ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

ARCES – ADVANCED RESEARCH CENTER ON ELECTRONIC SYSTEMS

# Antipodal random sequences with prescribed second-order statistics: application to Compressive Sensing and UWB systems based on DS-CDMA

**Mauro Mangia**

SUPERVISORS

Professor Riccardo Rovatti

Professor Gianluca Setti

COORDINATOR

Professor Claudio Fiegna

DOCTORATE ON INFORMATION TECHNOLOGY

JANUARY 2010 – DECEMBER 2012

XXV CYCLE – ING-INF/01 – 09/E3

FINAL EXAM 2013

“The language and concepts contained herein
are guaranteed not to cause eternal
torment in the place where the guy
with the horns and pointed stick
conducts his business"

*Frank Zappa*

# Preface

BEFORE starting my PhD it was usual for me to hear a strange short sentence: «*Random is better than...*». Why is randomness a good solution to a certain engineering problem? Which applications get benefits by randomization? I think that the obvious answers are: because there are many natural phenomena that are either completely random or present a partially random characterization; because there are many applications that work with classes of real signals modeled by mixed deterministic/random characterization, because many noise sources are modeled by random processes or because in many system models a partial random characterization guarantees useful properties.

There are many other possible answers, and all of them are related to the considered topic. In this thesis I will discuss about two crucial topics, focusing the attention on two specific applications that take advantage by randomizing some signals involved in intermediate ma-

nipulations. In particular, advantages are guaranteed by shaping the second order statistic of antipodal sequences involved in intermediate signal processing stages.

The first topic is in the area of analog-to-digital conversion, and it is named Compressive Sensing (CS). CS is a novel paradigm in signal processing that tries to merge signal acquisition and compression at the same time. Consequently it allows to direct acquire a signal in a compressed form, or to collect its information content without relying on classical digitalization.

For this reason it is usual to call this signal acquisition procedure Analog to Information Conversion (AIC). To obtain its primary goal an AIC needs *a priori* information about the signals subject to the information extraction. AICs based on CS assume that the acquired signals are sparse, i.e., when they are expressed on a proper basis, they presents only a few non zero coefficients.

For this class of signals, acquisition at a given time window can be done by a set of projections on a suitable set of sensing waveforms, while reconstruction is achieved by solving an optimization problem relying on the sparsity assumption.

This is a compressed acquisition scheme because the amount of needed projections in the considered time window is significantly smaller than the number of digital words generated by an analog-to-digital converter that follows the classical Nyquist/Shannon approach.

In this thesis, after an ample description of the CS methodology and its related architectures, I will present a new approach that tries to achieve high compression by design the second order statistics of the projection functions.

The second topic addressed in this thesis is about personal communication. In this case, the aim is to guarantee a link between people or process units that is, in principle, independent of the environment and of their relative positions. Modern communication systems must possess the ability to reach a user whatever his location and speed, thus forcing the consideration of unsupervised, asynchronous and shared radio channels.

A recognized strategy to obtain these features is the adoption of ultra-wideband (UWB) systems that fill large portions of the spectrum with low-power-density signals, possibly overlapping with different systems and services from whose point of view they tend to appear as an increase of the thermal noise floor. An option to produce and decode UWB signals is direct-sequence spreading with multiple access based on code division (DS-CDMA). Such a classical technique [1] may allow arbitrary spreading and is suitable for extremely simple implementations of both transmitter and receiver.

Focusing on this methodology, I will address the coexistence of a DS-CDMA system with a narrowband interferer that may be either an intentionally emitted jammer or the abstract resumption of the effects of a traditional non-UWB service. To do so, among all the possible performance figures that quantify quality of communication (e.g., multipath robustness, system capacity in terms of user number) I concentrate in the joint effect of both multiple access (MAI) and narrowband (NBI) interference on a simple matched filter receiver.

I will show that specific scenarios of NBI exist in which, when spreading sequence statistical properties are suitably designed, performance improvements are

possible with respect to a system exploiting chaos-based sequences minimizing MAI only.

In both introduced topics antipodal sequences with prescribed second order statistics will be considered for an improvement in term of global performance. This motivates the title of this thesis, but the diversity of the two considered topics suggests a separate discussion.

The first part of this thesis is about CS and a new proposed methodology on the projecting sequence generation, while the second one discusses the improvement produced in a UWB system based on DS-CDMA when the statistical properties of the spreading sequence are properly tuned.

# Contents

# List of Figures

PART

I

# Compressive Sensing

CHAPTER

# 1 Compressive Sensing

THIS chapter introduces the application of some recently developed signal-processing techniques to the sensing of physical quantities, i.e., to their conversion into a sequence of samples that can be processed by a digital electronic system for the most diverse purposes.

Conventional approaches to this task are based on the celebrated Shannon-Nyquist theorem [2, 3], stating that the sampling rate must be at least twice the highest frequency in the band of the signal (the so-called Nyquist frequency). This principle is the basis of almost all methods of acquisition used in nowadays audio and video consumer devices, in the processing of medical images, in the operation of radio receivers, etc;

*Compressed Sensing* (CS) is a recently introduced paradigm for the acquisition/sampling of signals that violates the Shannon-Nyquist theorem providing that additional assumptions can be made.

A bird's eye view of CS shows that it is based on

two general concepts: *sparsity*, which materializes the needed additional assumption, and *incoherence* between the sampling protocol and the signal structure.

Sparsity expresses the idea that the information content of a signal can be much less than what is suggested by its bandwidth, or, for a discrete-time signal, that the number of its true degrees of freedom may be much smaller than its time length. Actually, many natural signals are sparse in the sense that they have a very compact representation when expressed with respect to a suitable reference system and are therefore susceptible to CS.

Incoherence extends the concept of duality between time and frequency. It is used to formalize the fact that when two domains are incoherent, objects that have a sparse representation in the first of them spread their energy over a wide support when seen from the point of view of the other domain.

It is evident that the first domain is optimal when it comes to expressing and characterizing the signal, while the second is to be preferred for sensing operations since even a few scattered measurements have the chance of capturing the signal energy. This is exactly what happens, for example, when I want to acquire a sinusoidal profile of unknown frequency. Since such a signal is extremely sparse in the frequency domain, the only two non-zero components of its spectral profile are incredibly effective in representing it. Yet, nobody would randomly probe the frequency axis at few frequencies with the hope of coming across the one at which the signal is present, thus being able to recover the amplitude and phase. Instead, we know very well that only few samples in the time domain

are enough to capture all the signal features.

Generalizing all this, a CS acquisition architecture analyzes the target sparse signal by taking few measurements in the domain in which the energy is widespread and thus easy to collect. If this is done properly, the resulting samples can be subsequently processed by algorithmic means to reconstruct the small representation in the domain in which the signal is sparse.

The theoretical and practical machinery needed to perform CS in realistic conditions is being rapidly developed to obtain acquisition mechanisms that can be labeled as Analog-to-Information (AI) converters [4]. In fact, once the proper domain has been found in the form of a waveform basis along which signals can be expressed as a linear combination of few number of non-zero coefficients, the actual information being carried by the signals will be found in the positions and the magnitudes of those coefficients [5] [6] that are the true target of the CS systems.

To reach a more exhaustive description of the acquisition and reconstruction procedures, I now introduce the formalisms about both CS and its corner store concepts: sparsity and incoherence. Preamble of this discussion is a brief description of the well know analog signal acquisition based on the Nyquist-Shannon theorem. These two blocks wholly summarize this chapter which aims to explain what compressive sensing is.

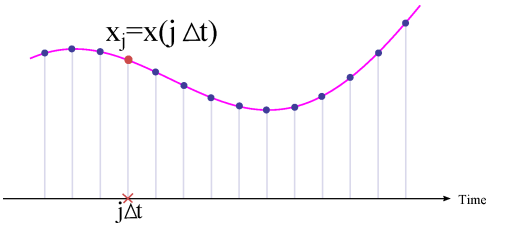



**Figure 1.1:** the classical acquisition procedure based on the Nyquist/Shannon theorem.

## 1.1 Nyquist/Shannon Theorem

As introduced before, the digitization of an analog signal is usually done by following the Nyquist-Shannon theorem [2, 3].

Let $x(t)$ a generic signal expressed in the time domain; $X(f)$ is its Fourier transform:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-\mathbf{i}2\pi ft} \mathrm{d}t$$

where **i** is the imaginary unit. By this definition, $x(t)$ is a *band-limited* signal if $X(f) = 0$ for all $|f| > B$, where $B$ is the signal band.

All band limited signals $x(t)$ can be represented by a stream of uniform samples at an interval $\Delta t$. The resulting sequence is denoted by $x_j$ for all possible integer values $j$, where $x_j = x(j\Delta t)$. At same time I define the sample rate $f_s = 1/\Delta t$. A visual example is shown in Figure 1.1.

A sufficient condition to perfect reconstruct $x(t)$ from $x_j$ is $f_s > 2B$, where the bound $2B$ is called the Nyquist rate. Under this assumption, the original signal can be reconstructed by a simple interpolation, i.e.,

$$x(t) = \sum_{j=-\infty}^{+\infty} x(j\Delta t) \cdot \mathrm{sinc}\left(\frac{t - j\Delta t}{\Delta t}\right)$$

It is important to highlight that this approach relies on a sufficient condition and so it is possible to think about different acquisition schemes capable of providing an advantage with respect to the Nyquist/Shannon approach. To reach this aim I formulate these questions:

- Is it possible to represent a generic signal in time without using sampling?
- Is a different acquisition procedure capable of collecting sufficient signal information to achieve perfect reconstruction?

In general an obvious answer to both questions is no. Both answer might become positive when some other hypotheses on the input signal are satisfied.

Compressive Sensing is a way to answer both previous questions. The signal information content is captured without sampling, while additional hypothesis on the existence of a sparse representation of the signal hold.

## 1.2 Compressive Sensing

In the classic approach, based on analog signal sampling, digital vales are obtained by a convolution between the input signal and a train of Dirac delta functions. A generic sample $x_j$ is obtained by

$$x_j = \langle x(t), \delta(t - j\Delta t) \rangle = \int_{-\infty}^{+\infty} x(t) \cdot \delta(t - j\Delta t) \mathrm{d}t = x(j\Delta t)$$

where $\langle \cdot, \cdot \rangle$ stands for dot product and $\Delta t$ is taken according to the Nyquist/Shannon theorem.

Now, let me consider a generic length $T$ time window, herein a fixed amount of samples $n$ is needed to represent the signal.

The basic idea of CS is a different choice of the projection functions needed to obtain the set of samples. Instead of Dirac delta functions, CS leads for set of functions able to extract the signal information by fewer projections with respect to the amount needed by traditional sampling.

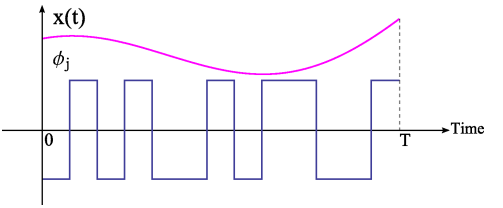
Let $\varphi_j$ with $j = \{1, 2, \ldots, m\}$ be the projecting functions (usually named sensing functions), and $y = (y_1, y_2, \cdots, y_m)$ the measurement vector.

$$y_j = \left\langle x(t), \varphi_j(t) \right\rangle = \int_{-\infty}^{+\infty} x(t) \cdot \varphi_j(t) \mathrm{d}t$$

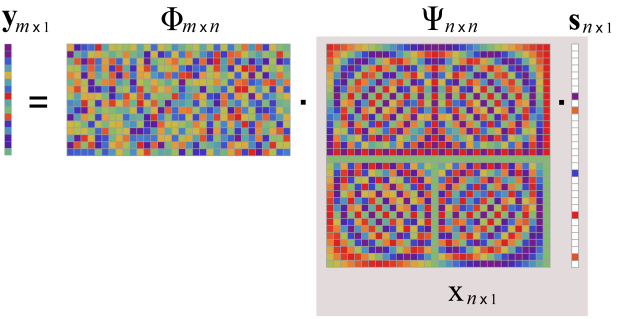
Compression is achieved when $m < n$, but in which conditions is it possible to obtain an accurate reconstruction of the analog signal with only $m < n$ measurements? How can I design the sensing functions to capture all the signal information required by the reconstruction stage. And also, how can I reconstruct the input signal?

Although it is possible to develop CS theory in continuous time/space, I restrict the analysis to the digital domain for a clearer description. So, $x(t)$ in a length $T$ time window is represented by an array $x \in \mathbb{R}^n$ that corresponds to the sampled signal at Nyquist/Shannon frequency and all sampling functions are represented by the sequences $\varphi_j = \{\varphi_{j,1}, \cdots, \varphi_{j,n}\}$ with $j \in \{1, \cdots, m\}$ so that the measurement vector is $y = \Phi x$ where $\Phi$, called sensing matrix, denotes an $m \times n$ matrix with the vectors $\varphi_j$ as rows.

Now, the input signal reconstruction can be achieved by recovering $x$ from $y = \Phi x$, but when $m < n$ it is an ill-posed inverse problem in general, i.e.,there exist infinite




(a)




(b)

**Figure 1.2:** (a)both the input signal $x(t)$ and a projecting function in a time window; (b) the matrix representation of the acquisition procedure based on Compressed Sensing.

$\hat{x} \in \mathbb{R}^n$ such that $\Phi\hat{x} = y$. To arrive at unique and correct solution of this inverse problem other hypotheses are required on the class of signals being acquired; this means that there exist some classes of signals for which a unique solution can be found, i.e., signal recovery is possible.

