Geometric Constraints in Sensing Matrix Design for Compressed Sensing

Geometric Constraints in Sensing Matrix Design for Compressed Sensing

C.H. Pimentel-Romero[a], M. Mangia[a,b], F. Pareschi[c,d], R. Rovatti[b,d], G. Setti[c,d]

[a] *Department of Engineering - University of Ferrara, 44124 Ferrara, Italy*
[b] *Electric, Electronic and Information Engineering Department - University of Bologna, 40136 Bologna, Italy*
[c] *Department of Electronic and Telecommunication - Politecnico di Torino, 10129 Torino, Italy*
[d] *ARCES Research Center - University of Bologna, 40125 Bologna, Italy*

Abstract

Compressed Sensing (CS) has been proposed as a method able to reduce the amount of data needed to represent sparse signals. Nowadays, different approaches have been proposed in order to increase the performance of this technique in each stage that composes it. Particularly, this paper provides a critical review of the state-of-the art of some CS adaptations in the sensing stage to identify the strengths and limitations of each of them. In addition, a new method is proposed (Nearly Orthogonal Rakeness-based CS) that aims to overcome limits of the CS adaptations covered in this work. After intensive numerical simulations on synthetic signals and electroencephalographic (EEG) signals, the proposed approach outperforms discussed state-of-the-art approaches in terms of compression capability required to achieve a target quality of service.

Email addresses: pmncrh@unife.it (C.H. Pimentel-Romero), mauro.mangia2@unibo.it (M. Mangia), fabio.pareschi@polito.it (F. Pareschi), riccardo.rovatti@unibo.it (R. Rovatti), gianluca.setti@polito.it (G. Setti)

Keywords: Compressed Sensing, rakeness, Electroencephalographic signals (EEG).

1. Introduction

In recent years, the trade-off between the increase of the quantity of data to manage and the electronic devices miniaturization has been a big challenge. The amount of data we currently store and share daily is growing impressively. To cope with this problem, new processing techniques have been developed along with the improvement of hardware devices. In the field of data compression, one of these new approaches is the Compressed Sensing (CS) [1; 2], that is a framework capable to simultaneously acquire and compress an input signal. Basically, CS exploits a very common feature of the signals called *sparsity*. This property is the mathematical quantification of the idea that a signal, when expressed in a proper basis, allows a much more compact representation than what one can obtain using a straightforward Nyquist-rate sampling approach. Due to this, the interest in CS has recently grown, and has led to the development of Analog-to-Information Converters (AICs) [3; 4; 5], that base the acquisition of an analog signal on the CS framework, with a potential reduction of energy consumption in comparison with the standard Analogic-to-Digital Converters (ADCs). The effectiveness of AICs has already been investigated in different areas, ranging from bio-medical prototypes [5; 6; 7] to AICs for the acquisition of radio-frequency signals [8; 9].

Mentioned prototypes are essentially CS-based encoder blocks. Compressed information is dispatched to a decoder stage, that is able to recover

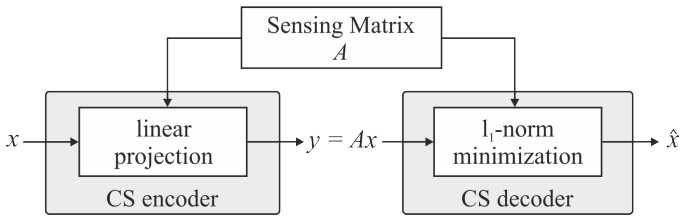

Figure 1: Std-CS based system block diagram.

the original signal exploiting the assumption that the input signal is sparse. The encoder/decoder processing scheme that follows the CS framework is highly asymmetric. The encoder compresses the input signal, once it has been segmented into contiguous and non overlapped chunks, by a linear projections on the rows of a properly designed sensing matrix A (usually drawn as instances of stochastic process) with a very low computational cost. This cost can be also further reduced with some constraints on the A. As an example, by asking that each element of A is an antipodal entry, i.e, whose value can be only $+1$ or -1, the linear projection is such that only signed sums are required to compute it. Since no performance degradation has been observed with this approach [10; 11], we will refer in this paper to approaches capable to generate an antipodal matrix A only.

Conversely, a much higher cost is paid by the decoder, where a convex optimization problem has to be solved. This fits well the Internet of Things paradigm, where the typical scenario is given by a plethora of sensors with a low computational power transmit acquired data on the cloud, where computational power is not an issue. Some proposed algorithms capable to recover the input signal are discussed in [12; 13]. However, for the sake of simplicity, l_1 minimization-based algorithms are frequently adopted [14; 15]. The block

diagram of a *Standard CS-based* (Std-CS) system is shown in Figure 1.

Many papers recently appeared in the literature proposed methods to improve performance of a CS system, either at the encoder or at the decoder block. Generally, the key point of all these works is the *adaptation* of some quantities characterizing the CS framework to special features of the input signal. An overview of these approaches can be found in [16]. Several works are focused on the decoder side. In details, they improve performance by adapting the reconstruction algorithms [17; 18] or the sparsity basis [19] to the input signal features. Many others works are focused in the encoder side [20; 21; 22; 23; 24; 25] by adapting the sensing procedure to the input signal peculiarities. This work belongs to the latter group.

1.1. State-of-the-Art and original contribution

The state of the art of the encoder-side optimization is focused on two different research lines. The first aims to optimize the sensing matrix A by minimizing the *mutual coherence* between the input signal and A [20; 21]. A second approach considers additional peculiarities of the input signals [22; 23; 25]. Both approaches are undoubtedly capable of improving the performance with respect to the Std-CS, even if better resuts are achieved by using the second approach.

This paper is based on the results appeared in [22] and [25], both taking advantage of an a priori knowledge about a property of the input signal called *localization*. In general terms, a signal is localized when its energy is not uniformly distributed over the whole signal space. Interestingly, the localization is not an unusual property, as almost all real-life signals are localized [26]. In [22] the *Rakeness-based CS* (Rak-CS) is proposed. In details, the correlation

profile of the rows of the sensing matrix is adapted with that of the input signal, still preserving the fundamental requirement of a randomly generated rows required by the Std-CS for the correct reconstruction. In [25] the authors proposed the *Nearly-Orthogonal CS* (NeO-CS), aiming to reduce any redundancy in the sensing stage through geometric constraints (i.e., by setting a minimum value for the angle between any couple of rows of A). The difference is that, whereas Rak-CS adapts, in average, the sensing stage to the input signal statistics, NeO-CS imposes a proper characterization to each possible CS sensing stage.

In particular, with respect to the two considered publications, our original contribution can be summarized as follows.

- We present a detailed analysis of both sensing matrix adaptation techniques (the Rak-CS and the NeO-CS). Strengths, and in particular limitations of each of them are discussed.

- A new innovative approach, indicated with *Nearly Orthogonal Rakeness-based CS* (NOR-CS) is proposed and discussed to overcome the limitation identified in our analysis.

In particular, if Rak-CS ensure the easy generation of the sensing matrix A, it cannot be trimmed to the maximum possible adaptation level in order to avoid reconstruction problems with instances of the input signal appearing with a lower probability. Conversely, the NeO-CS is always pushed to the maximum adaptation level, but the geometrical constraints introduced to ensure a good behavior with less frequently input signal instances may prevent the convergence of the process generating A. The proposed NOR-CS is

instead capable both to set adaptation to the maximum possible level, but also to ensure the convergence of the process generating A. This ensures performance results higher with respect to what obtained when applying both Rak-CS and NeO-CS, that are considered as reference cases. The comparison between the proposed approach and the two reference cases relies on both synthetic and real bio-signals case studies.

The rest of the paper is organized as follows. Section 2 gives a brief overview about the Compressed Sensing theory. Then, we provide a critical review of the Rak-CS and NeO-CS approaches. Once their advantages and limitations are discussed, Section 3 introduces in detail our proposed approach. In Section 4 the numerical results are presented testing our proposed method and those referred with synthetic signals and electroencephalograph (EEG) signals. Finally, Section 5 concludes this paper.

2. Adapted Compressed Sensing

The CS mechanism is designed to process chunks of the input signal that must be *sparse*. Let be $x = (x_0, \ldots, x_{n-1})^\top$ the input vector that contains n successive Nyquist samples, where $\cdot^\top$ stands for vector transpose. We say that x is k-sparse if its representation on the basis $S \in \mathbb{R}^{n \times n}$, i.e., $x = S\xi$, gives a coefficient vector $\xi \in \mathbb{R}^n$ that exhibits a number of non-zero components less than or equal to $k \ll n$ for any possible instance x [1; 2].

