a tailored algorithm that uses projected gradient descent and alternating projection.

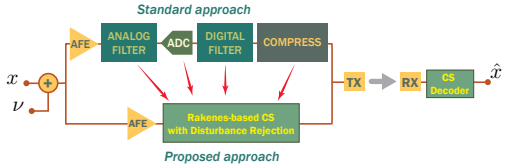

Fig. 1. CS system block diagram including the effect of noise.

The paper is organized as follows. Sec. II briefly introduces CS theory and its extension with rakeness and disturbance rejection. Sec. III describes the proposed algorithm in detail and compares its computation times with a general purpose solver, and in Sec. IV the effectiveness of the approach is shown with a toy case. Finally, we draw the conclusion.

II. Compressed Sensing, Rakeness and Disturbance Rejection

A. Standard CS Approach

We consider a discrete-time input modeled as a vector $x \in \mathbb{R}^n$ of signal samples composed of Nyquist rate samples in a certain window, while $\nu \in \mathbb{R}^n$ represents an additive disturbance. CS bases its effectiveness on the κ-*sparsity* prior: a basis $D \in \mathbb{R}^{n\times n}$ exists such that expressing x along D yields $x = D\xi$ for a vector of coefficients $\xi \in \mathbb{R}^n$ with at most $\kappa \ll n$ non-zero elements. Under this assumption, CS processes $x + \nu$ by projecting it onto $m < n$ sequences arranged as rows of a sensing matrix $A \in \mathbb{R}^{m\times n}$ to obtain a measurement vector $y \in \mathbb{R}^m$

$$y = A(x + \nu) = Ax + A\nu = AD\xi + A\nu \qquad (1)$$

Starting from y, a decoder may reconstruct $\hat{x} = D\hat{\xi}$ as the best approximation of x according to y and to the priors. The recovery procedure exploits the sparsity prior to select, among the infinite number of n-dimensional vectors ξ compatible with

[1]One exploits the statistical features of input signals to increase signal reconstruction quality by suitably design matched acquisition sequence. This is similar to what happens in (chaos-based) DS-CDMA communication, where chip waveforms, spreading sequence statistics and rake receivers taps are jointly selected to collect (rake) as much energy as possible at the received side [12] [13]. Readers interested in the methodology for sequence generation may see [15].

the values in y, the sparsest one $\hat{\xi}$. Mathematically, $\hat{\xi}$ is the solution of the convex optimization problem

$$\hat{\xi} = \arg\min_{\xi} \|\xi\|_1 \;\text{ s.t. } \|AD\xi - y\|_2^2 \leq \epsilon^2 \tag{2}$$

where $\|\cdot\|_p$ is the ℓ_p-norm, and ϵ accounts for the effect of disturbances. Classical CS theory (S-CS) [1] guarantees that the reconstruction error $\|\hat{x} - x\|_2$ is almost always vanishing if, for example, the elements of A are instances of independent and identically distributed (i.i.d.) Gaussian random variables and $m = O(\kappa \log(n/\kappa))$.

B. The Rakeness-based CS Approach

The classical approach has been extended in [7], [11] with the rakeness CS (R-CS) approach. Let $C^x = \mathbf{E}[xx^\top]$ be the correlation matrix of the input signal, and $\cdot^\top$ indicate the transpose operator. We define the *localization* of x as

$$\mathfrak{L}_x = \frac{\mathrm{tr}(C^{x2})}{\mathrm{tr}(C^x)^2} - \frac{1}{n} \tag{3}$$

where $\mathrm{tr}(\cdot)$ is the matrix trace. Mathematically, $\mathfrak{L}_x$ is an indication of the deviation of C^x from the $n \times n$ identity matrix I_n characterizing a white signal, for which $\mathfrak{L}_x = 0$.