CS takes into account sparse signals. As mentioned before, sparse means that all signal instances can be expressed as a linear combination of few functions in a proper base. Let $\Psi$ be an $n \times n$ matrix whose columns are vectors of the basis where the signal is always sparse. The input signal can then be written as $x = \Psi s$, where

$s \in \mathbb{R}^n$ is a vector that contains the coefficients of the signal projection on the columns of $\Psi$.

$$\Psi = \left( \psi_1 \middle| \psi_2 \middle| \cdots \middle| \psi_n \right)$$

$$s = \{s_j\}_{j=1,\cdots,n} \quad \text{where} \quad s_j = \left\langle x, \psi_j \right\rangle = \sum_{i=1}^{n} \Psi_{i,j} x_i$$

Note that sparsity is not a hard assumption: there are many classes of signals[1] that present this property when expressed on a proper basis.

Quantify the sparsity level of a proper class of signals means identifying the maximum amount of simultaneous non zero coefficients on $s$ for a generic instance. When a class of signals is $K-$sparse, for each realization $s$ contains at most $K$ non null elements. Considering sparse signals, it is now possible to link the measurement vector with the sparse representation in $\Psi$, so that

$$y = \Phi x = \Phi \Psi s = As$$

where $A \in \mathbb{R}^{m \times n}$ denotes the linker operator between the measurements contained in $y$ and the sparse vector $s$. A visual representation of this procedure is shown in Figure 1.2 (b).

The hypotheses of sparsity can now be used to develop the reconstruction stage. I am looking for a vector $\hat{s}$ such that $A\hat{s} = y$, and on all possible vectors $\hat{s}$, I am looking for the sparsest. This corresponds to the next optimization problem:

$$\begin{aligned} \hat{s} &= arg\{\min_a \|a\|_0\} \\ &s.t.\ Aa = y \end{aligned} \tag{1.1}$$

[1] bio-signals, images, radar signals, ultrasound signals and so on.

yielding the reconstructed input signal by $\hat{x} = \Psi\hat{s}$.

$\|\cdot\|_0$ represents the $l_0$ norm that correspond to the amount of non null elements in the basis expansion of $x$ over $\Psi$.

Solving an optimization problem based on the minimization of the $l_0$ norm is a very hard topic since it is combinatorial in nature, although there are some contributions that demonstrate an equivalence between 1.1 and a convex optimization problem based on the minimization of the $l_1$ norm [7], that is

$$\begin{aligned} \hat{s} &= arg\{\min_a \|a\|_1\} \\ &s.t.\ Aa = y \end{aligned} \tag{1.2}$$

where the $l_1$ norm of a generic length $P$ vector, $p$, is equivalent to $\|p\|_1 = \sum_{i=1}^{P} |p_i|$.

This new optimization problem can easily be solved with respect to 1.1. There are also many contributions in literature which present different techniques for the solution of 1.2. See [5, 8–12] for more details. A more detailed discussion about reconstruction procedures will be reported in Chapter 2.

As a conclusion of this introduction to CS I must explain the last missing piece of CS theory. How are the sampling sequences $\phi_j$ generated to guarantee a correct reconstruction? The answer is addressed by to two different arguments:

- **incoherence**: CS Theory is based on the features of two set of functions represented by $\Phi$ an $\Psi$. The coherence measure the largest correlation between

any two elements of the couple $\phi_j$, $\psi_k$ defined as:

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} \left| \langle \phi_j, \psi_k \rangle \right|$$

By working with normalized vectors, $\mu$ is bounded by $[1, \sqrt{(n)}]$. CS is thought to work with low coherence values, meaning that a signal with a sparse representation on the domain $\Psi$ has a dense representation when it is projected on $\Phi$ and so each projection is able to collect signal information which is carried out at the reconstruction stage.

- **restricted isometry property**: as discussed before $A = \Phi\Psi$ is the operator that maps the signal in the sparse domain to its measurements. Imposing the restricted isometry property on $A$ means working with an operator able to preserve the Euclidean length of $K$-sparse vectors in the measurement vector $y$. It also guarantees that $s$ is not in the null space of $A$. A more detailed discussion of this aspect is left to the next chapter.

A guideline for the sensing sequences $\phi_j$s to guarantee both requirements is: generate $\phi_j$ as instances of both random Gaussian variables or binary/antipodal i.i.d random variables [6]. An example of a sampling sequence that respects both requirements is shown in Figure 1.2 (a). In this case $\phi_j$ is taken as a realization of a set of i.i.d. antipodal random variables.

CHAPTER

# 2 Compressive Sensing Architectures

This chapter aims to highlight relative strengths and weaknesses of some of the recently proposed architectures for hardware implementation of Analog-to-Information Converters based on Compressive Sensing.

To do so, the most common architectures are analyzed when saturation of some building blocks is taken into account, and when measurements are subject to quantization to produce a digital stream. Furthermore, the signal reconstruction is performed by established and novel algorithms (one based on linear programming and the other based on iterative guessing of the support of the target signal), as well as their specialization to the particular architecture producing the measurements. Performance is assessed both as the probability of correct support reconstruction and as the final reconstruction error.

My results help highlighting pros and cons of various architectures and giving quantitative answers to some

typical design-oriented questions. Among these, I show: i) that the RMPI architecture and its recently proposed adjustments are probably the most versatile approach though not always the most economic to implement; iii) for each architecture, the trade-off between number of measurements and number of bits per measurements (given a fixed bit-budget); iv) pros and cons of the use of Gaussian *vs.* binary random variables for signal acquisition.

## 2.1 Introduction

Here I will show you *design guidelines* with the aim to reduce the gap between theoretical development of CS [5] [6] for AICs [4] and their hardware implementation.

More precisely, following [5], I will assume that I wish to acquire analog signals possessing a property called *sparsity*, which amounts to know that a basis exists with respect to which the representation of a signal has a small number of non-null coefficients as described in Chapter 1. These represent the *information content* of the signal, so that sparsity implies that the latter is limited and, for example, substantially smaller than the waveforms bandwidth. Furthermore, the above assumption paves the way for developing AICs [4], i.e. more efficient circuits/systems for translating the analog signals into a certain number of bits that are sufficient to reproduce their information content and which do not follow the classical Analog-to-Digital conversion framework based on Shannon-Nyquist theory. In other words, AICs must be able to identify the above coefficients relying on a number

of samples (more generally, "measurements") which is lower than what the Nyquist criterion would dictate, so that the number of bits needed to encode the original signal may be less than what would be entailed by a straightforward Pulse-Code-Modulated (PCM) stream with a number of bit-per-sample sufficient to match the same accuracy requirements.

The fact that the number of bits used is less than what a conventional ADC would produce to obtain the same quality by means of a PCM justifies the name "compressive" and is also the key feature that *makes AICs a promising technique in all the cases in which the resources allocated to acquisition and/or transmission of the acquired data are limited*.

Starting from the first example in [4], several architectures of AICs have been proposed in the literature [13]- [14], all of which enjoy the same theoretical ground, i.e. the possibility of performing a sparsity-informed conversion in digital form of the information carried by the input signal. Needless to say, any real implementation of any of the above architectures would violate at least some of those assumptions since, first of all, the unavoidable saturations and quantizations in real circuits prevent a truly linear processing of the signal. It is also fundamental to stress that the effect on AIC performance of the above implementation non-idealities can be properly defined only if one takes into account the result of the corresponding "decoding"/reconstruction operation, i.e. the one which that takes the bit stream at the AIC output and is able to compute the non-null coefficients in the basis expansion. More specifically, differently with respect to classical ADCs where this operation is performed by

low-pass filtering[1], one needs to take into account that when "encoding" is performed via an AIC, reconstruction must be obtained by inverting an under constrained linear system of equations when the corresponding matrix obeys some assumptions [10] [11] and by adopting specifically designed algorithms [8] [9].

All these ad-hoc solutions are difficult to analyze from a fully theoretical point of view though they are demonstrated to work at least in some cases by the researchers proposing and testing them.

What I do here is a survey of some of the most common architectures in order to establish a common framework in which a fair comparison can be made by simulating them in various operating conditions, thus establishing pros and cons for all of them.

## 2.2 Mathematical Model

For an $n$ dimensional vector $a = (a_0, \ldots, a_{n-1})^\top$ I define the support of $a$ as

$$supp(a) = \left\{ j = 0, \ldots, n-1 | a_j \neq 0 \right\}$$

its sparsity $spar(a)$ (sometimes indicated as $L_0$ norm) as the cardinality of $supp(a)$ and its usual $p$-norm as

$$\|a\|_p = \left( \sum_{j=0}^{n-1} |a_j|^p \right)^{1/p}$$

Sampled in time, the signal $x(t)$ that I want to acquire is represented by an $N$-dimensional vector of real numbers $x$. Following [5] I will assume that a suitable basis exists

[1] Plus, possibly, decimation filtering, in case of oversampled ADCs

whose vectors are the columns of the $N \times N$ matrix $\Psi$, and that the signal of interest are $K$-sparse, which means that for any instance of $x$ there is an $N$-dimensional vector $a$ such that $x = \Psi a$ and $spar(a) \leq K$.

Starting from $x$, the acquisition system produces $M$ measurements by means of a linear projection, i.e., by correlating the elements of $x$ with some coefficients. By arranging those coefficients as the rows of the $M \times N$ matrix $\Phi$ I obtain an $M$-dimensional measurement vector $y = \Phi x$. Since the input signal will be corrupted by noise, what I will acquire is the signal $\tilde{x} = x + v$, so that the measurement vector becomes $y = \Phi\tilde{x} = \Phi(\Psi a + v)$, where $v$ is the vector of noise coefficients superimposed to the samples in $x$, whose magnitude relative to the signal $x$ can be accounted for by defining an Intrinsic Signal-to-Noise Ratio (ISNR) $\|x\|_2 / \|v\|_2$.

Since the sparsity basis is a feature of the signal to acquire, the matrix $\Psi$ is known, so that different AICs architectures can be distinguished based on how they build and apply the matrix $\Phi$.

The aim of a proper design of $\Phi$ is to guarantee that the original signal $x = \Psi a$ can be reconstructed starting from the measurements, i.e. that setting $\Theta = \Phi\Psi$ and $\xi = \Phi v$ the equation

$$y = \Theta a + \xi \tag{2.1}$$

can be "solved" for $a$ once that $y$ is known.

In addition to the fact that the matrix $\Theta$ is $M \times N$ with $M \ll N$ as it is assumed in any purely theoretical framework [8,10], finding a solution of (2.1) can in practice additionally complicated by the fact that `i`) $v$ is unknown, `ii`) that $y$ is usually available after its analog-to-digital

conversion and is therefore subject to quantization errors, and iii) that the multiply-and-accumulate operations needed to compute correlations may suffer from limited signal range and thus be corrupted by saturation.

To counter all these problems one classically relies on the a-priori knowledge that $spar(a) \leq K$ [10], though other priors can be added and exploited to increase reconstruction performance, as recently suggested in [15, 16] and as I will describe in Chapters 3, 4 and 5. This is possible since priors help defining a subset of $\mathbb{R}^N$ containing all instances of $x$, so that the acquisition mechanism should map this subset into the measurement space $\mathbb{R}^M$ "quasi-bijectively" in a sense that will be clarified in the following.

More precisely [10], when sparsity is one of the priors, if $\Theta$ can be thought of as a realization of a random matrix with independent entries drawn according to a variety of distributions, then mapping by means of $\Theta$ provides, with high probability, the needed "quasi-bijection". Formally speaking, I say that a matrix $\Theta$ is a *Restricted Isometry* (RI) [10] when there is a constant $0 \leq \delta_K < 1$ such that

$$(1 - \delta_K) \|a\|_2^2 \leq \|\Theta a\|_2^2 \leq (1 + \delta_K) \|a\|_2^2 \tag{2.2}$$

whenever $spar(a) \leq K$. Hence, even if the dimensionality $M$ of the co-domain of a restricted isometry is less than the dimensionality $N$ of its domain, the mapping of $K$-sparse vectors leaves lengths substantially unaltered.

If $\Theta$ is made of independent random entries characterized by a sub-Gaussian distribution [17, 18], then with an overwhelming probability, it is a restricted isometry with a constant $\delta_K$ that decreases as $M$ increases [11].

If $\Theta$ is a restricted isometry, once that $supp(a)$ is known, I may restrict $\Theta$ to that domain and obtain an injective mapping. If the measurements in $y$ additionally encode information on which of the $\binom{N}{K}$ possible supports must be chosen, the overall mapping can be reversed to yield the whole $a$.

This is why a constant ingredient in the recipes for all compressive sensing architectures is *randomness as a mean of capturing information that is known to be sparse*. What is usually done is to overlook the fact that theory puts conditions on the statistical structure of $\Theta$ and design a system in which $\Phi$ is random and (hopefully) transfers its beneficial properties to $\Theta = \Phi\Psi$.

An important side-effect of this assumption (widely verified in practice) is that one does not design the acquisition matrix $\Phi$ depending on the specific $\Psi$ but relies on randomness to implicitly "scan" all possible sparsity bases.

## 2.3 Reconstruction Algorithms

Once that an "encoding" mapping allowing reconstruction has been devised, its "inversion" must be obtained by algorithmic means every time a measurement vector $y$ comes in.

Though reconstruction mechanisms should be designed jointly with the architectures producing the measurements, they are classically addressed as separate components of the overall acquisition system. Their development and analysis is a flourishing field that has recently produced strong and general results and taxonomies [8].

I will here take a very pragmatic approach, concentrate on the most frequently adopted methods, and note that those techniques fall in one of two categories: *optimization-based reconstruction* and *iterative support-guessing reconstruction*.

Both types of technique are commonly devised and set up in the noiseless and idealized case (i.e., for $\xi = \Phi v = 0$) and with neither quantization nor saturation, and are proved (or simply seen) to work in more realistic settings.

The key fact behind optimization-based methods is that, among all the possible counterimages $a$ of the vector $y = \Theta a$ the one that I am looking for is the "most sparse", i.e., the one for which $spar(a)$ is minimum.