For this class of signals, a CS encoder generates m dimensional measurement vector $y = (y_0, y_1, \ldots, y_{m-1})^\top$ by projecting x on the rows of a *sensing matrix* $A \in \mathbb{R}^{m \times n}$, such that $y = Ax = AS\xi$ where $m < n$ ensure data compression.

At the decoder side, CS theory insures that a signal can be recovered from the measurement vector y by exploiting sparsity. It is done by solving the following optimization problem

$$\begin{aligned} \hat{\xi} = \arg\min_{\xi \in \mathbb{R}^n} \ & \|\xi\|_1 \\ \text{s.t.} \ & \|AS\xi - y\|_2^2 \leq \varepsilon^2 \end{aligned} \tag{1}$$

where the 1-norm $\|\xi\|_1$ is used as sparsity-promoting function, $\|AS\xi - y\|_2^2$ counts the noise affecting y (bounded by $\varepsilon \geq 0$), while the final reconstructed signal is $\hat{x} = S\hat{\xi}$.

To achieve a correct reconstruction, CS theory suggests matrices A composed by instances of Gaussian or sub-Gaussian random variables where the number of rows is in the order $m = \mathcal{O}(k \log n)$ [14; 15]. Conveniently for hardware applications [7; 5; 8; 9], A can still be random but it is forced in the set $A \in \{-1, +1\}^{m \times n}$, i.e., A is an antipodal matrix. This setting is chosen as reference, and from now on, it is implicitly assumed that A is antipodal. As a remark, encoder and decoder need to share the information of A, while the knowledge of S is required only at the decoding stage.

Despite the advantages that the CS presents, it is proved that an appropriate design of the sensing matrix A can significantly increases the performance. In [20] and [21] two deterministic methods for the matrix A optimization has been proposed that are able to outperform the Std-CS. However, methods that adapts the sensing stage to the statistics of the input signal can further increase the reconstruction performance, as for both the method in [23] and the Rak-CS approach. Nevertheless, limiting to the adoption of antipodal sensing matrix, we consider the Rak-CS approach as representative for the state of the art in the Adapted CS framework [16].

2.1. Rakeness-based CS

The reconstruction of the input signal in the CS framework is theoretically guaranteed if using a number of measurement $m = \mathcal{O}(k \log(n/k))$ and a few conditions are satisfied. Among them, one of the most important and known is the Restricted Isometry Property (RIP) [14]. Indeed, CS performance can be improved (i.e., either by increasing the signal reconstruction quality at a given m, or by allowing a correct reconstruction with a value of m smaller than what expected by the general theory) if some other constraints on the input signal can be exploited.

In this paper we focus on the *localization* property, introduced in [22]. A class of signal is localized if the expected energy distribution of instances is not uniform over the whole signal space. Mathematically, let us consider the $n \times n$ input signal correlation matrix $\mathcal{X} = \mathbf{E}[xx^\top]$. In terms of eigenvalues $\mu_0 \geq \mu_1 \geq \mu_2 \cdots \geq \mu_{n-1} \geq 0 \ \forall j \in \{0, 1, \ldots, n-1\}$ corresponding to the orthonormal eigenvectors $u_0, u_1, \ldots, u_{n-1}$, we have $\mathcal{X} = \sum_{j=0}^{n-1} \mu_j u_j u_j^\top$. Localization appears when the eigenvalues μ_j are not equal to each other, and can be quantified by computing the deviation of each eigenvalue from the isotropic case by

$$\mathcal{L}_x = \sum_{j=0}^{n-1} \left(\frac{\mu_j}{\mathrm{tr}\,(\mathcal{X})} - \frac{1}{n} \right)^2 = \frac{\mathrm{tr}\,(\mathcal{X})^2}{\mathrm{tr}^2\,(\mathcal{X})} - \frac{1}{n} \tag{2}$$

where $\mathrm{tr}(\cdot)$ stands for matrix trace. The value of $\mathcal{L}_x$ ranges from 0 (when considering instances of a white random stochastic process, where the signal presents an uniform energy distribution) to $\mathcal{L}_x = 1 - 1/n \approx 1$ (when $\mu_0 > 1$ and $\mu_j = 0,\ j = 1, 2, \ldots, n-1$, i.e., when the signal energy is always concentrated in a single direction).

An example of a localized signal is a low-pass signal: in the frequency domain, energy is concentrated on the lower part of the spectrum, whereas the energy that can be found on the upper part of the spectrum is much smaller. Note that localization is a property that does not depend on the specific basis used to express the input signal x. In the above definition of $\mathcal{X}$ we have implicitly considered the canonical, leading to eigenvalues μ_j and eigenvectors u_j. Even if we consider the input signal as expressed in a different basis (such as the Fourier basis as in the example), its correlation matrix still gives rise to the same eigenvalues, and so the same $\mathcal{L}_x$ according to (2). This observation is of particular importance, at the light of the fact that the sparsity property, required by the CS framework, indeed depends on a particular basis. As a consequence, for a given input signal, the sparsity property and the localization property are independent from the each other.

The Rak-CS approach discussed in [22] and better formalized in [24] exploits the observation that the reconstruction quality in a CS framework can be improved by increasing the energy captured by the projection operation $y = Ax$. If x is localized, it is easy to understand that this energy is maximized by aligning the rows of the matrix A with the directions where the signal is most probably concentrated, and identified from $\mathcal{X}$. Nevertheless, it is still important to explore less energetic directions over the whole space signal to guarantee a correct reconstruction of any possible instance x and to not go against to the fundamental precepts of the standard CS theory.

The trade-off between focus-exploration is not a trivial issue. To formalize it, let us define a generic row of A as $a = (a_0, a_1, \ldots, a_{n-1})^\top$, with a correlation matrix $\mathcal{A} = \mathbf{E}[aa^\top]$. The aim of Rak-CS is to maximize the *rakeness*,

defined as the average energy collected by a generic entry of the measurement vector $a^\top x$ and expressed as $\rho = \mathbf{E}_{a,x}[(a^\top x)^2]$, under the assumption that the rows of A are still random enough and ensured by imposing a cap on the localization $\mathcal{L}_a$ of the generic row a by means of a localization scaling parameter $l \geq 0$, whose typical value is $l = 0.5$ [26]. Mathematically, the rakeness optimization problem [22] is analytically solved in terms of eigenvalues λ_j and eigenvectors v_j of $\mathcal{A}$:

$$\begin{aligned} v_j &= u_j \\ \lambda_j &= \frac{1}{J}\left[1 + \frac{J\mu_j - \Sigma_1(J)}{\sqrt{\frac{\Sigma_2(J) - \frac{1}{J}\Sigma_1^2(J)}{\gamma - \frac{1}{J}}}}\right] \end{aligned} \tag{3}$$

with $j = 0, 1, 2, \ldots, J-1$, and $\lambda_j = 0 \quad \forall j \geq J$. J is the largest integer for which, according to (3), it is still possible to achieve $\lambda_{J-1} > 0$. The two terms $\Sigma_1(J)$ and $\Sigma_2(J)$ are defined as the partial sums $\Sigma_1(J) = \sum_{j=0}^{J-1} \mu_j$ and $\Sigma_2(J) = \sum_{j=0}^{J-1} \mu_j^2$, whereas γ is defined according to the localization scaling parameter l as

$$\gamma = \frac{1}{n} + \frac{l^2 \mathcal{L}_x}{(1 - n\mu_j)^2} \tag{4}$$

Finally, the correlation matrix $\mathcal{A}$ is given by

$$\mathcal{A} = \sum_{j=0}^{J-1} \lambda_j v_j v_j^\top \tag{5}$$

while the corresponding process generation rows of A possesses a localization

equal to[1]

$$\mathcal{L}_a = \left(\frac{l}{1 - n\mu_{n-1}} \right)^2 \mathcal{L}_x.$$

The limitation of this approach is that the correlation profile of all the rows of A is adapted to the one of the input signal. So, dealing with highly localized signals could result in redundancies in terms of information of the measurement vector. In this case, the rows that conform A tends to be very similar to each other, resulting in measurements with almost the same information. This scenario has to be avoided by lowering the value of l with respect to the typical values suggested in [22; 25].

2.2. Nearly-Orthogonal CS

As previously described, the key feature to deal with localized signals is the balance between focusing the projections to the most energetic directions and the exploration of the whole signal domain. Another solution to handle this trade-off is the NeO-CS approach discussed in [25] where the rows of the matrix A possess the same statistical distribution of x. This implies $\mathcal{L}_a = \mathcal{L}_x$.