Under the additional (typically verified [11]) assumption that the class of input signal is localized, CS performance can be improved by designing A according to the R-CS. To sketch it, let $a^\top \in \mathbb{R}^n$ be a generic row of A. The idea behind R-CS is to increase the energy of the signal-related part Ax of $y = A(x + \nu)$, i.e. $\rho = \mathbf{E}\left[\left(a^\top x\right)^2\right] = \mathrm{tr}\left(C^a C^x\right)$. This is done by drawing a as a random vector with a correlation matrix $C^a = \mathbf{E}[aa^\top]$ that depends on C^x and thus on the $\mathfrak{L}_x$. Mathematically, one has to solve

$$\rho^* = \max_{C^a} \;\mathrm{tr}\left(C^a C^x\right) \tag{4a}$$

$$\text{s.t. } {C^a}^\top = C^a,\; C^a \succeq 0 \tag{4b}$$

$$\text{s.t. } \mathrm{tr}(C^a) = 1,\; \mathfrak{L}_a \leq \tau \mathfrak{L}_x \tag{4c}$$

where (4b) ensure that C^a is a correlation matrix (i.e., it is symmetric and positive semidefinite), and the first in (4c) normalizes the average energy of a. The role of $0 \leq \tau \leq 1$ in the second of (4c) is to limit the localization $\mathfrak{L}_a$ of a to a fraction of $\mathfrak{L}_x$ thus avoiding overadaptation, a common setting is $\tau = 1/4$. Problem (4) has been solved analytically [7], [11]:

$$\rho^* = \mathrm{tr}(C^x)\left(\frac{1}{2}\mathfrak{L}_x + \frac{1}{n}\right) \tag{5a}$$

$$C^a = \frac{1}{2}\left(\frac{C^x}{\mathrm{tr}(C^x)} + \frac{I_n}{n}\right) \tag{5b}$$

where I_n is the $n \times n$ identity matrix. In order to improve performance, one can simply generate the a according to the correlation matrix in (5b) [15].

C. Rakeness-based CS With Disturbance Rejection

If $C^\nu = \mathbf{E}[\nu\nu^\top]$, it has been shown in [14] that reducing the average energy $\mathbf{E}\left[\left(a^\top \nu\right)^2\right] = \mathrm{tr}\left(C^a C^\nu\right)$ then the disturbance-related part $A\nu$ in $y = A(x + \nu)$ results in an attenuation of the noise term in the reconstructed signal $\hat{x}$. In other terms, we have a *disturbance rejection* effect without the need of any additional filter.

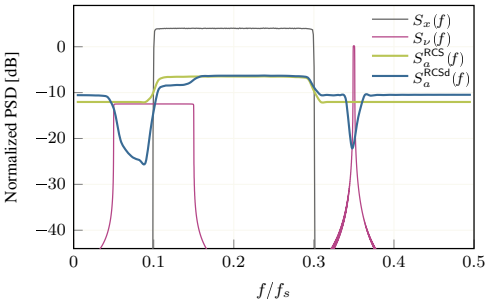

Fig. 2. With $n = 64$, normalized Power Spectral Densities of the considered class of input signals ($S_x(f)$) and disturbances ($S_\nu(f)$) along with profiles for both R-CS ($S_a^{\mathsf{RCS}}(f)$) and R-CSd with $r = 0.95$ ($S_a^{\mathsf{RCSd}}(f)$).

Being able to increase $\mathrm{tr}\left(C^a C^x\right)$ and, at the same time, to decrease $\mathrm{tr}\left(C^a C^\nu\right)$ will help the reconstruction stage to recover x (as for R-CS) and, at the same time, will reduce the contribute of the disturbance. This results in a trade-off as it is not possible to maximize $\mathrm{tr}\left(C^a C^x\right)$ and at the same time minimize $\mathrm{tr}\left(C^a C^\nu\right)$. The authors of [14] propose an approach R-CSd based on the following minimization problem.

$$\min_{C^a} \;\mathrm{tr}\left(C^a C^\nu\right) \tag{6a}$$

$$\text{s.t. } (4b)\,(4c)$$

$$\text{s.t. } \mathrm{tr}(C^a C^x) \geq r\rho^* \tag{6b}$$

where $0 \leq r \leq 1$ is a parameter whose value is typically close to 1 and sets how much we are allowed to depart from the solution of the rakeness problem (4).

The ratio behind (6) is to minimize the energy collected from the disturbance under the constraints ensuring that C^a is a proper correlation matrix, and providing that the energy collected from the signal is not too small with respect to what could be collected when the disturbance was not considered.