Since I usually have $spar(a) \leq K \ll N$ this assumption is sensible. Moreover, it leads to some elegant results on the possibility of recovering $a$ by means of simple optimization problems [10].

More formally, it can be shown that, if $\Theta$ is a restricted isometry with constant $\delta_K \leq \sqrt{2} - 1$ then, for any $\epsilon > 0$, the $\hat{a}$ solution of the optimization problem

$$\begin{aligned} &\min \ \|\hat{a}\|_1 \\ &\text{s.t.} \ \|\Theta\hat{a} - y\|_2 \leq \epsilon \end{aligned} \tag{2.3}$$

is such that

$$\|\hat{a} - a\|_2 \leq C\epsilon$$

for some constant $C > 0$.

Hence, if I use $\epsilon$ to bound the maximum magnitude of the disturbances involved in the measurement process (for instance by setting it proportional to the variance of the noise plus that of the quantization error) I can guarantee

that the reconstruction error vanishes when disturbances go to zero.

Though not impossible, the straightforward application of the above result, depends on a reliable estimation of the parameter $\epsilon$ that quantifies the maximum foreseeable deviation between the unperturbed measurement and its actual value in presence of a mixture of known (e.g., quantization) and unknown (e.g., noise) disturbances.

It is therefore quite common to substitute $\|\Theta\hat{a} - y\|_2 \leq \epsilon$ with $\Theta\hat{a} = y$ by implicitly assuming that the system is working in a relative low-disturbance regime that allows to assume $\epsilon \simeq 0$. Within this approximation, it is convenient to re-express the resulting optimization problem within the framework of linear programming by defining $u = (1, \ldots, 1)^\top$ and by introducing the auxiliary unknown vector $\ell = (\ell_0, \ldots, \ell_{n-1})^\top$ to write

$$\begin{aligned} \min\ & u^\top \ell \\ & \Theta\hat{a} = y \\ \text{s.}t.\ & \ell \geq 0 \\ & -\ell \leq \hat{a} \leq \ell \end{aligned} \tag{2.4}$$

where vector inequalities are meant to hold componentwise.

The equality constraints in (2.4) can be adjusted to cope with specific features of a given architecture or to take into account quantization or saturation.

In particular, due to quantization, I know that the true value of the $j$-th component of $y$ is somewhere in the interval $[y_j - \Delta y_j/2, y_j + \Delta y_j/2]$ with $y_j$ being the value known to the algorithm and $\Delta y_j$ the corresponding quantization step. Hence, in presence of a coarse quantization, it is sensible to substitute the equality constraints $\Theta\hat{a} = y$ in (2.4) with

$y - \Delta y/2 \leq \Theta\hat{a} \leq y + \Delta y/2$, where $\Delta y = (\Delta y_0, \ldots, \Delta y_{m-1})^\top$. Though it surely models the acquisition procedure with greater accuracy, this adjustment does not necessarily lead to improvements and is commonly employed only when one may expect the various $\Delta y_j$ to be substantially different one from the other.

I will indicate the algorithm I use as representative of optimization-based reconstruction as MinL$_1$.

It is interesting to note that MinL$_1$ works without any knowledge of the exact value of $K$ further to that implicit in the number of measurements that must be enough to allow reconstruction. This may be a plus in situations where $K$ cannot be exactly determined in advance. Regrettably, this positive feature is balanced by the fact that, in general, linear programming solution is computationally more expensive that other kinds of iterative reconstruction.

As far an iterative support-guessing reconstruction is concerned, note that, if $supp(a)$ were known I could drop the columns in $\Theta$ that are surely multiplied by 0 and the corresponding entries in $a$ to obtain an $M \times K$ matrix $\Theta_{supp(a)}$ and a $K$-dimensional vector $a_{supp(a)}$ for which $y = \Theta_{supp(a)} a_{supp(a)}$. Since $M > K$, this is an over constrained problem that may be effectively (even "optimally" in case of Gaussian disturbances [19] inverted by using the Moore-Penrose pseudo-inverse $\Theta^\dagger_{supp(a)}$ and computing $a_{supp(a)} = \Theta^\dagger_{supp(a)} y$.

Iterative support-guessing methods are, in general, procedures that alternate a rough, non-necessarily sparse, solution of $y = \Theta a$ from which an estimate of $supp(a)$ is inferred (for example by thresholding on the magnitudes of the components of the temporary solution) that is

then exploited in a pseudo-inverse-based step refining the value.

Though more sophisticated alternatives exists, a reference algorithm within this class is CoSaMP [9] that has some definite advantages. First, it works for matrices $\Theta$ that are restricted isometries and, if $K$ is known and the isometry constant $\delta_{2K}$ for vectors with $2K$ non-zero components can be bounded by $\delta_{2K} \leq c \leq 0.1$, then, given a tolerance $\epsilon > 0$, the reconstructed vector $\hat{a}$ satisfies

$$\|\hat{a} - a\|_2 \leq C \max\left\{\epsilon, \frac{\|a'\|_1}{\sqrt{K}} + \|\Phi v\|_2\right\}$$

where $a'$ is the vector that can be obtained by $a$ by setting to zero its $K/2$ largest entries.

The resulting algorithm is provably fast and, beyond the above formal guarantee on its performance, it is usually extremely stable and effective in recovering the original signal. These favorable properties are paid with the additional assumption that the sparsity of $a$ is known and that the isometry constant $\delta_{2K}$ must be quite low.

In analogy to what happens for optimization-based reconstruction, CoSaMP can be tailored to specific architectures. This can be done, for example, if it is known that errors in the magnitudes of the entries of $a$ are correlated by an implicit filtering in the acquisition scheme. Such an effect can be exploited by inserting a filtering step when passing from support-guessing to pseudo-inversion. More detail on how this may affect performance reconstruction will be given in the following.

## 2.4 Compressive Sensing Encoders

From the previous discussion, I get that to define a compressive sensing system I need to describe two stages

- **encoder**: a hardware system (actually implementing the AIC) and performing some mixed analog-digital operations on the incoming signal to produce a stream of bits. The mixed analog-digital operations are modeled as instance $\Phi$ of a random matrix linking the signal samples to the measurements whose quantization yields the stream of bits transferred from the encoder to the decoder;

- **decoder**: an algorithm that takes the incoming bits and, based on the knowledge of $\Phi$, reconstructs the original signal.

Note that, in a practical implementation, one may not want to communicate $\Phi$ to the decoder and thus most often exploits pseudo-random generators with a common initialization to yield matrices that can be simultaneously known at both stages.

Saturation and quantization are unavoidable in the signal path since the communication between encoder and decoder happens along a digital channel thus implying an ADC block with a finite range (i will assume $[-V^{max}, V^{max}]$ for a certain $V^{max}$) and a finite number of levels.

In the following I will consider the number $B$ of bits generated by the encoder corresponding to the acquisition of the input signal over a given time interval. This can be considered as a "bit budget" since it may be partitioned

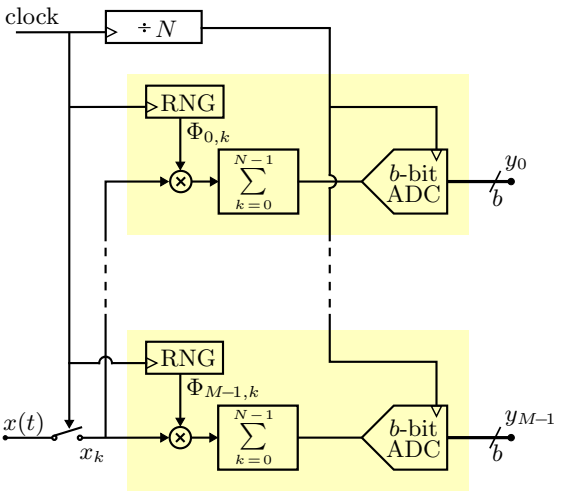



**Figure 2.1:** Block scheme of an RMPI encoder.

into digital words of different depths corresponding to different measurements.

### 2.4.1 Random Modulation Pre-Integration – RMPI

This is the first and probably the most straightforward implementation of an AIC [4].

With reference to Figure 2.1 the samples of the incoming signal $x_k$ are multiplied by the quantities $\Phi_{j,k}$ for a given $j$ and then fed into an accumulation stage to yields the value of the $j$-th measurement $y_j$ that is then quantized by an $b$-bit ADC and aggregated with all the other quantized measurements into the stream of bits that is passed to the decoding stage.

The implementation of the analog blocks preceding the ADC offers several options.

The structure of the multiplier depends on the quantities $\Phi_{j,k}$: some classical approaches adopt Gaussian random variables (Gaussian RMPI) and force the deployment of complete four-quadrant analog multipliers, while more aggressive approaches suggest to constrain $\Phi_{j,k} \in \{-1, +1\}$ (antipodal RMPI) so that multiplication can be implemented by simple switching.

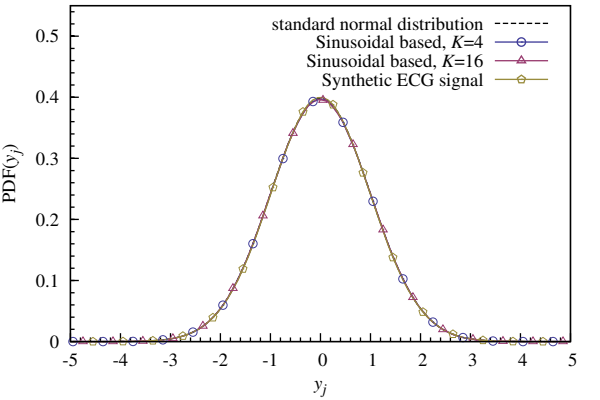
The accumulation stage may be implemented either as a continuous time integrator or as a switched capacitor subcircuit that implicitly matches the discrete-time operation of the multiplier. In any case, the output of the accumulating device will be subject to saturation. Referring to a discrete-time implementation, where $y_j = \sum_{k=0}^{N-1} \Phi_{j,k} x_k$, and relying on the following assumptions: i) the elements of $x$ and $\Phi$ are independent random variables; ii) the elements $\Phi_{j,k}$ of $\Phi$ are *independent* and *identically distributed* (either Gaussian or binary antipodal) random variables, with zero mean and unity variance, i.e., $\mathrm{E}\left[\Phi_{j,k}\right] = 0$ and $\mathrm{E}\left[\Phi_{j,k}^2\right] = 1$; iii) the energy of $x$ in the accumulation time window is unitary, i.e., assuming that the variance of each $x_k$ is the same, $\sum_{k=0}^{N-1} \mathrm{E}[x_k^2] = 1$. Let me consider the random variable $\xi_{j,k} = \sqrt{N}\,\Phi_{j,k}\,x_k$, so that $y_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \xi_{j,k}$. Since the random variables $\xi_{j,k}$ are independent with zero-mean, thanks to i) and ii)), and have unity variance, due to i)-iii), for $y_j$ the hypotheses for applying the central limit theorem are satisfied, so that for large $N$, it converges to the standard normal distributed random variable. Note also that this results is *signal independent*, i.e. it holds no matter the mean value and the variance of the $x_k$.

Furthermore, the speed of convergence is large enough to guarantee that the practical approximation of $y_j$ with a normally distributed random variable is satisfied for reasonable values of $N$. As an example, Figure 2.2(a) reports the observed distribution of the $y_j$ for the antipodal RMPI with $N = 256$ and where the input signals are a 4-sparse and a 16-sparse signal with respect to the Fourier basis, as well as a synthetic ECG signal, generated as in [20]. All signals have been imposed to satisfy the condition $\mathrm{E}\left[x_k^2\right] = 1/N$, so that, according to my model, $\mathrm{E}\left[y_j^2\right] = 1$. In all cases, the observed distributions are superimposing with the standard normal one.

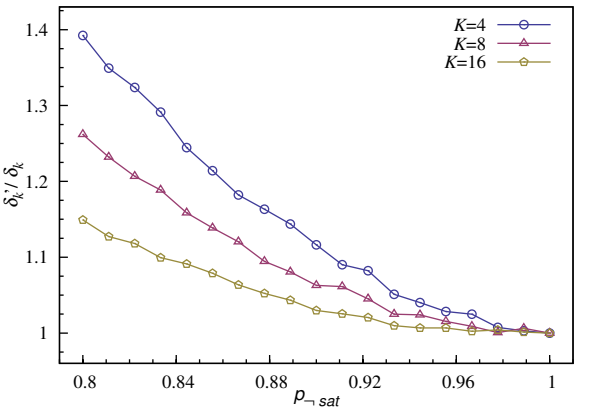
Of course, since a normal distribution is not limited, wherever the input range of the ADC is set, there is an unavoidable non-zero probability that $y_j$ falls out of the ADC conversion range. I refer to this as *static saturation*. Finally, it is immediate to extend the above mathematical model to the generic time step $k$ (instead of considering only the final one at $k = N - 1$), where *dynamic saturation* occurs whenever the intermediate accumulated value is out of the accumulator linear region.

Some theoretical considerations on the effect of static saturation can be found in [21] while a discussion of both static and dynamic saturation in the more realistic model I adopt here has been first proposed in [22]. From the latter, I get in particular that it is sensible to design the accumulation stage so that its saturations coincide with the extrema of the input range of the ADC that I identify with the interval $[-V^{max}, V^{max}]$.

To model the system behavior, I say that, while in the process of computing the $j$-th measurement, the circuit produces the $k$-th intermediate result $y_{j,k} = \sum_{l=0}^{k} \Phi_{j,l} x_l$




(a)




(b)

**Figure 2.2:** (a) Probability distribution density (PDF) of the value of $y_j$ for different input signals in the antipodal RMPI case, obtained with 1000 trials of a system with $M = 96$; (b) Average variation of restricted isometry constant $\delta'_K$ for the sampling matrix $\Phi'$ with respect to the reference value $\delta_K$ given by $\Phi$ in an antipodal RMPI system with $N = 256$, $M = 96$ and $K \in \{4, 8, 16\}$.

when accumulating the $k$-th sample, to finally produce $y_j = y_{j,N-1}$.