To reduce redundancy in measurements, a lower bound is imposed on the angles between all couples of rows that compose the matrix A. Formally, indicate with a_j and a_k two rows of the matrix A. The cosine of the angle $\alpha = \widehat{a_j a_k}$ between them is $\cos(\alpha) = \frac{a_j^\top a_k}{||a_j||_2 ||a_k||_2}$. NeO-CS restricts the matrix A to be composed by couples of rows such that $|\cos(\alpha)| \leq c$ only. In [25] authors suggest to generate all m rows of A iteratively. Each new row is randomly obtained; if it satisfies the geometric constraint with respect to all

[1] $\mathcal{L}_a$ is computed as in (2), where λ_j replaces μ_j and where $\mathrm{tr}(\mathcal{A}) = n$ to be compliant with the generation of antipodal sequences.

rows already generated, it is added to A. Otherwise, it discarded, and a new row is generated and tested.

An approximation of the probability to generate a new row whose angle with another row has a cosine smaller than c is explained in detail in [25] and given by

$$\Pr\{|\cos(\alpha)| \leq c\} = \operatorname{erf}\left(\frac{c}{\sqrt{2\left(\mathcal{L}_a + \frac{1}{n}\right)}}\right). \tag{6}$$

The probability to generate a new row that satisfies the geometric constraint when more than one row has already been generated rapidly decreases with the number of rows.

Let us indicate with Z the number of iterations necessary to generate all m rows of A. According to [25], when we deal with a high localized signal, a better performance in the reconstruction stage is achieved when a smaller value of c is imposed; however, the difficulty to generate the matrix A also increases considerably, and so, the expected value of Z. Although the matrix A can be generated offline and locally stored, it is important to consider if the increase in the performance justifies a possible huge computational effort. For instance, Figure 2 shows the average number of iterations $\mathbf{E}[Z]$ required to obtain a new row of A as a function of the total number of rows m with $c = 0.1875$ and for $n = \{128, 256\}$, obtained by montecarlo simulations. The figure also reports profiles obtained by data extrapolation (solid line) that evidences that $\mathbf{E}[Z] > 10^6$ is necessary to get a compression ratio $CR = n/m = 2$. If we limit Z to be less than a certain Z_{max}, NeO-CS could hardly be applicable with certain c values.

From a geometric point of view, in the case of $\mathcal{L}_a = 0$, antipodal vectors

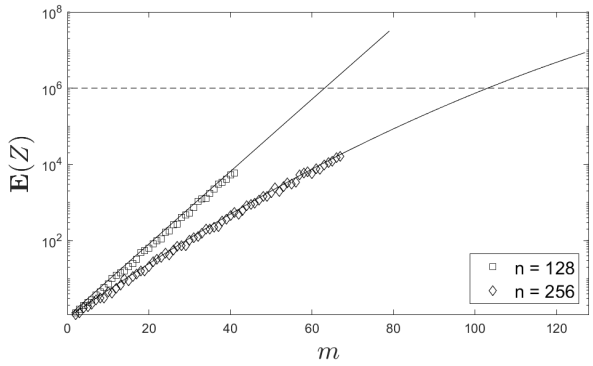

Figure 2: Average of the number of iterations $\mathbf{E}[Z]$ to generate the rows of the matrix A using NeO-CS with $c = 0.1875$ and two different values of n. The threshold value $\mathbf{E}[Z] = 10^6$ is also indicated.

can be represented by points uniformly distributed on the surface of a multi-dimensional sphere. As a result, when n increases, the angle between a couple of rows increases up to $\frac{\pi}{2}$. However, for $\mathcal{L}_a > 0$, the points are not more uniformly distributed on the surface and their distribution concentrates according to the assigned correlation matrix. The observable effect is that the average value of α, $\mathbf{E}[\alpha]$, slowly approaches the angle $\frac{\pi}{2}$. This is the reason why when $\mathcal{L}_a$ increases the difficult to generate the rows of A increases as well.

The impact of both localization and n is shown in Figure 3, that depicts the $\mathbf{E}[\alpha]$ profiles as a function of n for $\mathcal{L}_a = (0.001, 0.002, 0.005, 0.01)$. As one can observe, as n increases also $\mathbf{E}[\alpha]$ increases, but the speed of convergence of $\mathbf{E}[\alpha]$ to the $\frac{\pi}{2}$ asymptotic limit strongly depends on the localization value.

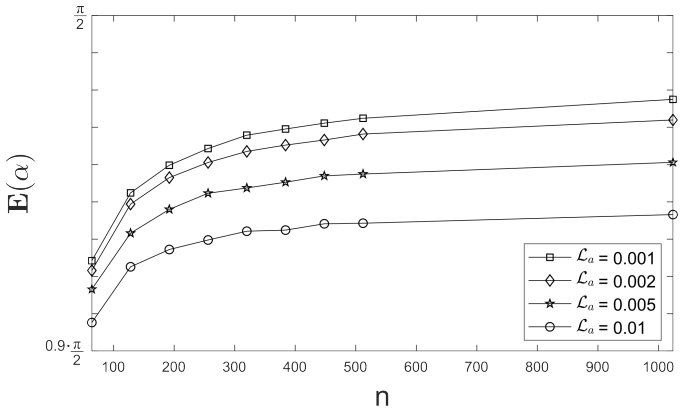

Figure 3: Average of the angle between two rows as a function of n.

3. Nearly Orthogonal Rakeness Based CS

With the aim to propose a new focus/exploration trade-off, we introduce the Nearly Orthogonal Rakeness-based (NOR-CS) approach. This method exploits the geometric constraints that characterize NeO-CS, and mitigate the hardness in generating the m rows of A by exploiting the localization as in the Rak-CS approach mathematically described by (3), i.e., by re-introducing the l parameter that scales the localization imposed to the A rows.

With NeO-CS we limit the maximum number of iterations that are allowed to generate a new row of A by Z_{max}, NOR-CS tries to exploit as much as possible the same geometric constraint introduced before where the hardness of a new generation is imposed on average, i.e., $\mathbf{E}[Z] \leq Z_{max}$. If this cap cannot be ensured, the localization of the next generated row is reduced by scaling the value of l.

For a known class of signal, i.e., for a known $\mathcal{L}_x$, the hardness to generate

the $\hat{m}$-th row of A with a given value of c has been modeled by the probability p that a process with a correlation profile evaluated with (3)(4) will generate a row for which the geometric constraint holds with the $\hat{m}-1$ already generated rows. Mathematically, fixed a-priori the values of $\mathcal{L}_x$ and n, p is a function of: the cosine value c, the localization scaling factor l and the number of steps $\hat{m}$. Due to the complexity in computing a closed-expression for p, we evaluated it by means of a look-up table $p = F_p(c, l, \hat{m})$ that has been estimated by Montecarlo simulations for the two values $n = 128$ and $n = 256$. This means that for a fixed n a system configuration is defined by the value of $\mathcal{L}_x$ and c, while the values of p for each possible $\hat{m}$ is estimated by following the procedure described here.

1. The first row of the matrix is not geometrically constrained so that it is generated as for the NeO-CS framework.
2. To obtain a new row, the method generates candidates with the same correlation profile until the geometric constraint is satisfied, i.e., $|\cos(\alpha)| \leq c$, where α is the angle between the current candidate and the first generated row. We use this approach to estimate p from the number of generated candidates.
3. The rows generation continue as in the last step until $\hat{m} = m_{max}$, where m_{max} is the maximum number of rows considered in the look-up table and at each new step, the geometric constraint must hold with all the already generated rows.

The p values in the look-table refer to Montecarlo simulation on 10000 different p estimation. As is shown in the example in Figure 4b, the hardness to obtain a new row $\hat{m}$ increases significantly as $\mathcal{L}_x$ increases

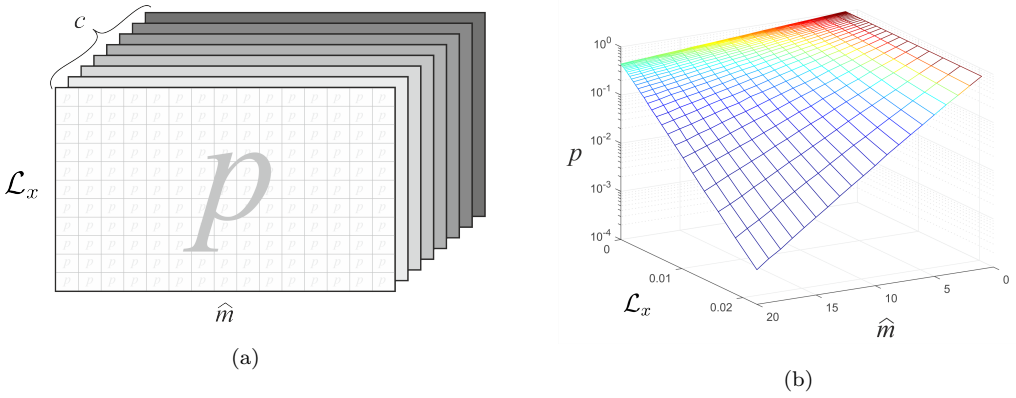

Figure 4: a) Representation of the look-up table. b) Look-up table for $n = 128$ and $c = 0.3125$.