Figure 2 shows normalized power spectral densities for a toy case with $n = 64$. Here a band-pass signal is perturbed by a disturbance composed by two contribution, one passband with a partial overlap with the input signal, and a second narrowband signal. Figure reports also spectra for the R-CS problem as well as for the R-CSd problem. The $S_a^{\mathsf{RCSd}}(f)$ profile is high in the same band of the signal, as for $S_a^{\mathsf{RCS}}(f)$, while it rapidly decreases when the $S_\nu(f)$ is high.

III. Projected-Gradient-Descent and Alternating Projections Approach

The optimization problem in (6) is well-formed as it entails a linear objective function and convex constraints. Hence, it may be solved by means of standard convex solvers like CVX [20]. Yet, general purpose solvers tend to suffer from the increase of problem dimensionality that, in our case, is $O(n^2)$ as we have to determine the entries of an $n \times n$ correlation matrix. Regrettably, quite often the application of CS to signal acquisition entails signal dimensionalities in the hundreds (or thousands) and the resulting problem sizes may be completely out of the reach of non-specialized solvers. This is why we develop here a tailored iterative procedure for solving (6) in

which the complexity of each iteration is $O(n^2)$ and can be easily executed on standard workstations.

The first step is reformulating (6) in terms of convex subsets of the space of $n \times n$ matrices. Let $\mathbf{C}_{\mathrm{PSD}}$ be the subset of all symmetric and positive semidefinite $n \times n$ matrices. With this, constraints (4b) can be simply written as $C^a \in \mathbf{C}_{\mathrm{PSD}}$. Let also be $\mathbf{C}_{\mathrm{En}}$ the class of matrix satisfying the energy constraint (first in (4c)), $\mathbf{C}_{\tau-\mathrm{Loc}}$ satisfying the second constraint in (4c) and in (4b), and $\mathbf{C}_{r-\mathrm{Rak}}$ satisfying (6b) and the second constraint in (4b).

With this, the R-CSd problem in (6) translates into

$$\begin{aligned} \min_{C^a} \ & \mathrm{tr}\,(C^a C^\nu) \\ \text{s.t.} \ & C^a \in \mathbf{C}_{\mathrm{PSD}} \cap \mathbf{C}_{\mathrm{En}} \cap \mathbf{C}_{\tau-\mathrm{Loc}} \cap \mathbf{C}_{r-\mathrm{Rak}} \end{aligned} \tag{7}$$

that, observing that the gradient $\nabla_{C^a} \mathrm{tr}\,(C^a C^\nu) = C^\nu$ is a constant matrix, is a classical case of a bounded-gradient objective function subject to convex constraint. Such a problem can be effectively tackled by a *projected-gradient-descent* approach [16]. In order to apply this method, let $\Pi_{\mathbf{C}}$ be the projection operator on a set of matrices $\mathbf{C}$, i.e., the operator that takes any $n \times n$ matrix X and maps it into the matrix $Y = \Pi_{\mathbf{C}}(X)$ such that $Y \in \mathbf{C}$ and $\|X - Y\|$ is minimum.

Then, starting from an initial matrix $C^a_{(0)}$ that satisfies the constraints in (7), one can compute iteratively

$$C^a_{(t+1)} = \Pi_{\mathbf{C}_{\mathrm{PSD}} \cap \mathbf{C}_{\mathrm{En}} \cap \mathbf{C}_{\tau-\mathrm{Loc}} \cap \mathbf{C}_{r-\mathrm{Rak}}} \left(C^a_{(t)} + \delta_{(t)} C^\nu \right) \tag{8}$$

for a suitably non-increasing sequence $\delta_{(0)} \geq \delta_{(1)} \geq \dots$. It can be proved that the sequence of $C^a_{(t)}$ converges to the optimal solution as $t \to \infty$. A good choice for $C^a_{(0)}$ is the C^a from (5b), i.e., the solution of the R-CS problem.