Yet, depending on $\Phi$ and on the signal samples, there may be time instants $k$ in which saturation takes place and causes $y_{j,k} \neq y_{j,k-1} + \Phi_{j,\kappa_j} x_k$. If $\kappa_j$ is the first of these instants (i.e., I know that saturation did not affect the system before $\kappa_j$ for the $j$-th measurement) I have that $y_{j,\kappa_j-1} = \sum_{k=0}^{\kappa_j-1} \Phi_{j,k} x_k$ and I can also approximate $y_{j,\kappa_j} \simeq \sum_{k=0}^{\kappa_j} \Phi_{j,k} x_k \simeq \pm V^{max}$, but certainly $y_j = y_{j,N-1} \neq \sum_{k=0}^{N-1} \Phi_{j,k} x_k$. In other words, if saturation takes place, the corresponding measurement is corrupted and, in general, brings little information on the original signal.

An important design parameter for this particular AIC is the probability $p_{\neg sat}$ of no static saturation when dynamic saturation did also previously not occur. From the above consideration, it readily follows that since $y_j$ is normally distributed, then $p_{\neg sat} = 1 - erfc(V^{max})$, where $erfc(\cdot)$ is the complementary error function. Furthermore, despite $p_{\neg sat}$ is formally defined in absence of dynamic saturation, as it will be confirmed in Section 2.6, it can be considered as a sensible indicator of the relative match between the signal actual range and the system saturation levels, even though dynamic saturation may first happen at some $\kappa < N - 1$.

Given the above definitions, the first straightforward design consideration is that, since $N$ is typically in the order of tens if not hundreds, no design strategy may actually aim at obtaining $p_{\neg sat} \simeq 1$ since this would require to increase $V^{max}$ to satisfy $-V^{max} \leq \sum_{k=0}^{N-1} \Phi_{j,k} x_k \leq V^{max}$ for any $N$ and $x$. To cope with this, a fairly simple approach would be what is referred to in [21] as *"democracy"* of the set of measurements, i.e., the fact that, under mild

conditions, one may assume that the information content of each measurement is identical. If this were true, simply discarding saturated measurement would produce a graceful performance degradation since the acquisition system would behave as if it were designed to use a number of measurement equal to the number of the non-saturated ones. Moreover, non-degraded performance could be restored by simply taking further measurement until the original number is reached. Following [22], I will name this approach as RMPI-SPD since it concretizes in Saturated Projection Dropping.

Regrettably, perfect democracy only holds between measurements that are taken as linear combinations of the samples. Saturation acts as a selector discarding those that have a larger value while keeping the smaller ones. From the point of view of the SRN this is clearly not a democratic behavior and causes non-saturated measurements to be less useful than those that have to be dropped and cannot be perfectly replaced by simply trying more measurements.

To cope with this, one may alternatively think [22] of exploiting all the information that is still available as a linear combination of the samples $x_k$, i.e., the fact that, if saturation first occurs at $\kappa_j$, I have $y_{j,\kappa_j} \simeq \sum_{k=0}^{\kappa_j} \Phi_{j,k} x_k \simeq \pm V^{max}$. This allows to replace $\Phi$ with a modified matrix $\Phi'$ in which, the rows corresponding to saturated measurements have zeros in correspondence of time instants after the one in which saturation occurs, i.e. more formally I may set

$$\Phi'_{j,k} = \begin{cases} \Phi_{j,k} & \text{for } k = 0, \ldots, \kappa_j - 1 \\ 0 & \text{for } k = \kappa_j, \ldots, N-1 \end{cases} \tag{2.5}$$

where $\kappa_j = N$ correspond of the case of no dynamic or static saturation, and

$$y'_j = \begin{cases} y_j & \text{if no saturation occurred} \\ V^{max} & \text{if positive saturation occurred} \\ -V^{max} & \text{if negative saturation occurred} \end{cases} \quad (2.6)$$

so that the AIC output can be expressed as $y' = \Phi'\Psi a = \Theta' a$.

Note that this solution makes the matrix $\Phi'$ used for reconstruction a function of the signal samples $x_k$ that caused saturation. Hence, the measurement vector passed to the decoder must contain also the information needed to construct $\Phi'$ from the signal-independent $\Phi$. Assuming that $b$ bits are assigned to the $j$-th component of the measurement vector, I decide to reserve 1 bit to indicate whether saturation happened and use the remaining bits depending on its value: either $b-1$ bits for the quantized encoding of a non-saturated measurement, or 1 bit to discriminate between positive and negative saturation along with $b-2$ bits dedicated to the encoding of $\kappa_j$. I will indicate this approach as RMPI-SPW, standing for Saturated Projection Windowing.

Even if I cannot provide exact results on the RIP for $\Phi'$, it is worth taking a pragmatic look at this issue by comparing the RI constant $\delta'_K$ associated to the $\Phi'$, with the RI constant $\delta_K$ associated to the original matrix $\Phi$ by means of extensive Montecarlo simulations. To do so, I use the approach described in [15] which involves to compute the eigenvalues of the matrices $\Psi^*\Phi^*\Phi\Psi$ and $\Psi^*\Phi'^*\Phi'\Psi$, where $\cdot^*$ stand for transpose conjugate.

Figure 2.2(b) reports the average variation over 500 trials of the ratio $\delta'_K/\delta_K$ as a function of $p_{\neg sat}$ varying from

1 to a value as low as 0.8, when a Fourier basis matrix has been chosen for $\Psi$ and where $N = 256$, $M = 96$ and $K \in \{4, 8, 16\}$, so that the associated $\Phi$ is known to be a RI matrix[2]. As it can be noticed, when $p_{\neg sat} \in [0.8, 1]$, so that the number of entries in the matrix $\Phi'$ which are changed to zero is not too large, $\delta'_K$ smoothly increases with respect to $\delta_K$ in all considered examples. This can be intuitively accepted as a good indication that one can reasonably expect that $\Phi'$ satisfies condition (2.2) when $p_{\neg sat}$ is not far from 1.

I would like to propose here a further observation. It is know that RIP property is only a sufficient condition to CS, but it is not necessary [5]. Once the projection matrix $\Phi'$ is obtained at the decoder, signal reconstruction happens by usual means though guarantees on the resulting $\delta'_K$ are not as strict as for the original $\Phi$.

In fact, the trailing zeros that may appear at certain rows of $\Phi'$ potentially degrade its ability to behave almost as an isometry for $K$-sparse vectors. Yet, as a partial relief from this concern, I know that $\Phi'$ is signal-dependent and I get from (2.5) and from the definition of $\kappa_j$ that when a row of $\Phi'$ contains trailing 0s it is because the projection of the signal along the initial part of the same row was large enough to cause saturation.

Hence it is sensible to assume that, with high probability, most of the energy of the signal had already been extracted by the first part of the row that still enters the reconstruction procedure.

[2]The value $M = 96$, for all considered $K$ values, obeys the theoretical inequality imposed by RI, i.e., $M > CK \log(\frac{N}{K})$ with $C = 4$ [10]

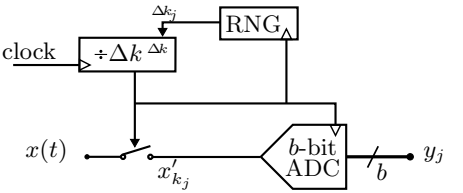



**Figure 2.3:** Block scheme of an random sampling encoder.

### 2.4.2 Random Sampling – RSAM

In classical acquisition systems, samples of the signal are taken regularly on the time axis at a given rate (usually not less than the Nyquist one). AICs relying on random sampling avoid this regularity to produce a number of measurements that, on the average, are less than those produced by Nyquist sampling, while still allowing the reconstruction of the whole signal thanks to sparsity and other priors.

In principle, sampling instants can happen anywhere along the time axis. Yet, a straightforward implementation chooses them among regularly spaced time points that can be selected by digital means. The result is schematized in Figure 2.3 where a backward counter is pseudo-randomly re-loaded each time it reaches zeros, triggering conversion. Grid spacing, and thus clock rate, depends on the resolution with which one wants to place the sampling instants and thus may be expected to be larger that Nyquist rate.

To translate the above block scheme into formulas, say that such the clock identifies a vector $x' = (x'_0, \ldots, x'_{\upsilon-1}, \ldots, x'_{\upsilon(N-1)}, \ldots, x'_{\upsilon N-1})^\top$ that oversamples a

band limited $x(t)$ by a factor $\upsilon$ with respect to $x = (x_0 \ldots, x_{N-1})^\top$ containing the Nyquist samples. The two vectors $x'$ and $x$ are linked by $x' = \Upsilon_{RSAM} x$, being $\Upsilon_{RSAM}$ an upsampling matrix.

With this, the $M \times N$ matrix $\Phi$ is nothing but the product $\Phi = \Sigma\Upsilon_{RSAM}$, where $\Sigma$ is the random sampling matrix defined by the $M$ time instants $k_0 < k_1 < \cdots < k_{M-1}$ at which the counter reaches 0 as in

$$\Sigma_{j,k} = \begin{cases} 1 & \text{if } k = k_j \\ 0 & \text{otherwise} \end{cases}$$

The resulting sampling follows a so-called renewal-process in which all the inter-measurement intervals $\Delta k_j = k_{j+1} - k_j$ are drawn as independent integer random variables exponentially distributed in the interval $[\Delta k_{\min}, \infty]$.

The minimum inter-measurement gap $\Delta k_{\min} \geq 1$ depends on the speed of the ADC, which must be ready for a new conversion each time a measurement is taken so that, by increasing $\Delta k_{\min}$ I loosen the constraints on the ADC implementation. The exponential trend is then tuned to have an average inter-measurement gap equal to $\frac{N}{M}$ so that (at least for large $N$) I expect an average of $M$ measurements.

Each of these measurements is commonly quantized by means of a $b$-bit ADC to yield the bit stream passed to the decoder to yield the most straightforward RSAM-ADC option for an AIC (equivalent to the implementation proposed in [14]).

Yet, one may note that, since the time between two measurements changes, the amount of time potentially available for conversion also changes. This can be

exploited, for example, by deploying a Successive Approximation Register that, relying on an hardware much simpler than that of a full $b$-bit ADC, exploits the available time to increase conversion accuracy.

This translates in a RSAM-SAR system [13] in which each measurement contributes to the bit stream passed to the decoder with a number of bits that increases as the time before the next measurement increases.

Since, in principle, I increase accuracy by one bit for every iteration of the SAR, the $j$-th measurement $y_j$ is quantized with $\Delta k_j$ bits. In this situation, it is sensible to exploit the possibility of embedding quantization into the equations used by minimization-based decoders. This can be done by defining $\Delta y_j = 2V^{max}2^{-\Delta k_j}$, $\Delta y = (\Delta y_0, \ldots, \Delta y_{m-1})^\top$, and, as already highlighted in Section 2.3, use $y - \Delta y/2 \leq \Theta\hat{a} \leq y + \Delta y/2$ instead of equality constraints in laying formulating the minimization problem (2.4).

Though RSAM-ADC and RSAM-SAR treat quantization in a different way, they are both subject only to the static saturation due to the finite input range of the conversion stage. This poses no problem since it can be tackled at design time by simply rescaling the signal input range as in conventional acquisition systems.

### 2.4.3 A note on hardware complexity

It is clear that the above described architectures have different complexities from an hardware point of view. Furthermore, given an architecture, many circuital tradeoffs can be arranged by the designer depending on the target application. Since I want to keep my comparison

at a high abstraction level, I can consider only a generic amount of resources required, which can be used as an indicator either for circuit area, for power consumption, or for both. An additional hardware parameter which needs to be addressed is the design effort required from the designer to implement the AIC. This is an indicator of how many commercially available components or IP blocks can be used to design the converter, which has strong influence on the the design cost of the circuit and on its possible time-to-market.

Most notably, RMPI is the one requiring probably the least established circuitry (continuous-time or discrete-time analog multiply-and-accumulate blocks) and the largest amount of resources (though serial implementations can be conceived, their performance is largely impaired by projecting disjoint time windows of the signals and that option is not considered here).

On the other extreme, RSAM-ADC is surely the least complicated to design and implement since they employ standard blocks (a multi-bit ADC) to which a randomization stage is prepended.

Yet, differences in implementation complexity and cost are difficult to assess without reference to constraints of a specific application and are intentionally left out of the scope of the paper.

What I am doing here is to put the various options at trial in a common arena to highlight respective strengths and weaknesses.

Though it is easily expected (and confirmed) that the strengths-vs-weaknesses balance may favor solutions entailing a larger use of resources, I must leave any final judgment to the application-aware designer that will

be able to manage the trade-off between performance and implementation effort by also addressing issues like analog/digital partitioning, power budget, etc.

## 2.5 Simulation Setting

I simulated the previously described AIC architectures and corresponding decoding stages in a normalized configuration.

Each input $x$ is generated starting from a sparse vector of coefficients $a$ with $N = 256$ entries, only $K$ of which are non-null at the same time, and assuming that the probability of each of the $\binom{N}{K}$ choices of the non-null components is equally probable. These non-null elements are realizations of independent random variables uniformly distributed in the set $[-1, -\frac{1}{2}] \cup [\frac{1}{2}, 1]$. This distribution has been selected in order to avoid components in $supp(a)$ which are too close to zero.