Visually, the look-up table can be represented by the Figure 4a as a collection of tables containing the probabilities p to generate each row $\hat{m}$ for a given $\mathcal{L}_x$. One table for each value of c.

The algorithm we propose to generate A according to the NOR-CS approach is briefly described as follows, and summarized in Algorithm 1, where the look-up table is used to map the function F_p. Furthermore, we limit the (expected) number of iterations in the generation of each row with Z_{max}. The key feature of the algorithm is that the value of l is initially set to the desired value l_0. We use $l = l_0$ until the hardness of generating a single row exceeds a threshold p_{min} (computed according to the desired Z_{max}). After that, by means of the proposed look-up table, l is reduced to the maximum value that still ensure $p > p_{min}$. In detail:

1. First, c and the hardness p_{min} to generate the rows are fixed.
2. The first antipodal row is generated with an initial correlation profile evaluated by (3)(4) where l is equal to l_0.

$n = 128$, $\hat{m} = 35$

	$c = 0.1094$	$c = 0.1563$	$c = 0.2031$	$c = 0.25$
$l = 0.163$	$5.48 \cdot 10^{-4}$	0.0897	0.5199	0.8731
$l = 0.281$	$2.82 \cdot 10^{-4}$	0.0609	0.4657	0.8345
$l = 0.415$	$1.00 \cdot 10^{-4}$	0.0360	0.3466	0.7711
$l = 0.490$	$4.55 \cdot 10^{-5}$	0.0229	0.2874	0.7147
$l = 0.616$	$9.61 \cdot 10^{-6}$	$8.50 \cdot 10^{-3}$	0.1779	0.5753
$l = 0.711$	$2.51 \cdot 10^{-6}$	$3.90 \cdot 10^{-3}$	0.1155	0.4763
$l = 0.817$	$4.72 \cdot 10^{-7}$	$1.30 \cdot 10^{-3}$	0.0621	0.3425
$l = 1.010$	$2.34 \cdot 10^{-8}$	$1.21 \cdot 10^{-4}$	0.0154	0.1620
$l = 1.370$	$1.36 \cdot 10^{-10}$	$5.07 \cdot 10^{-7}$	$4.68 \cdot 10^{-4}$	0.0185
$l = 1.493$	$8.23 \cdot 10^{-12}$	$1.25 \cdot 10^{-7}$	$1.07 \cdot 10^{-4}$	$7.10 \cdot 10^{-3}$

$n = 128$, $\hat{m} = 50$

	$c = 0.1094$	$c = 0.1563$	$c = 0.2031$	$c = 0.25$
$l = 0.163$	$1.94 \cdot 10^{-5}$	0.0308	0.3882	0.8223
$l = 0.281$	$7.43 \cdot 10^{-6}$	0.0173	0.3342	0.7697
$l = 0.415$	$1.69 \cdot 10^{-6}$	$8.30 \cdot 10^{-3}$	0.2159	0.6911
$l = 0.490$	$5.36 \cdot 10^{-7}$	$4.30 \cdot 10^{-3}$	0.1666	0.6190
$l = 0.616$	$5.63 \cdot 10^{-8}$	$1.00 \cdot 10^{-3}$	0.0831	0.4501
$l = 0.711$	$8.02 \cdot 10^{-9}$	$3.38 \cdot 10^{-4}$	0.0448	0.3423
$l = 0.817$	$7.06 \cdot 10^{-10}$	$7.05 \cdot 10^{-5}$	0.0183	0.2135
$l = 1.010$	$9.52 \cdot 10^{-12}$	$2.16 \cdot 10^{-6}$	$2.50 \cdot 10^{-3}$	0.0733
$l = 1.370$	$6.01 \cdot 10^{-15}$	$7.36 \cdot 10^{-10}$	$1.60 \cdot 10^{-5}$	$3.20 \cdot 10^{-3}$
$l = 1.493$	$1.01 \cdot 10^{-16}$	$1.06 \cdot 10^{-10}$	$1.90 \cdot 10^{-6}$	$8.24 \cdot 10^{-4}$

$n = 256$, $\hat{m} = 60$

	$c = 0.1094$	$c = 0.1563$	$c = 0.2031$	$c = 0.25$
$l = 0.154$	$8.00 \cdot 10^{-3}$	0.4887	0.9338	0.9972
$l = 0.273$	$2.50 \cdot 10^{-3}$	0.3373	0.8789	0.9869
$l = 0.405$	$2.75 \cdot 10^{-4}$	0.1500	0.7363	0.9606
$l = 0.500$	$3.45 \cdot 10^{-5}$	0.0670	0.5501	0.9147
$l = 0.600$	$2.75 \cdot 10^{-6}$	0.0191	0.3752	0.8306
$l = 0.708$	$1.22 \cdot 10^{-7}$	$4.00 \cdot 10^{-3}$	0.1867	0.6700
$l = 0.823$	$2.30 \cdot 10^{-9}$	$4.71 \cdot 10^{-4}$	0.0696	0.4362
$l = 0.972$	$1.17 \cdot 10^{-11}$	$1.93 \cdot 10^{-5}$	0.0131	0.2027
$l = 1.386$	$4.00 \cdot 10^{-19}$	$2.12 \cdot 10^{-10}$	$1.05 \cdot 10^{-5}$	$3.70 \cdot 10^{-3}$
$l = 1.603$	$2.05 \cdot 10^{-20}$	$8.35 \cdot 10^{-14}$	$7.30 \cdot 10^{-8}$	$1.68 \cdot 10^{-4}$

$n = 256$, $\hat{m} = 90$

	$c = 0.1094$	$c = 0.1563$	$c = 0.2031$	$c = 0.25$
$l = 0.154$	$6.82 \cdot 10^{-4}$	0.3413	0.9026	0.9958
$l = 0.273$	$1.19 \cdot 10^{-4}$	0.1946	0.8244	0.9798
$l = 0.405$	$4.23 \cdot 10^{-6}$	0.0569	0.6309	0.9409
$l = 0.500$	$1.83 \cdot 10^{-7}$	0.0169	0.4036	0.8739
$l = 0.600$	$4.01 \cdot 10^{-9}$	$2.50 \cdot 10^{-3}$	0.2272	0.7564
$l = 0.708$	$3.69 \cdot 10^{-11}$	$2.46 \cdot 10^{-4}$	0.0792	0.5472
$l = 0.823$	$8.91 \cdot 10^{-14}$	$9.74 \cdot 10^{-6}$	0.0181	0.2859
$l = 0.972$	$2.91 \cdot 10^{-17}$	$7.79 \cdot 10^{-8}$	$1.50 \cdot 10^{-3}$	0.0902
$l = 1.386$	$1.20 \cdot 10^{-28}$	$2.38 \cdot 10^{-15}$	$3.13 \cdot 10^{-8}$	$2.16 \cdot 10^{-4}$
$l = 1.603$	$1.94 \cdot 10^{-30}$	$1.47 \cdot 10^{-20}$	$1.66 \cdot 10^{-11}$	$2.10 \cdot 10^{-6}$

Figure 5: Values extracted from the $F_p(c, l, \hat{m})$ look-up table with $\mathcal{L}_x = 0.01$.

3. The probability p to generate the next row is obtained from the look-up table $p = F_p(c, l, \hat{m})$. If $p > p_{min}$, the row is generated with the same value of l as the previous row. Otherwise, a maximum value of l is calculated by exponential interpolations of the values in the look-up table such that $p > p_{min}$ holds. With this new value of l, (3)(4) is evaluated and the row is generated.