The iteration in (8) needs to compute the projection on a convex set that is the intersection of a number of elementary convex sets. Such a projection operator can be computed by means of the alternating projection method [17], [18]. To explain how it works, let us pretend that we have only two convex sets $\mathbf{C}_0$ and $\mathbf{C}_1$ and that we want to compute $\Pi_{\mathbf{C}_0 \cap \mathbf{C}_1}(C)$ for some matrix C. The projection is computed iteratively producing a sequence of matrices $P_{(j)}$ and using two auxiliary sequences of offset matrices $\Delta P_{0,(j)}$ and $\Delta P_{1,(j)}$. Initialization sets $P_{(0)} = C$ and $\Delta P_{0,(0)} = \Delta P_{1,(0)} = 0$. Then, the s-th step of the procedure computes

$$\begin{aligned} P_{(2s+1)} &= \Pi_{\mathbf{C}_0} \left(P_{(2s)} - \Delta P_{(2s)} \right) \\ \Delta P_{0,(s+1)} &= P_{(2s+1)} - P_{(2s)} - \Delta P_{0,(s)} \\ P_{(2s+2)} &= \Pi_{\mathbf{C}_1} \left(P_{(2s+1)} - \Delta P_{1,(s)} \right) \\ \Delta P_{1,(s+1)} &= P_{(2s+2)} - P_{(2s+1)} - \Delta P_{1,(s)} \end{aligned}$$

As $s \to \infty$ one has $P_{(2s+1)} \to P_{(2s)} \to \Pi_{\mathbf{C}_0 \cap \mathbf{C}_1}(C)$. The scheme straightforwardly generalizes to more than two elementary convex sets in the intersection, and reveals that the overall projection in (8) can be computed starting from the individual projection operators on $\mathbf{C}_{\mathrm{PSD}}$, $\mathbf{C}_{\mathrm{En}}$, $\mathbf{C}_{\tau-\mathrm{Loc}}$ and $\mathbf{C}_{r-\mathrm{Rak}}$.

Assuming that C can be spectrally decomposed as $C = E \Lambda E^\top$, where E is the matrix of orthonormal eigenvectors and Λ is the diagonal matrix of the corresponding eigenvalues, then

TABLE I
COMPUTATIONAL TIME RATIO, $\mathrm{CPU}_{\mathrm{ratio}}$, BETWEEN GENERAL PURPOSE SOLVER AND THE PROJECTED-GRADIENT-DESCENT ALGORITHM

n	32	64	96	128
$\mathrm{CPU}_{\mathrm{ratio}}$	3.3	5.9	5.8	15.9

$$\Pi_{\mathbf{C}_{\mathrm{PSD}}}(C) = E \max\{\Lambda, 0\} E^\top$$

The energy constraint (4c) and the minimum rakeness constraint (6b) define two linear subspaces and the corresponding projection operators are simple affine transformations following the direction of the gradient of the constraint

$$\Pi_{\mathbf{C}_{\mathrm{En}}}(C) = C + \frac{1 - \mathrm{tr}(C)}{n} I_n$$

$$\Pi_{\mathbf{C}_{r-\mathrm{Rak}}}(C) = C + \begin{cases} \frac{r\rho^* - \mathrm{tr}(CC^x)}{\mathrm{tr}(C^{x\,2})} C^x & \text{if } \mathrm{tr}(CC^x) < r\rho^* \\ 0 & \text{otherwise} \end{cases}$$

Assuming $\mathrm{tr}(C) = 1$, the limit on localization (second constraint in (4c)) is a homogeneous quadratic constraint and projection can be obtained by scaling

$$\Pi_{\mathbf{C}_{\tau-\mathrm{Loc}}}(C) = C \begin{cases} \sqrt{\frac{\tau \mathfrak{L}_x + 1/n}{\mathrm{tr}(C^2)}} & \text{if } \mathrm{tr}(C^2) > \tau \mathfrak{L}_x + 1/n \\ 1 & \text{otherwise} \end{cases}$$

To assess the computational resources needed by this approach, Table I reports the CPU time ratio between the general purpose CVX solver and the proposed method for n values ranging from 32 to 128. As expected, the tailored approach amply outperforms a general purpose solver, where for $n = 128$ the CPU time of the proposed approach is 5.31 seconds[2]. In addition to this, for values of n larger than 128 the CVX solver is not able to solve the problem.

IV. SETTINGS AND NUMERICAL RESULTS

To assess the effectiveness of the proposed method we consider time windows composed by $n = 512$ successive Nyquist samples, i.e., a case for which the CVX solver fails.