The base matrix $\Psi$ is taken from the family of matrices $\Lambda^{(n,\ell)}$ defined by

$$\Lambda^{(n,\ell)} = \begin{cases} \Gamma^{(n)} & \text{if } \ell = 0 \\ \begin{pmatrix} \Lambda^{(n/2,\ell-1)} & 0 \\ 0 & \Lambda^{(n/2,\ell-1)} \end{pmatrix} & \text{if } \ell > 0 \end{cases}$$

where $\Gamma^{(n)}$ is the $n \times n$ matrix of the orthonormal sinusoidal base for vectors of $n$ samples, i.e.,

$$\Gamma^{(n)}_{j,k} = \frac{1}{\sqrt{n}} \begin{cases} 1 & \text{for } j = 0 \\ \sqrt{2}\cos\left(2\pi\frac{jk}{n}\right) & \text{for } j = 1, \ldots, n/2 - 1 \\ (-1)^k & \text{for } j = n/2 \\ \sqrt{2}\sin\left(2\pi\frac{jk}{n}\right) & \text{for } j = n/2 + 1, \ldots, n - 1 \end{cases}$$

Most of the simulations are done with $\Psi = \Lambda^{(N,0)}$ that is a common choice to present new architectures and compare performance, yet some numerical evidence for $\Psi = \Lambda^{(N,\ell)}$ with $\ell > 0$ is also reported to assess how much each architecture is sensitive to the concentration of energy along the time axis, a phenomenon that is known to impair acquisition based on compressive sensing (see, e.g., [23]).

In order to allow a fair comparison, the sample vector $x = \Psi a$ is normalized coherently with the design needs of each architecture. Roughly speaking, I am assuming that the input signal level is the one which allows the AIC to achieve the best performance. As in conventional ADCs, when the input signal level is too high saturation occurs, conversely, when the level is too low only a small ratio of the dynamic range is used and the quantization noise is increased. Note that this assumption allows me also to avoid the introduction of a new parameter representing the ratio between the optimum and the actual maximum level of the input signal, thus avoiding to unnecessarily complicate notation.

In particular, signals entering the RMPI encoders are scaled to have unitary energy in the integrating window, since this makes unitary also the variance of the $y_j$ and thus helps assessing the probability that such an accumulation saturates the ADC range. Signals fed into an RSAM encoder are scaled to have unit peak amplitude thus matching the ADC conversion range.

Finally, $x$ is perturbed by Gaussian noise with a power controlled by the ISNR parameter, and results for ISNR ranging from 20d$B$ to 60d$B$ are presented.

When a vector $x'$ oversampled by a factor $\upsilon$ is needed,

it is computed as $x' = \Upsilon x$, where

$$\Upsilon_{j,k} = sinc\left(\frac{j}{\upsilon} - k\right)$$

The number of iteration used in CoSaMP was set to 200. This value was chosen since it presents a good compromise between accuracy and execution time in my simulation set-up.

Performance is evaluated by matching the reconstructed vector $\hat{a}$ with the original vector $a$ and using two merit figures: the Probability of Support Reconstruction (PSR) and the Average Reconstruction Signal-to-Noise Ratio (ARSNR), i.e.,

$$PSR = \Pr\left\{supp(a) \subseteq supp_{\min\{a\}/5}(\hat{a})\right\}$$

$$ARSNR(\mathrm{d}B) = \mathbf{E}\left[\mathrm{d}B\left(\frac{\|a\|_2^2}{\|a-\hat{a}\|_2^2}\right)\right] = \mathbf{E}\left[\mathrm{d}B\left(\frac{\|x\|_2^2}{\|x-\hat{x}\|_2^2}\right)\right]$$

where the thresholded support is conventionally defined as

$$supp_{\tau}(a) = \left\{j = 0, \ldots, n-1 | |a_j| \geq \tau\right\}$$

for some suitably small $\tau$. Probabilities and expectations are estimated by Montecarlo simulations.

Given the achieved ARSNR of the AIC, it may be sensible to compare the number of bit $B$ used in the AIC with the number of bit required by a standard PCM coding to achieve an output SNR equal to the ARSNR.

In a PCM coding based on Nyquist rate sampling (i.e., with $N$ samples for each time window) and a total number of bit equal to $B_{PCM}$, I can roughly estimate the output SNR by computing the average number of bits for each sample

$$SNR_{PCM}(\mathrm{d}B) = \frac{B_{PCM}}{N}6.02$$

I define the equivalent compression rate $\gamma$ as the ratio between $B_{PCM}$ and $B$ when ARSNR=$\text{SNR}_{PCM}$, i.e.

$$\gamma = \frac{B_{PCM}}{B} = \frac{\text{ARSNR}(\text{d}B)}{6.02}\frac{N}{B}$$

Clearly, it is convenient to apply CS techniques only if they result in $\gamma \geq 1$ since $\gamma < 1$ implies that the same quality in signal reconstruction could be obtained by a simple PCM entailing a smaller number of total bits. As such, this merit figure is also presented in some of the pictures in Section 2.6.

Montecarlo simulations are taken over respect to 5000 trials.

## 2.6 Numerical Evidence

The standardized environment described above can be employed to test different options for encoding and decoding in a variety of operating conditions as far as sparsity $K$, ISNR, total bit budgets $B$ and its partition into $M$ measurements.

With that, I may explore some areas of the design space of an AIC and highlight some interesting issues and guidelines.

### 2.6.1 Comparison between decoding strategies

A first set of simulations was dedicated to compare performance of the algorithms that I took as representatives of the two different approaches to signal reconstruction, i.e., minimization based procedures (MinL$_1$) and iterative support guessing procedures (CoSaMP).

The support of $a$ is never explicitly computed in minimization-based approaches, which exploit the sparsity promoting features of the $\|\cdot\|_1$ norm to simultaneously obtain an estimate of the support and of the non-null entries.

On the contrary, as already mentioned in Section 2.3, iterative support-guessing methods concentrate on the computation of the support that, once known, is used to transform the under-constrained inversion problem (in which the matrix is $M \times N$ with $M < N$) into an over-constrained problem (in which the matrix is $M \times K$ with $M > K$) that can be pseudo-inverted. Since pseudo-inversion of over-constrained problems is very effective in rejecting small disturbances, if the support is correctly guessed then the final estimation $\hat{a}$ is extremely good.

Figure 2.4 shows how well support reconstruction works for the two algorithms. The PSR is plotted against the number of measurements for different levels of sparsity (i.e., different difficulty of reconstruction). All the trends clearly highlight a minimum value of $M$ under which reconstruction is ineffective. Yet, as $K$ increases, the optimization-based method exhibits better performance since it achieves the same PSR with a smaller number of measurements (and thus of total bits). Note that this happens independently of the encoder and thus may sensibly be ascribed to the decoding strategy.

Further to that, I investigate whether this difference in PSR affects ARSNR and how much the performance of CoSaMP are affected by possible errors on the value of $K$ with which it is made to work. To do so I consider an RSAM encoder providing a variable number $M$ of measurements, each quantized by $b = 8$ bits, of a signal

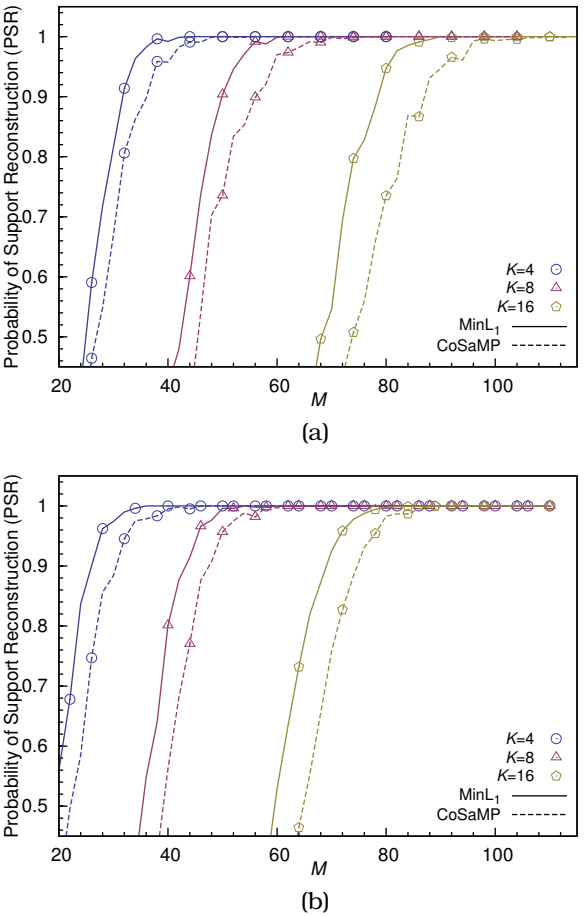



**Figure 2.4:** The PSR achieved by MinL1 and CoSaMP as a function of the number of measurements $M$ for different levels of sparsity $K$ and ISNR=40d$B$ matched by $b = 8$ bits per measure. The bit stream comes: (a) from an RMPI encoder, (b) from an RSAM encoder.

with ISNR= 40d$B$ and $K = 16$.

For each $M$ I run CoSaMP instructing it to assume $spar(\hat{a}) = K = 16$, $spar(\hat{a}) = 1.5K = 24$ and $spar(\hat{a}) = 2K = 32$ (no underestimation of $spar(a)$ is considered since it causes CoSaMP to always perform very poorly).

Figure 2.5-(a) shows the ARSNR when I focus on an RSAM encoder (qualitative trends are identical for the RMPI) and a CoSaMP decoder based on different assumptions for $spar(\hat{a})$. The solid curve shows the ARSNR associated to MinL$_1$ decoding, and can be used as a reference level. CoSaMP is capable to achieve better results in terms of ARSNR with respect to MinL$_1$ in all considered cases, but at same time it suffers form an higher sensitivity in terms of PSR, which is shown in Figure 2.5-(b).

It is interesting to note that the reference MinL$_1$ curve perfectly obeys the theoretical CS framework prediction that the input signal can be recovered from compressed measurement if $M > CK\log_{10}(N/K)$. In fact, if I can consider the signal recovered, once, for example, PSR> 0.99, by setting $C = 4$ as suggested in [10], I get $M > 77.06$ and simulation results from Figure 2.5-(b) confirms that for $M = 77$ the PSR is equal to 0.991.

When $spar(a) = K = 16$, I get from Figure 2.4-(b) that the PSR of CoSaMP is dominated by that of MinL$_1$. Notwithstanding this, a comparison between the two methods in terms of ARSNR is clearly in favor of CoSaMP when its PSR is at least 0.5 (meaning that $supp(a)$ is correctly recovered at least 50% or the times). This is due to the pseudo-inversion step in CoSaMP that is dedicated to (and extremely effective in) computing the values of the non-null components of $\hat{a}$ once that they are the same a

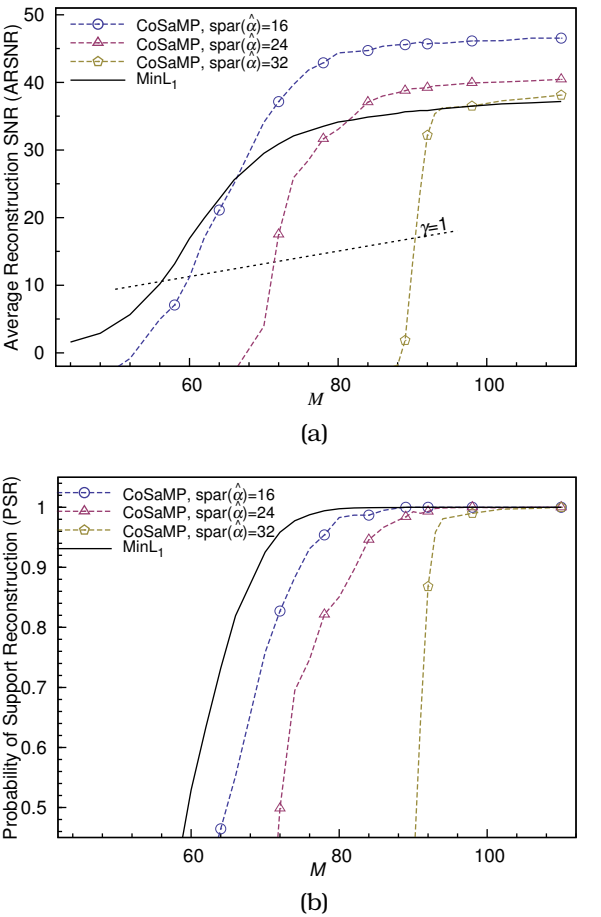



**Figure 2.5:** Performance of CoSaMP as a function of the number of measurements $M$ for $K = 16$ and ISNR=40d$B$ matched by $b = 8$ bits per measure when the bit stream comes from an RSAM encoder and CoSaMP is run with different assumptions on s$par(\hat{a})$. The case of MinL$_1$ decoding is taken as a reference solid curve.

those of $a$.

This property holds until problems in the support-guessing step of CoSaMP prevail. This is evident by matching Figure 2.5-(b) with Figure 2.5-(a) for $spar(\hat{a}) > K$. As an extreme case, when $spar(\hat{a}) = 2K$, MinL$_1$ with no knowledge of $K$ is able to consistently outperform CoSaMP in both PSR and ARSNR.

The evidence I collect pushes towards a general guideline favoring CoSaMP-like methods when support-guessing is relatively easy while suggesting MinL$_1$-like methods when support recovery is the critical issue, for example in biomedical applications.

### 2.6.2 Gaussian or antipodal modulation

The statistics of the random modulating symbols is a degree of freedom in the design of architectures that entail a random modulation, such as RMPI.

Such a degree of freedom may be effectively exploited to match priors further to sparsity that may be available in certain applications (see, e.g., [16]) but also to ease implementation.

In particular, generating, storing and applying to the sample stream a sequence of Gaussian random variables implies an expensive circuitry that can be avoided by resorting to much simpler antipodal random variables taking values in $\{-1, +1\}$.