Input: l_0, c, p_{min}, m.
l_0: initial value of l.
c: geometric constraint value.
p_{min}: probability threshold.
m: total of rows to be generated.
$l \leftarrow l_0$
$\mathcal{A} \leftarrow$ equation (3)(4)
$A(1,:) \leftarrow$ generate the first antipodal row with $\mathcal{A}$
for $i = 2$ **to** m **do**
  $p = F_p(c, l, i)$
  if $p < p_{min}$ **then**
    $l \leftarrow \arg\max_{\hat{l}} F_p(c, \hat{l}, i) > p_{min}$
    $\mathcal{A} \leftarrow$ equation (3)(4)
  end
  $A(i,:) \leftarrow$ generate antipodal row with $\mathcal{A}$
end
Output: A

Algorithm 1: NOR-CS pseudocode

As an example, an instance of the matrix A generated by each method is showed in the Figure 6 trying to match the localization of a low-pass input signal x. The matrix generated with the Std-CS (Figure 6a) shows a purely random distribution. In the case of the matrix generated with

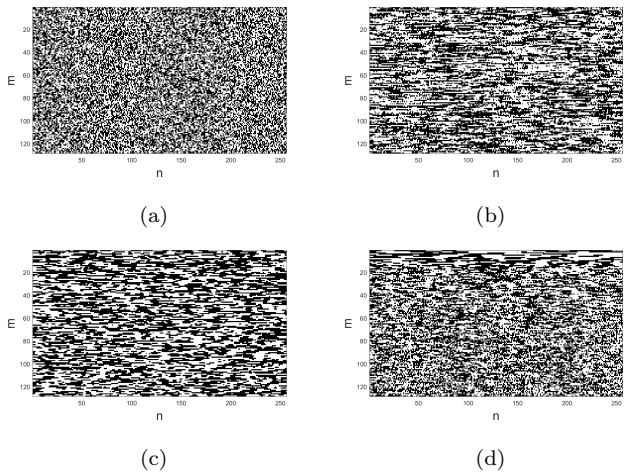

Figure 6: Matrix A generated with the different approaches: a) Std-CS, b) Rak-CS, c) Neo-CS and d) NOR-CS. Black dots correspond to +1, white dots to -1.

Rak-CS and NeO-CS (Figure 6b and 6c respectively), a low-pass profile is clearly identifiable. Yet, it is possible to see that sequences generated by NeO-CS is more localized than that obtained with Rak-CS due to higher localization. The Figure 6d shows the matrix generated by NOR-CS, where is possible to visualize how first rows are highly localized and then, gradually the localization decreases every certain numbers of rows.

Note that, among the many methods known to generate antipodal sequences with a certain correlation profile [27; 28; 29], the Rak-CS, the NeO-CS and the NOR-CS cases in the above example have been generated using the clipping of Gaussian instances [30; 31]. Basically, a $n \times n$ correlation matrix $\mathcal{G} = \sin\left(\frac{\pi}{2}\frac{n\mathcal{A}}{\mathrm{tr}(\mathcal{A})}\right)$ is used to generate a zero mean Gaussian vector g, such that antipodal a_j are computed by clipping the elements of g. As a result, also the obtained antipodal sequences are zero mean.

4. Numeric Results

Performance of all the approaches described above is provided through Montecarlo simulations. For each instance, we compute the Reconstructed Signal to Noise Ratio (RSNR) as the ratio between the energy of the input signal x and the energy of the difference between x and the reconstructed signal $\hat{x}$ expressed in dB. Starting from these data we evaluate the Average RSNR (ARSNR)

$$\text{ARSNR} = \mathbf{E}_{A,x} \quad \left[20 \log_{10} \left(\frac{||x||_2}{||x - \hat{x}||_2} \right) \right] \tag{7}$$

and the Probability of Correct Reconstruction (PCR)

$$\text{PCR} = \Pr\{\text{RSNR} \geq \text{RSNR}_{\min}\}, \tag{8}$$

this estimates the probability that the RSNR exceeds a minimum value.

4.1. Synthetic Low-pass Sparse signals test

To prove the performance of the NOR-CS and compare it with the other approaches mentioned above, the methods are tested with synthetic low-pass signals. Basically, n-dimensional instances x, localized and k-sparse in a certain basis S are generated starting from an instance of a random vector x' with zero mean and correlation matrix $\mathcal{X}'$. Computing $\xi' = S^{-1}x'$, we obtain ξ by keeping only the k higher absolute values in ξ'. Finally, the synthetic signal is obtained as $x = S\xi$. The correlation matrix $\mathcal{X}'$ is expressed as a Toepliz matrix $\mathcal{X}'_{i,j} = r^{\frac{\beta}{n}|(i-j)|} \ \forall i, j \in \{0, 1, \ldots, n-1\}$, with $r \in [0, 1]$, x' is a chunk of a stationary stochastic process with low pass profile. The factor β is empirically imposed to prevent an abrupt decay of the profile when n

is large. In this work, the Discrete Cosine Transform (DCT) is used as the orthonormal basis S, the setting of the parameters to generate the synthetic signals are reported in the Table 1 including the corresponding values of $\mathcal{L}_x$.

Table 1: Parameter settings to generate the synthetic signals.

Sparsity Basis	n	r	k	β	$\mathcal{L}_x$
DCT	128	0.7	12	150	0.026
DCT	256	0.7	25	150	0.024

For the Rak-CS approach we use the typical value $l = 0.5$ [26], with NeO-CS, for each value of c, we have limited the number of iterations to $Z_{max} = 10^5$ to generate a new row $\hat{m}$. As p_{min} for NOR-CS we adopt 10^{-3}, i.e, the expected average value of trials to be used for each row generation is 10^3. Once the parameters are established, for each method, 100 instances of the matrix A were drawn, each of them encoding 20 different instances of x. Also, non-idealities were modeled adding white Gaussian noise to each input vector x adapted to an Intrinsic Signal to Noise Ratio (ISNR) equivalent to 60 dB. For the estimation of the PCR in (8), we consider a value $\text{RSNR}_{min} = 55$ dB. Finally, the instances are reconstructed solving (1) with the SPGL1 toolbox[2]. Tuning of c for both NeO-CS and NOR-CS is done in order to reduce as much as possible the number of measurements required to obtain a PCR = 0.95. We refer to each of them value as m_{min}.

Performances are shown in the Figure 7, where we can observe that the

[2]This tool is available online in https://www.cs.ubc.ca/∼mpf/spgl1/download.html

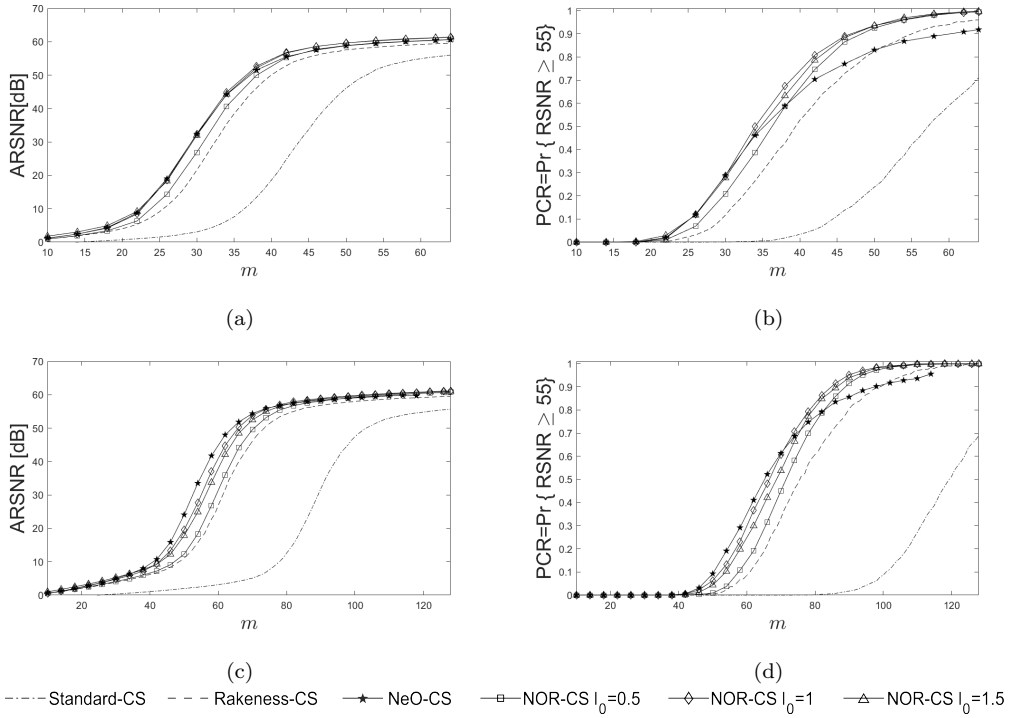

Figure 7: Performances in terms of ARSNR for a) n = 128 and c) n = 256 and PCR for b) n = 128 and d) n = 256 between the Std-CS and the optimized CS approaches. NeO-CS and NOR-CS are evaluated with the best values of c reported in the Table 2.