Spectral profiles for input signals and disturbances are the same used in the previous case while power spectra for both R-CS and R-CSd are reported in Figure 3-(a). Remarkably, the comparison of $S_a^{\mathsf{RCS}}(f)$ and $S_a^{\mathsf{RCSd}}(f)$ for $n = 64$ (Figure 2) with the ones in Figure 3-(a), where $n = 512$, shows how an increase in the signal dimensionality allows more effective spectral profiles.

As before, we use band pass input signals while disturbances spectral profile is composed by a bandpass contribution, concentrated in a region with a partial overlap with the input signal PSD, along with a narrowband disturbance. This means that both, signals and disturbances posses localization. In particular, each x is also κ-sparse with respect to the Fourier basis with $\kappa = 51$, while each ν ia an instance of a zero-mean Gaussian random vector. The mentioned spectral profiles determine the correlation matrices C^x and C^ν. These matrices

[2]All CPU times were evaluated on a standard PC with a Intel i7-6820HQ processor and 16 GB RAM

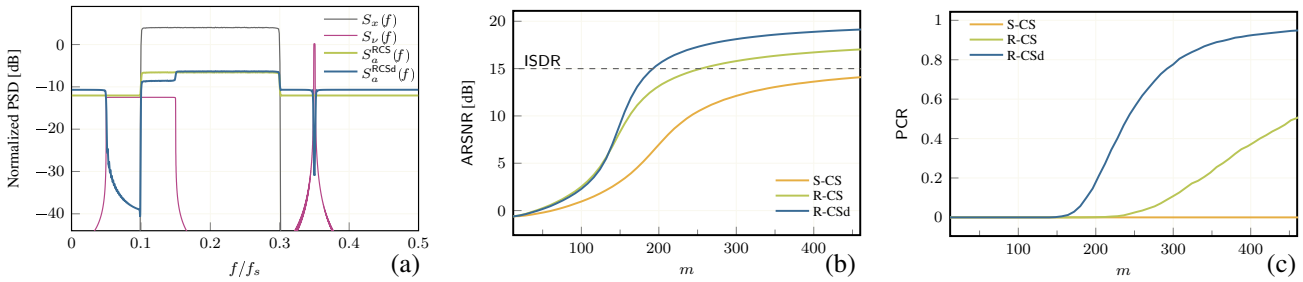

Fig. 3. Plot (a) reports the Normalized PSDs of x, ν and ones for sensing matrices rows according to R-CS and R-CSd with $r = 0.95$, while plots (b) and (c) depict ARSNR and PCR as functions of the number of measurement m, for the three considered CS approaches and with $\mathsf{ISDR} = 15\,\mathrm{dB}$.

are used in (5b) and in the algorithm we propose here for the solution of (6).

Signals are reconstructed by solving (2) via the SPGL1 toolbox [19] where the reconstruction error for each trial is assessed by the Reconstructed Signal to Noise Ratio (RSNR)

$$\mathsf{RSNR} = \frac{\|x\|_2}{\|x - \hat{x}\|_2}$$

For the overall system performance two different figures of merit were considered. The Average Reconstructed Signal to Noise Ratio (ARSNR), evaluated as empirical mean of the observed values of RSNR, along with the Probability of Correct Reconstruction (PCR), i.e., the probability that the observed RSNR exceeds a proper threshold $\mathsf{RSNR}_{\mathsf{min}}$. Both are estimated over 3000 Montecarlo trials using different signal and disturbances instances and where sensing matrices were generated in accordances with R-CS and R-CSd with $r = 0.95$. S-CS was also included and uses sensing matrices generated as instances of independent zero-mean unit-variance Gaussian random variables. Considered m values range from few units to more than 400.

Note that the effect of ν in the measurements computation in (1) is counted by the intrinsic signal-to-disturbance ratio (ISDR) defined as

$$\mathsf{ISDR} = \left(\frac{\|x\|_2}{\|\nu\|_2}\right)_{\mathrm{dB}}$$

where we set $\mathsf{ISDR} = 15\,\mathrm{dB}$ while $\mathsf{RSNR}_{\mathsf{min}} = 17\,\mathrm{dB}$.