Actually, in the RMPI case, the distinction between the two modulations appears to be quite artificial if one considers that what can be decided by the implementation is $\Phi$ that enters the definition of $\Theta = \Phi\Psi$. If $\Psi$ is orthonormal and $\Phi$ is made of independent entries, then

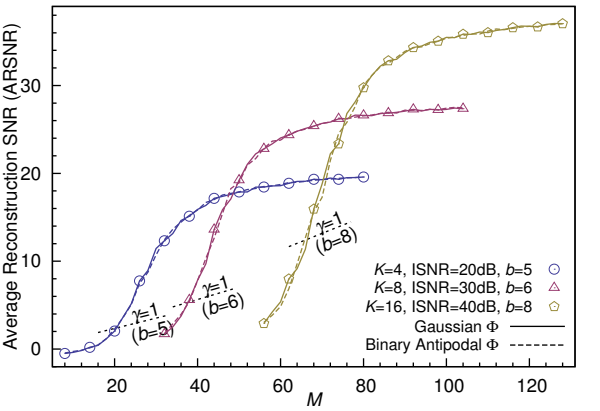



**Figure 2.6:** Performance of RMPI encoder with Gaussian or antipodal modulating symbols plotted against the number of measurements $M$ for different sparsities $K$ and ISNR, each sample is quantized with $b$ bits to approximately match the corresponding ISNR.

$\Theta$ is made of independent rows in which entries (that happens to be linear combinations of independent random variables with a normalized set of coefficients) distribute very similarly to Gaussian random variables. Since this happens independently of the original distribution, Gaussian and antipodal modulations perform in a substantially equivalent way.

Yet, the evidence I collect by simulation indicates that no true difference exists. Figure 2.6 reports the ARSNR obtained when decoding with CoSaMP a bit stream produced either by an RMPI encoder in various operating conditions. It is evident that antipodal modulation is always not worse than Gaussian modulation.

Hence, even if Gaussian modulation is often referred to even in implementation oriented contributions (see, e.g., [24, 25]), any optimized design may safely rely on

the more hardware-friendly (linear feedback shift registers instead of random-number generators, switches instead of multipliers) antipodal modulation and aim for the same final performance.

### 2.6.3 RMPI specific issues

The most distinctive feature of an RMPI architecture is the presence of multiply-and-accumulate blocks deployed to straightforwardly implement CS principles.

The discussion in Section 2.4.1 already highlighted that such operations are saturation-prone since tend to increase the dynamic range of the signals. When saturation happens the information content of the current measurement may be completely lost if no countermeasure is taken.

Actually, once that the input signal range is set (in terms of its average energy, i.e., variance), the probability of saturation is linked to the amplitude of the ADC range. Since the number of bits output by the ADC is finite, reducing its input range not only eases its implementation and paves the way for some resource saving (for example, the increment in the range of the ADC relaxes noise constraints in the analog stages, allowing a reduction of biasing currents) but also decreases the quantization intervals thus increasing resolution. Hence, joint design of saturation and quantization is a key issue for RMPI architectures that has been at least partially addressed, for example, in [22].

Figure 2.7 reports the performance of an RMPI architecture in which the range of the ADC varies thus varying the probability of non-saturation $p_{\neg sat}$ and, implicitly,

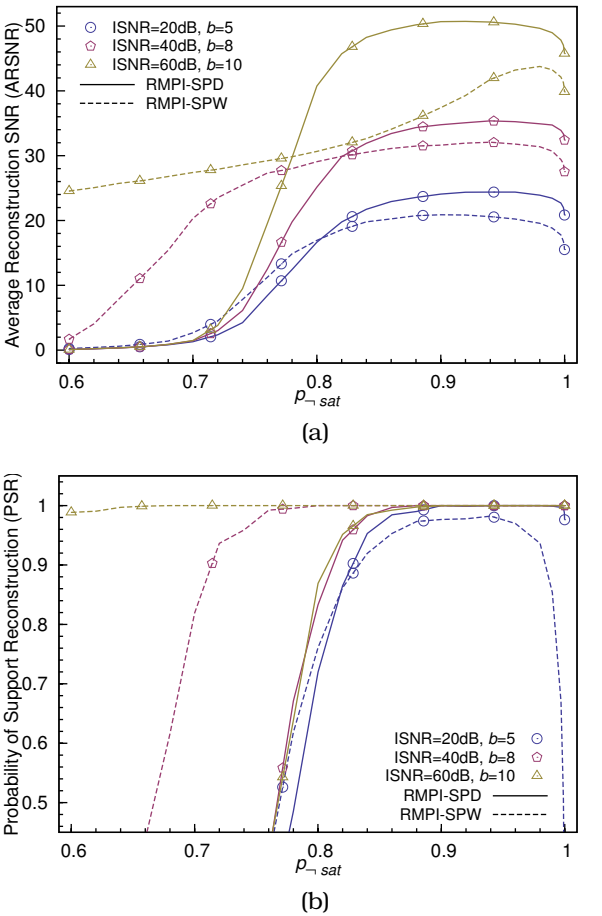



**Figure 2.7:** ARSNR (a) and PSR (b) of an RMPI producing $M = 96$ measurements decoded by CoSaMP for ranges of the ADC such that the probability of non-saturation ($p_{\neg sat}$) is 60% or above and in a signal with $K = 16$ and various operating conditions as far as ISNR and $b$ are concerned. The SPD curves are obtained by dropping the saturated measurements, while the SPW curves are obtained windowing the projections so that a non-saturated value is always obtained.

the resolution with which measurements are quantized. Decoding is performed by CoSaMP.

The SPD and SPW strategies described in Section 2.4.1 are compared when $M = 96$ measurements are taken in a signal with sparsity $K = 16$, and different condition as far as ISNR and bit-per-measurements $b$ are concerned.

Figure 2.7-(b) shows that as $b$ increases, SPW becomes more effective in countering the effect of saturation. In fact, when $b = 5$, SPW does not offer any significant improvement with respect to SPD. This is mainly due to the fact that the 5-bit word encoding each measurement should reserve 1 bit to signal whether a saturation has happened and either 4 bits to encode a non saturated measurement, or 1 bit to distinguish positive from negative saturation and 3 bits to encode the time at which saturation happened. This extremely coarse quantization of times significantly reduces the amount of information that can be squeezed from saturated measurements and makes SPW practically equivalent to SPD.

When $b = 8$ things are different since SPW is able to make both PSR and ARSRN raise to sensible levels for values of $p_{\neg sat}$, for instance in the range $[0.7, 0.8]$, significantly smaller than those needed for SPD, and thus for ADC ranges definitely narrower. With reference to the compression ratio $\gamma$, note that with $N = 256$, $M = 96$ and $b = 8$, CS is a viable opportunity (i.e., enjoys $\gamma > 1$) only if ARSNR$> 18.06$d$B$, a threshold that is achieved for $p_{\neg sat} \geq 0.68$ by SPW, and for $p_{\neg sat} \geq 0.78$ by SPD.

Things get drastically different for $b = 10$ since in this case SPW exploits the relative abundance of bits to ensure that every measurement carry a significant information. In this case, CS is sensible only for ARSNR$> 22.58$d$B$,

a level that can be achieved by SPW for any $p_{\neg sat} \geq 0.6$, while SPD still requires $p_{\neg sat} \geq 0.76$.

Notice also that, as $p_{\neg sat}$ increases (due to the fact that the ADC range is widened) SPW follows a trend that is eventually dominated by SPD due to the fact that even non-saturated measurement are encoded by $b - 1$ bits instead of $b$ bits.

In all cases, performance decreases as $p_{\neg sat} \to 1$. This is due to the trade-off between non-saturation and fine quantization. In fact, to increase $p_{\neg sat}$ one must enlarge the ADC range that eventually becomes so large that the finite number of bits spent to quantize it produces a too coarse subdivision.

### 2.6.4 RSAM specific issues

In principle, the random delay that RSAM architectures insert between measurements may be thought as a real quantity. Hence, the grid on which, for the convenience of the implementation, the $\Delta k_j$ are taken is potentially a design parameter.

Given this temporal grid and indicating with $\upsilon$ its oversampling ratio with respect to the entries of the signal vector $x$, the other key design parameter is $\Delta k_{\min}/\upsilon$. In fact, since no two points on the oversampled grids whose distance is less than $\Delta k_{\min}$ are taken as sampling instants, $\Delta k_{\min}/\upsilon$ defines how slower the ADC in an RSAM architecture can be with respect to one that must convert all the components of $x$ at their original rate.

Figure 2.8 reports the ARSN of an RSAM bit stream decoded by $\mathtt{MinL}_1$ for different oversampling $\upsilon$ and $\Delta k_{\min}/\upsilon$ where the encoded signals have ISNR= 20 and $K = 4$. The

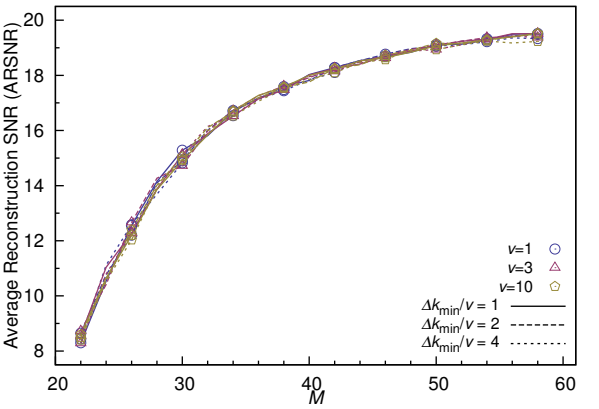



**Figure 2.8:** ARSNR of an RSAM-ADC achieved using $\mathrm{MinL_1}$ for different oversampling ratios $\upsilon$ and related different $\Delta k_{\min}$ on signals with $K = 4$ and ISNR= 20.

same profile can be observed changing either the decoding strategies or $K$ and ISNR.

The fact that performance appears to be invariant with respect to both parameters suggests that oversampling is actually unnecessary even if samples are not equally spaced in time and that a conspicuous relaxation of the requirements put on the ADC can be obtained until $\Delta k_{\min}/\upsilon$ is not too close to $N/M$.

### 2.6.5 Bit budget management

When the total bit budget $B$ is assigned, RMPI and RSAM are able to trade the number of measurements $M$ for the depth $b$ of the digital word encoding each of them, thus allowing some performance optimization.

Actually, the $M$-vs-$b$ trade-off may be seen as a further incarnation of an abstract trade-off showing itself in many

fields of information processing, like the exploration-vs-exploitation trade-off in nonlinear optimization strategies.

The key factor here is the ISNR since it can be intuitively accepted that there is little advantage in allocating bits that are then spent to encode the value of a random perturbation.

If this may yield a clear guideline in straightforward acquisition schemes, CS architectures manipulate the signal in different ways before quantization takes place, and the optimum granularity of the ADC may be less easy to anticipate.

Figure 2.9 shows what happens to the performance of an RMPI architecture when a certain bit budget is partitioned by using a different number of bits for each measure.

Figure 2.10 reports the same data for RSAM-ADC architecture.

Note that all PSR plots are non-increasing in $b$, indicating that support reconstruction is easier when the number of measurements $M = B/b$ is high, i.e., when most resources are devoted to "exploration" to cope with the intrinsically combinatorial problem of support finding.

It is then natural to see that, for the largest possible value of $b$ that maintains such an "exploration" highly effective, i.e., for the rightmost point of the PSR curves that still yields a PSR value close to 1, all ARSNR curves exhibit a maximum. That is the point at which the maximum effectiveness of the "exploration" pairs with the highest possible resolution of the measurement, i.e., with an intense "exploitation" of the available data.

As a first intuition, one may expect that the trade-off between $b$ and $M$ given $B$ which optimizes the ARSNR is

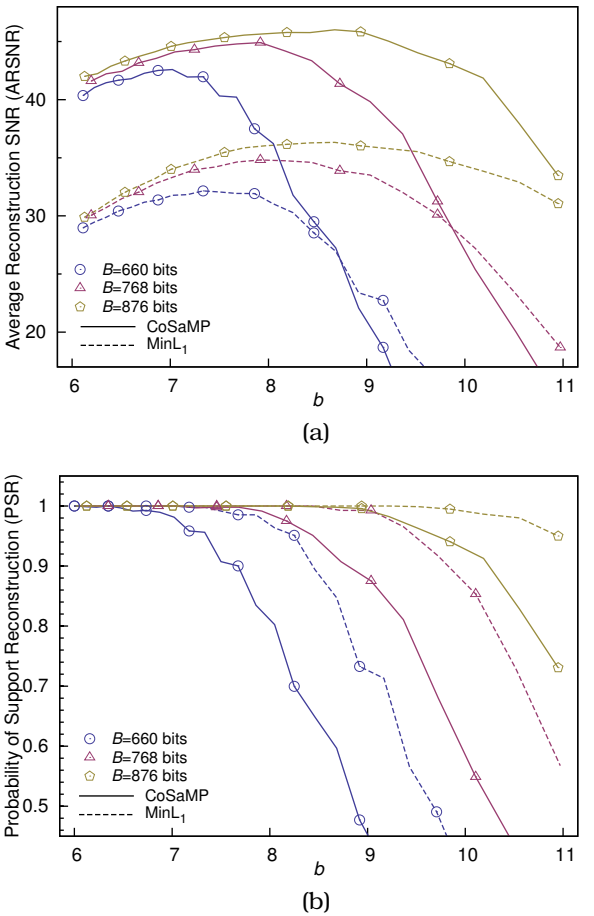



**Figure 2.9:** Performance of an RMPI architecture in terms of (a) ARSNR and (b) PSR plotted against the number of bits $b$ used to encode each measurement when the total number of bits is constrained by the bit budget $B$. In this case $K = 16$, ISNR=40d$B$, and $p_{\neg sat} = 0.9$.