ARSNR and the PCR of all the optimized methods mentioned above have a better performance than the Std-CS approach, that is why from now only the optimized methods will be considered. In Figures 7a and 7c we can observe a similar performance in ARSNR for NeO-CS and NOR-CS, both were evaluated with the best value of c reported in the Table 2. Also, we can observe in the figure that NeO-CS and NOR-CS slightly outperform the Rak-CS. Interestingly, the Figures 7b and 7d shown a notably better performance in terms of PCR for NOR-CS compared with the other methods. Same results are in the Table 2, where NOR-CS shows a very positive impact in the reduction of m_{min}.

Table 2: Best performances for NeO-CS and NOR-CS.

Approach	l_0	$n = 128$		$n = 256$	
		c	m_{min}	c	m_{min}
NeO-CS	1.0	0.2031	51	0.2031	104
	0.5	0.1250	39	0.1328	85
NOR-CS	1.0	0.1406	38	0.1484	81
	1.5	0.1406	39	0.1328	83

In the Figure 8a we can observe how the localization is reduced by the parameter l in the process to generate the matrix A by three different versions of NOR-CS. This localization profile corresponds to the setting that provides the better performances reported in the Table 2 with $n = 256$, the constant values of l for the NeO-CS and Rak-CS are included only as a reference. In addition, we can observe in the Figure 8b the average of the number of iter-

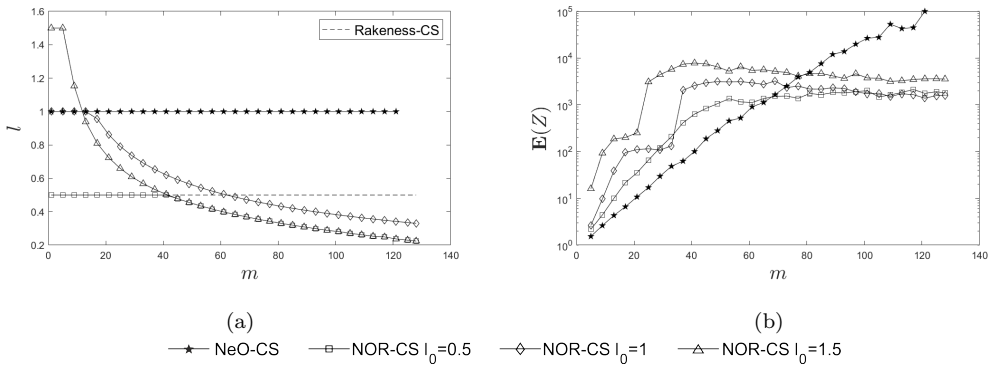

Figure 8: a) Values of l along the matrix A generation. b) Average of the number of iterations $\mathbf{E}[Z]$ of the optimized CS methods with geometric constraints. Both profiles correspond to the performances showed in the Figures 7c and 7d .

ations $\mathbf{E}[Z]$ as a function of m according with the localization profile showed in the Figure 8a. As was expected, the computational effort to generate a new row with NOR-CS is approximately constant after a certain number of row generation.

According with the Table 2 the NOR-CS with $l_0 = 1$ presents a slight outperformance from the other versions (at least dealing with synthetic signals). For this reason, from now we refer generally to NOR-CS implying $l_0 = 1$. In Table 3 we also report values of c that maximize the PCR for same reference values of m. This table includes the corresponding final l values as well as performance for the Rak-CS where the observed average values of c are included. As a result, the observed optimal values of c are not equal to the minimum tested values. This means that the imposed geometric constraint must be such that rows span different signal regions, but they are still inside the subspace where the signal concentrate its energy. This is also the case of the last proposed comparison in Figure 9. Here, c is the value

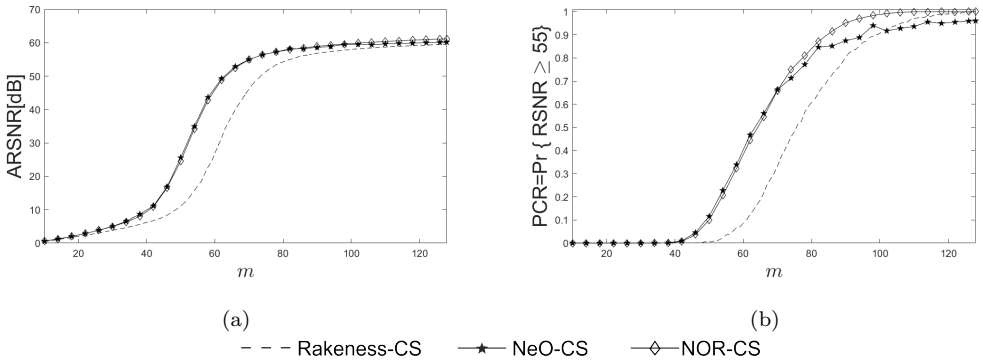

Figure 9: ARSNR for and PCR with the optimum values of c that maximize the performance in each row.

that maximize the performance for each considered number of rows in A. The Figure 9a shows that the NOR-CS approach presents a performance not different from NeO-CS in terms of ARSNR and both NeO-CS and NOR-CS have an outperforming with respect to the Rak-CS. However, the Figure 9b provides a considerable better performance of the NOR-CS in terms of PCR than NeO-CS and Rak-CS.

4.2. EEG signals Test

In this section, we propose the reconstruction of electroencephalograph (EEG) signals as an example to demonstrate the effectiveness of the NOR-CS. An EEG is a set of signals recorded from several electrodes on the scalp to analyze the brain activity. These signals provide information to identify different brain conditions and is useful to monitor the patient's health and diagnosis in the neuroscience, cognitive science and cognitive psychology areas. In particular we focus on Evoked Potentials (EPs), that consist in recordings of the electrical activity of the brain following a repetitive auditory stimulus

Table 3: Best value of c to obtain the best PCR in each row where: i) l values for NOR-CS indicate the observed final values with $l_0 = 1$; ii) performance for Rak-CS include observed average value of c, $\mu(c)$.

		NOR-CS			**NeO-CS**			**Rak-CS**		
m	CR	c	l	PCR	c	l	PCR	$\mu(c)$	l	PCR
80	3.20	0.1328	0.3621	0.8455	0.1875	1	0.7981	0.3699	0.5	0.6075
90	2.84	0.1484	0.4319	0.9520	0.1953	1	0.8745	0.3705	0.5	0.8050
100	2.56	0.1406	0.3439	0.9870	0.2031	1	0.9094	0.3710	0.5	0.9065
110	2.32	0.1484	0.3806	0.9995	0.2031	1	0.9359	0.3713	0.5	0.9705

with a time interval of 1 s between them to evaluate the auditory perceptual threshold [32; 33]. However, the individual responses of a spontaneous EEG are visually indistinguishable, for this reason the analysis of EPs suggests that the ensemble of the signals post-stimuli (epochs) has to be averaged to detect the response of the auditory stimulus. A possible scenario is showed in the Figure 10, where the EEG of the patient is collected by sensors of a battery-powered system. Due to the amount of information that each channel generates, the compression stage plays an important role for energy saving. It is proved that systems based in CS techniques are characterized by a low energy consumption in portable devices because the rate of the output data is reduced considerably [34; 35]. A properly designed sensing matrix A could greatly reduce the data to be transmitted or stored without compromising the quality of the input signal (in this case without compromise the correct EEG interpretation).

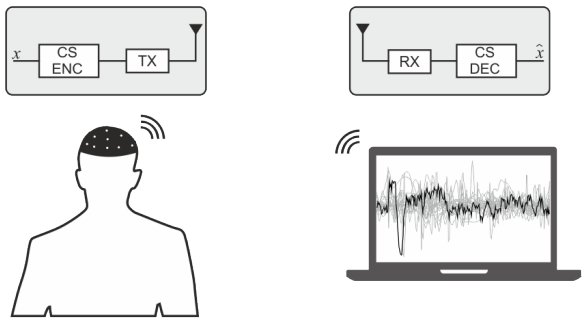

Figure 10: CS based system.

We use a data set that consists in EP recordings from a normal-hearing patient subjected to an auditory task. Basically, the test consists of listening speech syllables in one second intervals. The EEG acquisition was collected by 32 channels according the International 10/20 system of electrode placement and two channels used as noise reference to reject ocular artifacts. Each channel is organized in 700 epochs divided in two parts, the first 350 epochs are used as Training Set (TS) and it is employed in the correlation matrix estimation, while the second half named Data Set (DS) is used to test the CS-based approaches. According to [33] the channel C_z gives the strongest auditory response. Due to this, we compare the reconstructions of this channel only . The associated correlation matrix is calculated by (3) from the average denoised signal of the TS as in [36] using [37; 38; 39] to remove the artifacts.