Figure 3 shows system performance in terms of ARSNR (b) and PCR (c). For both figures of merit R-CSd clearly outperforms both R-CS and S-CS. These results confirm the ability of the R-CSd to partially reject disturbance affecting the input signal without any additional filtering stage. Remarkably, the signal dimensionality in this example is such that a traditional general purpose solver is not able to solve the optimization problem (6), while the algorithm discussed in Section III converge to a solution in a reasonable amount of time.

V. Conclusion

A Project-Gradient-Descent and alternating projections method to solve the rakeness minimizzation problem for disturbance rejection is presented. As expected, it drastically outperforms a general purpose solver in terms of CPU time and, more important, it is able to converge to the minimum point also for reasonably large values of n.

References

[1] D.L.Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 54, n. 4, pp. 1289–1306, 2006

[2] D. Gangopadhyay, et al. "Compressed Sensing Analog Front-End for Bio-Sensor Applications," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 2, pp. 426–438, 2014.

[3] F. Pareschi, et al. "Hardware-Algorithms Co-design and Implementation of an Analog-to-Information Converter for Biosignals based on Compressed Sensing," *IEEE Trans. on Biomed. Circuits and System*, vol. 10, no. 1, pp. 149-162, 2016

[4] J. Haboba, et al. "A Pragmatic Look at Some Compressive Sensing Architectures With Saturation and Quantization," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 443-459, 2012.

[5] D. Bortolotti, et al. "An Ultra-Low Power Dual-mode ECG Monitor for Healthcare and Wellness," *DATE 2015*, pp. 1611-1616, 2015

[6] D. E. Bellasi, R. Rovatti, L. Benini and G. Setti, "A Low-Power Architecture for Punctured Compressed Sensing and Estimation in Wireless Sensor-Nodes," in *IEEE Trans. on Circuits and Systems I*, vol. 62, no. 5, pp. 1296-1305, 2015.

[7] M. Mangia, at al. "Rakeness-based design of low-complexity compressed sensing," *IEEE Transactions on Circuits and Systems I*, vol. 64, no. 5, pp. 1201–1213, 2017.

[8] V. Cambareri, at al. "A Case Study in Low-Complexity ECG Signal Encoding: How Compressing is Compressed Sensing?," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1743-1747, 2015

[9] V. Cambareri, at al. "On Known-Plaintext Attacks to a Compressed Sensing-Based Encryption: A Quantitative Analysis," in *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2182-2195, 2015.

[10] L. Y. Zhang et al., "On the Security of a Class of Diffusion Mechanisms for Image Encryption," *IEEE Transactions on Cybernetics* vol. 48, no. 4, pp. 1163-1175, 2018

[11] 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*, vol. 59, no. 5, pp. 1001-1014, 2012

[12] R. Rovatti, G. Mazzini, and G. Setti, "Enhanced rake receivers for chaos-based ds-cdma," *IEEE Transactions on Circuits and Systems I*, vol. 48, pp. 818–829, 2001

[13] G. Setti, R. Rovatti, and G. Mazzini, "Performance of chaos-based asynchronous ds-cdma with different pulse shapes," *IEEE Communications Letters*, vol. 8, pp. 416–418, July 2004

[14] A. Marchioni, at al. "Disturbance Rejection With Rakeness-based Compressed Sensing: Method and Application to Baseline/Powerline," *2018 IEEE ISCAS2018*, May 2018

[15] R. Rovatti, G. Mazzini, and G. Setti,"Memory- antipodal processes: Spectral analysis and synthesis," *IEEE Transactions on Circuits and Systems I*, vol. 56, no. 1, pp. 156–167, 2009

[16] A.A. Goldstein, "Convex Programming in Hilbert Space," *Bullettin of the American Mathematical Society*, vol. 70, pp. 709–710, 1964

[17] R.L. Dykstra, "An Algorithm for Restricted Least Squares Regression," *Journal of the American Statistical Association*, vol. 78, n. 384, pp. 837–842, 1983

[18] J.P. Boyle, R.L. Dykstra, "A method for finding projections onto the intersection of convex sets in Hilbert spaces," *Lecture Notes in Statistics*, vol. 37, pp. 28—47, 1986

[19] E. Van den Berg, M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890-912, 2008

[20] M. Grant, S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control, Springer-Verlag Limited*, pp. 95-110,2008