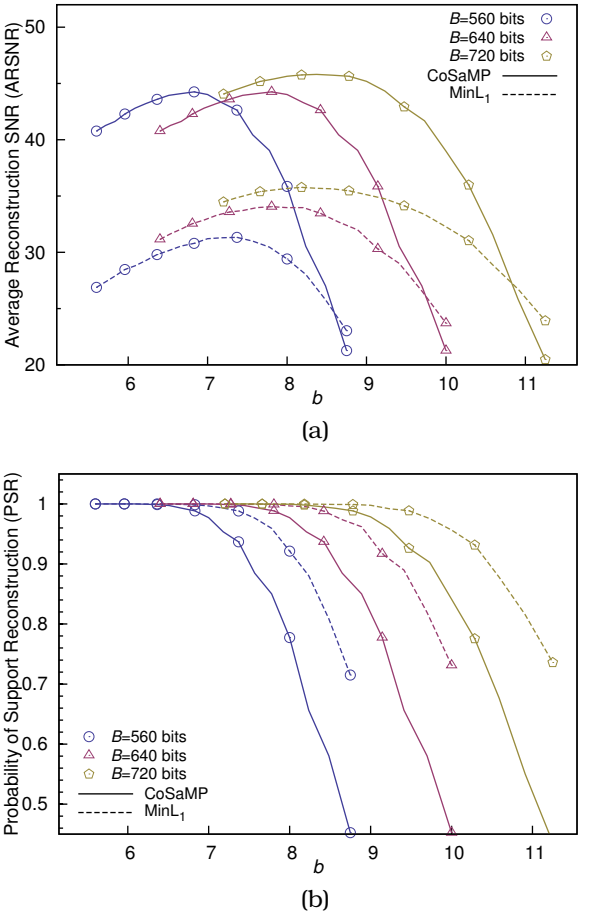



**Figure 2.10:** Performance of an RSAM-ADC architecture in terms of (a) ARSNR and (b) PSR plotted against the number of bits $b$ used to encode each measurement when the total number of bits is constrained by the bit budget $B$. In this case $K = 16$ and ISNR=40d$B$.

expected to be around $b$=ISNR/6.02, since the SRN of the measurements is equal to the ISNR. This is immediate for the RSAM-ADC; it can also be seen for the RMPI by applying the central limit theorem as in Section 2.4.1 to the noisy input signal as defined in Section 2.2. Yet, performance is influenced by many other parameters, which may move the optimum far from this expected point.

By looking at the numerical results in Figures 2.9 and 2.10 I can see how the optimum point it is affected by $B$. When $B$ is small it may be sensible to reduce $b$ and increase $M$ in order to better satisfies some CS constraints such as the restricted isometry property. On the contrary, when $B$ is large, one could get more advantages by increasing $b$ with respect to increasing $M$, since when the number of measurement is large enough to allow signal reconstruction it is certainly preferable increasing measurements quality instead of their quantity.

As a final remark, recall that while RSAM-ADC allows to optimize the trade off between $M$ and $b$, RSAM-SAR automatically assigns a different number of bits to every measurement.

To know whether this automatic assignment is beneficial to the general performance, or must be only suffered as a side-effect of a lighter architecture (a SAR compared to a full ADC) I compare RSAM-ADC and RSAM-SAR in Figure 2.11.

The performance of RSAM-ADC always dominates that of RSAM-SAR whose automatic bit allocation does not help reconstruction, especially for large ISNR values and a small $\Delta k_{min}$ that sets the minimum for the number of bits of each measurement. Note that the performance

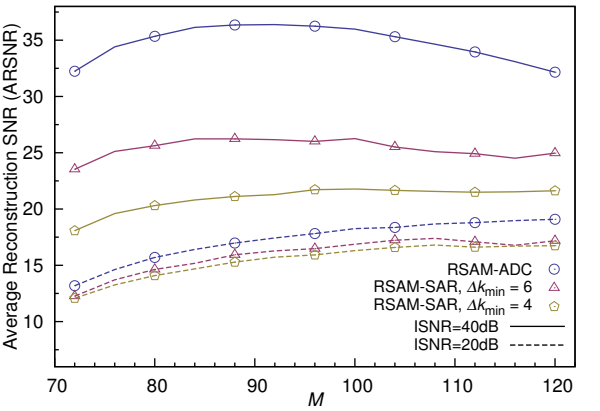



**Figure 2.11:** Performance of an RSAM-ADC and of an RSAM-SAR architecture in terms of ARSNR plotted against the number of measurements for different ISNR, with sparsity $K = 16$ and a the total bit budget $B = 768$. The number of bits for each measurement for the RSAM-SAR depends on the random inter-measurement interval and on $B$.

of the RSAM-SAR are improved by increasing $\Delta k_{min}$, thus pushing bit allocation towards the uniformity that characterizes the RSAM-ADC option.

For small values of ISNR the performance difference is strongly reduced. In the specific case of the figure $\Delta k_{min}$ is large enough to ensure a quantization noise smaller with respect to the ISNR, and both architectures reach an ARSNR value similar to the ISNR.

As a conclusion, the trade-off between $M$ and $b$ is a complex issue, and the optimum point may not be coincident with what is suggested by intuition, thus hinting the fact that the signal processing before quantization affects the trade-off. Furthermore, this optimization is strongly depended on the ISNR, making a signal independent analysis impossible.

### 2.6.6 General comparison and discussion

As a final step I compare the considered architectures to establish what is the best performing in a certain number of operative conditions.

Results are reported in Figure 2.12 for the case in which $K = 16$ and ISNR=40d$B$ assuming that $K$ is also known at the decoder so that CoSaMP is the best decoding choice.

For each bit budget $B$, I plot the maximum ARSNR that can be obtained by (numerically) optimizing $M$ (and thus $b = B/M$) for RMPI and for RSAM-ADC. RMPI is designed so that $p_{\neg sat} = 0.9$. This plot allows me to see which of the architectures achieve the best compression *vs* reconstruction quality trade-off.

By comparison with the $\gamma = 1$ line, it is evident that all considered architectures makes AIC convenient with respect to straightforward acquisition. It is also clear that performance increases as the resources spent in acquisition increase thus allowing an informed trade-off once that application specific constraints are known.

Though these performance evaluation are significative, one of the issues that is not yet clarified is how much they are "robust" with respect to the basis along which the original signal is sparse.

It can be intuitively accepted that the sinusoidal basis I use is a good representative of bases whose waveforms cover most of of the time interval in which the signal is observed. Yet, it is well known [23] that if the waveforms of the basis exhibits some kind of concentration along the time axis, then compressive sensing becomes increasingly more difficult (more formally, it requires an increasing

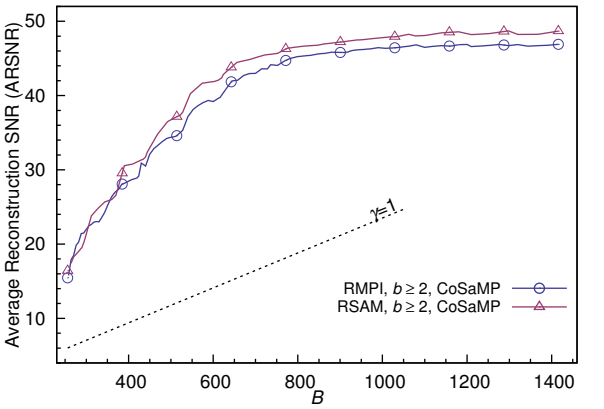



**Figure 2.12:** Largest achievable ARSNR for each architecture plotted against the total bit budget $B$.

number of measurements).

The formal argument quantifying the above fact can be paralleled by an independent intuitive argument. Despite the fact that for abstract signal processing basis choice is arbitrary, all operations eventually get implemented by time-domain operations. If these operations are not time-invariant then the amount of energy collected by the system from the incoming signals can vary depending on its support.

From this I get that strategies like RSAM are less robust with respect to energy concentration along the time axis since they heavily rely on non-time-invariant processing: RSAM samples at single points ignoring the signal even in a strict neighborhood of the sampling instant.

This is fully confirmed by the numerical evidence reported in Figure 2.13 in which I adopt $\Psi = \Lambda^{(256,\ell)}$ for $\ell = 0, 1, 2, 3, 4$. This produces bases whose waveforms

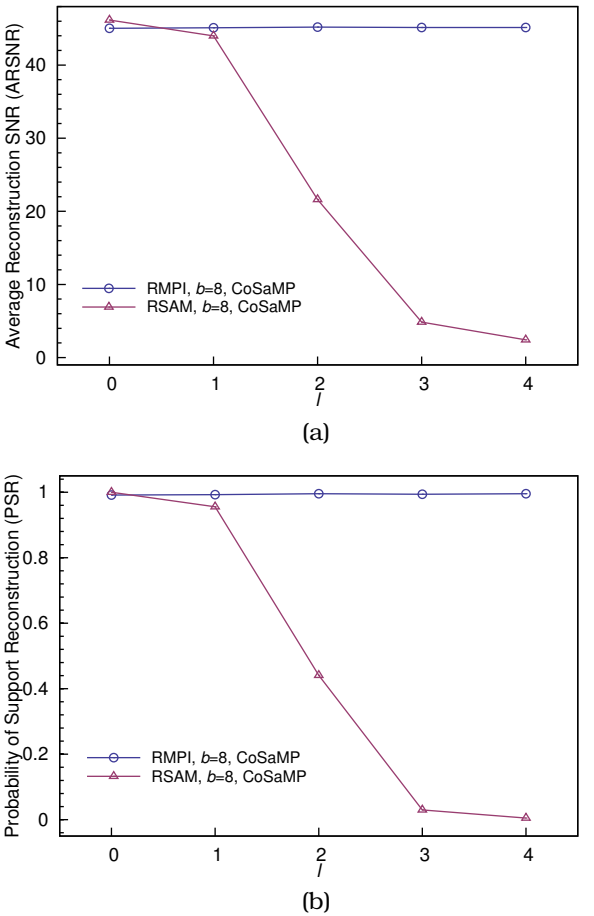



**Figure 2.13:** Performance of RMPI and RSAM-ADC when acquiring a signal characterized by ISNR=40d$B$ with $M = 96$ measurements quantized with $B = 768$ bits. The signal is sparse with respect to a basis whose waveforms have a support covering a fraction of the acquisition interval equal to $2^{\ell}$ for $\ell = 0, 1, \ldots, 4$.

| | RMPI | RSAM |
|---|---|---|
| Performance: | | |
| • compression *vs* quality | + | + |
| • sparsity base robustness | ++ | − |
| Hardware complexity: | | |
| • resources required | −− | + |
| • design effort | − | + |

**TABLE 2.1:** Summary of advantages and disadvantages of the different architecture analyzed in this paper, both from the signal processing point of view and from hardware complexity point of view.

cover $2^\ell$ of the total time range of $N = 256$ instants.

Simple visual inspection reveals that RMPI-based processing is completely insensitive to the support of the basis waveforms while the performance of non-time-invariant methods such as RSAM suffer a consistent performance decrease as that support decreases.

Obviously, this is a key feature to keep in mind when selecting an AIC architecture for a particular application since, for example, suggests the adoption of RMPI-based strategies for signals like ECGs (that are usually decomposed along Gabor atoms whose support is only a fraction of, say, an heartbeat time) [24–26], while allowing an extremely parsimonious acquisition of EEGs whose sparsity can be revealed in close-to-fourier bases [27].

## 2.7 Conclusion

I took some steps in clarifying relative strengths and weaknesses of some of the recently proposed architectures for hardware implementation of AICs based on CS.

Though theoretical support to CS techniques is in

rapid evolution and gives some fundamental guarantees on the fact that CS is feasible and may be advantageous over conventional acquisition strategies, coping with the most basic nonlinearity unavoidable in any real implementation (i.e., saturations and quantization of the measurements) and yield fair quantitative comparison required extensive simulations in an homogeneous environment.

Based on numerical evidence, I were able to discuss some typical issues in AIC design, such as architecture selection for a particular operative setting, or the partitioning of a total bit budget into digital words corresponding to individual measurements. Pros and cons of each architecture have been discussed, evaluating both the hardware complexity of all solutions (see Section 2.4.3) and the performance in terms of ARSNR and PSR (see Section 2.6.6). This comparison has been briefly summarized in Table 2.1.

I were also able to at least envision some more general guidelines such as, for example, that the RMPI architecture and its recently proposed adjustments are probably the most versatile approach though not always the most economic to implement.

CHAPTER

# 3 Energy driven CS and Rakeness

In this chapter I will show you a different point of view on sampling sequences characterization. The leveraging on sparsity has been recently paired [15] with another technique widely used by engineers to spot information content in signals, i.e. the uneven distribution of average energy along properly defined bases (that, in general, are different from those for which sparsity can be identified) [1]

## 3.1 CS for localized signals

In the following, I will indicate such an uneven distribution of energy with the term *localization* and I will observe that, in general, it provides a different a-priori information with respect to sparsity. As it is consequently naturally to expect, I will be able to show that these signal features allows improved sensing operations.

[1] A typical example is the class of band-pass signals, which are localized in the frequency domain, i.e., with respect to the Fourier basis.

The key assumptions under which this may happen are that (i) measurements are taken by projecting the signal onto a proper set of waveforms whose cardinality is smaller than the dimensionality of the signal, and (ii) the overall effect of disturbances in the sensing process (thermal noise, quantization errors, etc.) can be modeled as a projection-independent error.

When (i) and (ii) hold, a noise-tolerant reconstruction of the sparse signal from a number of measurements that is smaller than its dimensionality is commonly achieved by designing the projection operator so that it is a *restricted isometry* (RI), i.e. it approximately preserves the length of the sparse signal to which it is applied so that the ratio between the norm of such signal and that of its projection falls within an interval $[\sqrt{1-\delta}, \sqrt{1+\delta}]$ where the RI constant $0 \leq \delta \leq 1$ should be as small as possible [5] [6].