To estimate the reconstructions quality we use the Mean Squared Error (MSE) MSE $= \frac{1}{n}\sum_{i=0}^{n}(\mathbf{x_i} - \hat{\mathbf{x}}_\mathbf{i})^2$. Here, $\mathbf{x}$ is an n-dimensional vector computed by averaging all the EEG chunks in the DS while $\hat{\mathbf{x}}$ is another n-dimensional vector referring to the average profile of the corresponding reconstructed signal for a certain CS framework. The MSE is computed

Table 4: Performance of the CS methods on EEG signals.

		Std-CS	**Rak-CS**	**NeO-CS**		**NOR-CS**	
m	CR	MSE (μV^2)	MSE (μV^2)	MSE (μV^2)	c	MSE (μV^2)	c
16	16	1.3065	0.8274	0.6954	0.1875	0.5859	0.1641
32	8	1.2055	0.6793	0.4246	0.1953	0.3827	0.1875
64	4	0.7580	0.3175	0.1852	0.1641	0.1435	0.2266

for three different setting where CR $= \{4, 8, 16\}$, that is, $m = \{64, 32, 16\}$. Results are in Table 4 and confirm that all optimized CS approaches have a better performance than the Standard CS. Remarkably, approaches based on geometric constrains present a better performance in terms of MSE than the Rak-CS. However, the proposed NOR-CS approach with $l_0 = 1.5$ presents the best performance among the optimized CS approaches. Note that for MSE, lower values correspond to a higher quality.

To visualize the performance related to the acquisition and reconstruction of the EPs, let us first consider the case with $m = 16$ that correspond to $CR = 16$. We can note in Figure 11a that the matching between the average reconstructed EPs using the Std-CS is very poor. The matching is improved using the Rak-CS as we can observe in the Figure 11b; however, the Figure 11c shows that the quality of the reconstruction using the Rak-CS is outperformed by the NeO-CS. As highlighted by the Figure 11d, the best performance is given by the NOR-CS, where the peaks of the auditory stimulus are easier to distinguish than those given by the other approaches. Moreover, if the requirement in terms of compression ratio can be relaxed

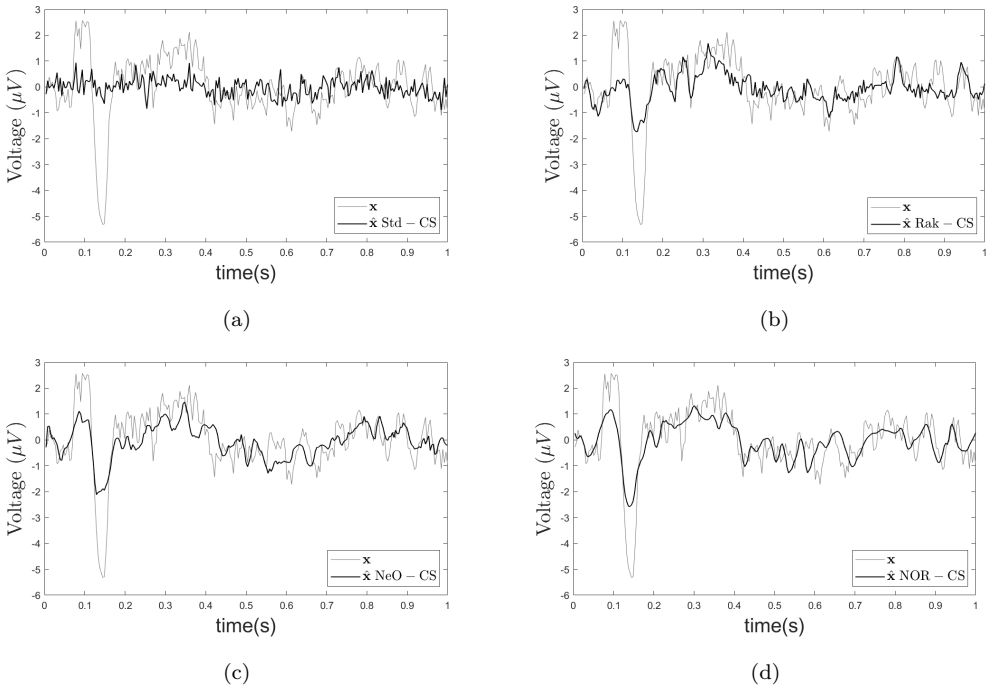

Figure 11: Comparison between the Average of the EEG raw signals $\mathbf{x}$ and the average of reconstructed signals $\hat{\mathbf{x}}$ with $m = 16$ that correspond to a compression ratio equal to 16. Results are for a) Std-CS, b) Rak-CS, c) NeO-CS and d) NOR-CS.

the difference between $\mathbf{x}$ and $\hat{\mathbf{x}}$ is strongly reduced. This is the case of Figure 12 where $m = 64$.

5. Conclusion

In this paper, we present a critical review of the state of the art of the optimized CS methods in the the sensing stage (Rakeness-based CS and Nearly Orthogonal-based CS). Advantages and limitations using these methods to generate the sensing matrix are discussed. In addition a new algorithmic solution named Nearly Orthogonal Rakeness-based CS is proposed with the aim

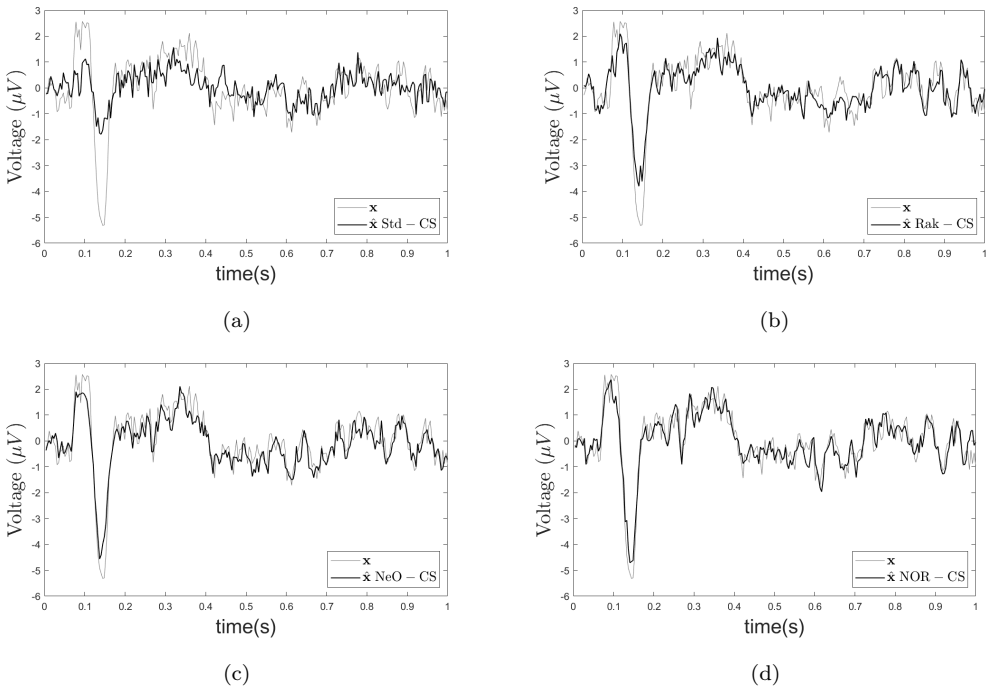

Figure 12: Comparison between the Average of the EEG raw signals $\mathbf{x}$ and the average of reconstructed signals $\hat{\mathbf{x}}$ with $m = 64$ that correspond to a compression ratio equal to 4. Results are for a) Std-CS, b) Rak-CS, c) NeO-CS and d) NOR-CS.

to overcome the limitations founded in the methods reviewed. This technique allows to exploit the geometric constraint as much as possible characterizing the sensing sequences of A with an adapted localization. After intensive simulations with synthetic low pass signals and electroencephalographic signals, the performances of all the methods discussed in this paper were compared. Results shows a remarkably better performance in therms of PCR of the proposed approach using synthetic low pass signals. Also, results demonstrate that electroencephalographic signals reconstructed with the proposed method present the best quality regardless the compression ratio. For this reason, using a high compression ratio, NOR-CS it can be considered as a suitable method to identify Evoked Potentials.

References

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

[2] E. J. Candès, M. B. Wakin, An introduction to compressive sampling, Signal Processing Magazine, IEEE 25 (2) (2008) 21–30.

[3] H. Mamaghanian, N. Khaled, D. Atienza, P. Vandergheynst, Design and exploration of low-power analog to information conversion based on compressed sensing, IEEE Journal on Emerging and Selected Topics in Circuits and Systems 2 (3) (2012) 493–501. doi:10.1109/JETCAS.2012.2220253.

[4] S. Liu, M. Zhang, W. Jiang, J. Wang, P. Qi, Theory and hardware implementation of an analog-to-information converter based on compressive

sensing, in: 2013 IEEE 10th International Conference on ASIC, 2013, pp. 1–4. doi:10.1109/ASICON.2013.6812033.

[5] F. Pareschi, P. Albertini, G. Frattini, M. Mangia, R. Rovatti, G. Setti, Hardware-algorithms co-design and implementation of an analog-to-information converter for biosignals based on compressed sensing, IEEE Transactions on Biomedical Circuits and Systems 10 (1) (2016) 149–162. doi:10.1109/TBCAS.2015.2444276.

[6] D. Gangopadhyay, E. G. Allstot, A. M. R. Dixon, K. Natarajan, S. Gupta, D. J. Allstot, Compressed sensing analog front-end for bio-sensor applications, IEEE Journal of Solid-State Circuits 49 (2) (2014) 426–438. doi:10.1109/JSSC.2013.2284673.

[7] M. Shoaran, M. H. Kamal, C. Pollo, P. Vandergheynst, A. Schmid, Compact low-power cortical recording architecture for compressive multichannel data acquisition, IEEE Transactions on Biomedical Circuits and Systems 8 (6) (2014) 857–870. doi:10.1109/TBCAS.2014.2304582.

[8] J. Yoo, S. Becker, M. Loh, M. Monge, E. Candès, A. Emami-Neyestanak, A 100mhz-2ghz 12.5x sub-nyquist rate receiver in 90nm cmos, in: 2012 IEEE Radio Frequency Integrated Circuits Symposium, 2012, pp. 31–34. doi:10.1109/RFIC.2012.6242225.

[9] X. Chen, E. A. Sobhy, Z. Yu, S. Hoyos, J. Silva-Martinez, S. Palermo, B. M. Sadler, A sub-nyquist rate compressive sensing data acquisition front-end, IEEE Journal on Emerging and Selected Topics in Circuits and Systems 2 (3) (2012) 542–551. doi:10.1109/JETCAS.2012.2221531.

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

[11] U. Dias, M. E. Rane, Comparative analysis of sensing matrices for compressed sensed thermal images, in: 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013, pp. 265–270. doi:10.1109/iMac4s.2013.6526420.

[12] D. L. Donoho, A. Maleki, A. Montanari, Message-passing algorithms for compressed sensing, Proceedings of the National Academy of Sciences 106 (45) (2009) 18914–18919. doi:10.1073/pnas.0909892106.

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

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

[15] E. J. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Transactions on Information Theory 52 (2) (2006) 489–509. doi:10.1109/TIT.2005.862083.

[16] M. Mangia, F. Pareschi, R. Rovatti, G. Setti, Adapted compressed sens-

ing: a game worth playing, IEEE Circuits and System Magazine(to appear) (2020).

[17] J. Zhang, Z. Gu, Z. L. Yu, Y. Li, Energy-efficient ecg compression on wireless biosensors via minimal coherence sensing and weighted ℓ_1 minimization reconstruction, IEEE Journal of Biomedical and Health Informatics 19 (2) (2015) 520–528. doi:10.1109/JBHI.2014.2312374.

[18] Z. Zhang, T. P. Jung, S. Makeig, B. D. Rao, Compressed sensing for energy-efficient wireless telemonitoring of noninvasive fetal ecg via block sparse bayesian learning, IEEE Transactions on Biomedical Engineering 60 (2) (2013) 300–309. doi:10.1109/TBME.2012.2226175.

[19] Y. Suo, J. Zhang, T. Xiong, P. S. Chin, R. Etienne-Cummings, T. D. Tran, Energy-efficient multi-mode compressed sensing system for implantable neural recordings, IEEE Transactions on Biomedical Circuits and Systems 8 (5) (2014) 648–659. doi:10.1109/TBCAS.2014.2359180.

[20] M. Elad, Optimized projections for compressed sensing, IEEE Transactions on Signal Processing 55 (12) (2007) 5695–5702. doi:10.1109/TSP.2007.900760.

[21] J. M. Duarte-Carvajalino, G. Sapiro, Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization, IEEE Transactions on Image Processing 18 (7) (2009) 1395–1408. doi:10.1109/TIP.2009.2022459.

[22] M. Mangia, R. Rovatti, G. Setti, Rakeness in the design of analog-to-information conversion of sparse and localized signals, IEEE Transac-

tions on Circuits and Systems I: Regular Papers 59 (5) (2012) 1001–1014. doi:10.1109/TCSI.2012.2191312.

[23] S. Pazos, M. Hurtado, C. H. Muravchik, A. Nehorai, Projection matrix optimization for sparse signals in structured noise, IEEE Transactions on Signal Processing 63 (15) (2015) 3902–3913. doi:10.1109/TSP.2015.2434328.

[24] M. Mangia, F. Pareschi, V. Cambareri, R. Rovatti, G. Setti, Rakeness-based design of low-complexity compressed sensing, IEEE Transactions on Circuits and Systems I: Regular Papers 64 (5) (2017) 1201–1213. doi:10.1109/TCSI.2017.2649572.

[25] M. Mangia, F. Pareschi, R. Rovatti, G. Setti, Adaptive matrix design for boosting compressed sensing, IEEE Transactions on Circuits and Systems I: Regular Papers 65 (3) (2018) 1016–1027. doi:10.1109/TCSI.2017.2766247.

[26] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, A rakeness-based design flow for analog-to-information conversion by compressive sensing, in: 2013 IEEE International Symposium on Circuits and Systems (ISCAS2013), IEEE, 2013, pp. 1360–1363. doi:10.1109/ISCAS.2013.6572107.

[27] R. Rovatti, G. Mazzini, G. Setti, S. Vitali, Linear probability feedback processes, in: Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on, IEEE, 2008, pp. 548–551. doi:10.1109/ISCAS.2008.4541476.

[28] R. Rovatti, G. Mazzini, G. Setti, Memory-m antipodal processes: spectral analysis and synthesis, Circuits and Systems I: Regular Papers, IEEE Transactions on 56 (1) (2009). doi:10.1109/TCSI.2008.920986.

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

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

[31] J. H. Van Vleck, D. Middleton, The spectrum of clipped noise, Proceedings of the IEEE 54 (1) (1966) 2–19. doi:10.1109/PROC.1966.4567.

[32] B. M. Savers, H. A. Beagley, W. R. Henshall, The mechanism of auditory evoked eeg responses, Nature 247 (5441) (1974) 481–483. doi:10.1038/247481a0.

[33] M. Scherg, J. Vajsar, T. W. Picton, A source analysis of the late human auditory evoked potentials, Journal of Cognitive Neuroscience 1 (4) (1989) 336–355. doi:10.1162/jocn.1989.1.4.336.

[34] F. Pareschi, M. Mangia, D. Bortolotti, A. Bartolini, L. Benini, R. Rovatti, G. Setti, Energy analysis of decoders for rakeness-based compressed sensing of ecg signals, IEEE Transactions on Biomedical Circuits and Systems 11 (6) (2017) 1278–1289. doi:10.1109/TBCAS.2017.2740059.

[35] A. Marchioni, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, Sparse sensing matrix based compressed sensing in low-power ecg sensor nodes, in: 2017 IEEE Biomedical Circuits and Systems Conference (BioCAS), 2017, pp. 1–4. doi:10.1109/BIOCAS.2017.8325155.

[36] N. Bertoni, B. Senevirathna, F. Pareschi, M. Mangia, R. Rovatti, P. Abshire, J. Z. Simon, G. Setti, Low-power eeg monitor based on compressed sensing with compressed domain noise rejection, in: 2016 IEEE International Symposium on Circuits and Systems (ISCAS), 2016, pp. 522–525. doi:10.1109/ISCAS.2016.7527292.

[37] A. de Cheveigné, J. Z. Simon, Denoising based on time-shift pca, Journal of Neuroscience Methods 165 (2) (2007) 297–305. doi:10.1016/j.jneumeth.2007.06.003.

[38] A. de Cheveigné, J. Z. Simon, Sensor noise suppression, Journal of Neuroscience Methods 168 (1) (2008) 195–202. doi:10.1016/j.jneumeth.2007.09.012.

[39] A. de Cheveigné, J. Z. Simon, Denoising based on spatial filtering, Journal of Neuroscience Methods 171 (2) (2008) 331–339. doi:10.1016/j.jneumeth.2008.03.015.