Roughly speaking this means that, if the measurements come from a RI, the original signal energy is not lost in the projection and, when acquisition error is added, the signal-to-noise ratio (SNR) of the samples remains high enough to perform reconstruction.

This approach and its pairing with localization can be intuitively explained with reference to a simplified, low-dimensionality setting in which the signal to acquire $a$ has three components $(a_0, a_1, a_2)$ and is sparse since only one of its components is non-vanishing in each realization. More formally I may assume that $a_j \neq 0$ with probability $p_j$ for $j = 0, 1, 2$ and, obviously, for $p_0 + p_1 + p_2 = 1$. Furthermore, when it is non-zero, the $j$-th component of the signal is a realization of a random variable with variance $\sigma_j^2$ for $j = 0, 1, 2$.

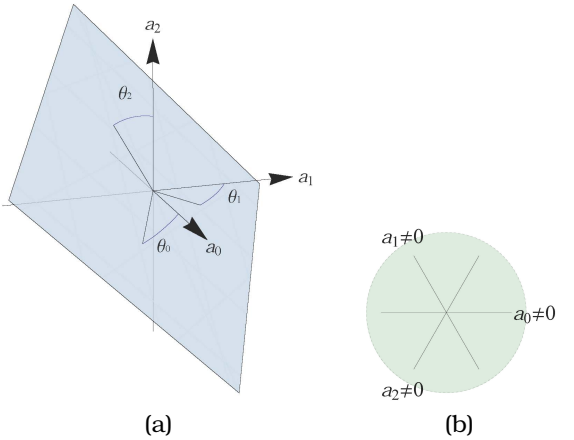



**Figure 3.1:** A simple CS task using a projection plane designed by considering only the restricted isometry property (a). A graphical evaluation of the corresponding restricted isometry constant (b).

It is worth stressing that the possibility of $p_0\sigma_0^2 \neq p_1\sigma_1^2 \neq p_2\sigma_2^2$ implies that localization and sparsity are two separate concepts. In fact, though $a$ is sparse by construction, its average energy concentrates on the axis whose associated $p_j\sigma_j^2$ is larger and this concentration depends on the unbalance between the probability-variances products.

Since $a$ is sparse, it can be reconstructed by measuring its projection on a two-dimensional plane. To define it, refer to Figure 3.1-(a) and note that the generic projection plane passing through the origin defines an angle $\vartheta_j \in [0, \pi[$ with each axis $a_j$ for $j = 0, 1, 2$. These angles are such that $\sum_{j=0}^{2} \sin^2(\vartheta_j) = 1$.

Any set of angles $\vartheta_j \neq \pi/2$ for $j = 0, 1, 2$ is a feasible choice. This is shown in Figure 3.1-(b) that reports a

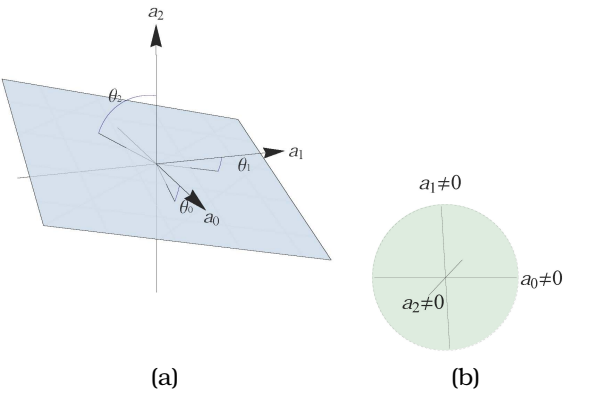



**Figure 3.2:** A simple CS task using an optimized projection plane designed by merging rakeness and restricted isometry (a). A graphical evaluation of the corresponding restricted isometry constant (b).

unit radius circle on the projecting plane, along with the projections of unit-length segments centered in the origin and aligned with each of the three axis.

As long as the three projections have no point in common but the origin (so that, in general, the projections of the coordinate axes are distinct straight lines) the retrieval of the original signal in noiseless conditions can be ensured without complicated algorithms.

When noise comes into play, classical CS theory looks for planes corresponding to projection operators that are good RI. To do so, note that the ratio between the length of a segment aligned with the axis $a_j$ and that of its projection on the plane is $\cos(\partial_j)$. Hence, to minimize the RI constant I should choose each $\cos(\partial_j)$ as close as possible to 1, i.e. $\partial_0 = \partial_1 = \partial_2 = \sin^{-1}\sqrt{\frac{1}{3}}$ that is actually the case reported

in Figure 3.1.

This choice clearly disregards the actual values of the probabilities $p_j$ and signal powers $\sigma_j^2$ for $j = 0, 1, 2$ and may be suboptimal.

To take these further information into account note that, since disturbances are introduced in acquiring the projections, they are independent from the plane. Hence, I may improve the SNR by choosing a plane that is able to *rake* a larger fraction of the signal power. I call this property *rakeness* and, in this case, to maximize it I have to maximize the power of the projection $\sigma^2 = p_0\sigma_0^2\cos^2(\vartheta_0) + p_1\sigma_1^2\cos^2(\vartheta_1) + p_2\sigma_2^2\cos^2(\vartheta_2)$.

With my assumption and by setting $\xi_j = \cos^2(\vartheta_j)$ for $j = 0, 1, 2$ this amounts to maximizing $\sigma^2 = p_0\sigma_0^2\xi_0 + p_1\sigma_1^2\xi_1 + p_2\sigma_2^2\xi_2$, subject to the constraint on the $\vartheta_j$ that becomes $\xi_0 + \xi_1 + \xi_2 = 2$. Assuming that $p_0\sigma_0^2 > p_1\sigma_1^2 > p_2\sigma_2^2$, this criterion leads to $\xi_0 = \xi_1 = 1$ and $\xi_2 = 0$, i.e., a projections plane that coincides with the coordinate plane spanned by $a_0$ and $a_1$.

Clearly, the sheer maximization of the rakeness is not acceptable since any realization of $a$ in which $a_2 \neq 0$ would not be captured by the system or, in terms of the RI property, $\delta = 1$ since the $a_2$ axis belongs to the null-space of the projection operator.

This toy case highlights that RI and rakeness may be suboptimal as a design criterion when considered alone and that improvements may be sought addressing the trade-off between RI enforcement and rakeness maximization.

Such a trade off can be addressed both in a deterministic and in a statistical way.

Pursuing the deterministic path, one may choose a

projecting plane like the one in Figure 3.2-(a) that still allows signals along $a_2$ to have a non-zero projection but clearly favors directions with the largest expected power. Figure 3.2-(b), that is analogous to Figure 3.1-(b) for the new plane choice, shows that this is detrimental in terms of the RI constant since the length of the projection of the segment along $a_2$ is substantially reduced. Yet, the lengths of the projections of the segments along $a_0$ and $a_1$ are increased and since these are the occurrences carrying more power on the average, the overall average acquisition quality may be improved.

The same improvement may be pursued in statistical terms by assuming that the projecting plane is chosen randomly at each measurement. In this case, the statistic of plane choices can be biased so that planes collecting larger energy are more probable, but planes allowing the acquisition of less important components are still possible.

This second setting is particularly interesting since random projections are already employed to guarantee good RI properties [17] and the main aim of this contribution is to show that the trade-off between RI and rakeness can be addressed by proper design of the statistical distribution of the projecting directions.

The rest of the chapter is organized as follows. Section 3.2 will define the conversion architecture and lay down its mathematical model. Section 3.3 introduces more formally the rakeness and its use as a design criterion. In doing this, to focus this exposition on application-oriented considerations, I accept that maximizing the energy of acquisitions is the right direction to go, thus postponing the statement of the formal chain of results

starting from a mathematical definition of localization to a future contribution. This accepted, Section 3.4 describes a design path addressing the RI/rakeness trade-off when the signals to acquire are localized in the frequency domain, that is by far the most common domain for signal analysis. Section 3.5 expands that view to include a generic adaptive domain, that is able to reveal localization in a large class of signals. A couple of lengthy derivations are reported in the 3.6.

## 3.2 System Definition

I will concentrate on systems that perform AI conversion of sparse and localized signals by means of Random Modulation Pre-Integration (RMPI) [4].

This scheme sketched in Figure 3.3 acts on signals of the kind $a(t)$ where $t$ is most usually time but may also be any other indexing variable.

A "slice" of the signal $a$ (say for $-T/2 \leq t \leq T/2$ for some $T > 0$) is processed by multiplying it by a waveform $b(t)$ with a correspondingly sized support. The waveform $b(t)$ is made by amplitude modulated pulses whose modulating symbols are chosen from a certain set.

The most hardware friendly choices are rectangular pulses with antipodal ($\{-1, +1\}$) symbols since multiplication can be implemented by a simple arrangement of switches that nicely embeds, for example, into switched-capacitor implementations [28]. I had also shown you that antipodal symbols in place of real Gaussian values do not produce disadvantages in the decoding stage (Subsection 2.6.2).

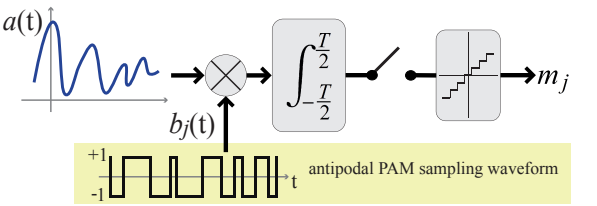



**Figure 3.3:** Block diagram of RMPI architecture: the signal to acquire is multiplied by the $j$-th antipodal PAM waveform and fed into an integrator whose output is sampled and quantized to produce the digital conversion of the $j$-th measurement

The resulting waveform is then integrated or low-pass filtered to obtain a single value that is converted into a digital word by conventional means. Note that this further step also nicely fits into a switched-capacitor implementation that naturally manages charge integration.

Despite the fact that the rate (for time-indexed signals) or density (for generic signals) of the pulses may be very high and even larger than what a Nyquist-obeying acquisition would require, only the integrated values are actually converted.

This multiply-and-integrate operation materializes the scalar product $\langle a(t), b(t) \rangle$ and can be performed $M$ times (either serially or in parallel), each time considering a different waveform $b_j(t)$ ($j = 0, \ldots, M-1$) that is characterized by a set of modulating symbols drawn at random with a certain statistic.

The resulting projections $m_j = \left\langle a(t), b_j(t) \right\rangle$ for $j = 0, \ldots, M-1$ can be aligned in a measurement vector $\underline{m} = (m_0, \ldots, m_{M-1})^\top$. Figure 3.3 exemplifies the signals and the operations entailed by the acquisition of $m_j$ using antipodal PAM waveforms.

For what concerns signal reconstruction, I assume that $a$ is $K$-sparse, i.e., that there is a collection of $N$ waveforms $u_j(t)$ for $j = 0, \ldots, N-1$ such that every realization of $a$ can be written as

$$a(t) = \sum_{j=0}^{N-1} a_j u_j(t) \tag{3.1}$$

for certain coefficients $a_j$ such that at most $K < N$ of them can be non-zero at any time.

Plugging (3.1) into the definition of $m_j$ I get $m_j = \sum_{k=0}^{N-1} a_k \left\langle u_k(t), b_j(t) \right\rangle$. By defining the vector $\underline{a} = (a_0, \ldots, a_{N-1})^\top$, the $M \times N$ projection matrix $\underline{P} = [\underline{P}_{j,k}] = [\left\langle u_k(t), b_j(t) \right\rangle]$, and the vector $\underline{v} = (v_0, \ldots, v_{M-1})^\top$ accounting for the total noise affecting the projections, I have that

$$\underline{m} = \underline{P}\,\underline{a} + \underline{v} \tag{3.2}$$

is the reconstruction equality to be solved for the unknown $\underline{a}$ with the aid of its $K$-sparsity. In principle, this could be done by selecting, among all the vectors $\underline{a}$ satisfying (3.2), the one with the minimum number of non-zero entries. Since this is, in general, a problem subject to combinatorial explosion, many alternative theoretical and algorithmic methods have been developed allowing efficient and effective reconstructions [29] [12] [30]. Among all these possibilities, I will exploit the algorithm described in [12] in my experiments in Chapter 4.

Note that $\underline{v}$ takes into account at least the intrinsic thermal noise affecting the analog processing of $a(t)$ and the quantization noise due to digitalization. Since thermal noise is additive white and Gaussian (AWGN), its contribution to $\underline{v}$ is independent of the projecting waveforms $b_j(t)$ as long as they have constant energy.

I assume that quantization noise is also approximately white and independent on the quantized input so that condition (ii) discussed in Section 3.1 is satisfied.

## 3.3 RI and Rakeness

To cope with the noise in $\underline{v}$, RI-based design [10] tries to make the RI constant $\delta$ of the projection operator as low as possible. This can be checked directly from the matrix $\underline{P}$. In fact, since the projection is applied to $K$-sparse vectors, I should consider each of the $\binom{N}{K}$ matrices $\underline{P}'$ that are built selecting $K$ of the $N$ columns of $\underline{P}$. If $\lambda_{\underline{P}'}^{min}$ and $\lambda_{\underline{P}'}^{max}$ are respectively the minimum and maximum among the singular values [31] of $\underline{P}'$ I have

$$\delta = \max_{\underline{P}'} \left\{ \max \left[ 1 - \lambda_{\underline{P}'}^{min}, \lambda_{\underline{P}'}^{max} - 1 \right] \right\}$$

To go further, I define the average rakeness $\rho$ between any two processes $a$ and $\beta$ as

$$\rho(a, \beta) = \kappa_\rho \mathbf{E}_{a,\beta} \left[ |\langle a, \beta \rangle|^2 \right] \tag{3.3}$$

where the constant $\kappa_\rho$ is used to switch the meaning of $\rho$ from "average *energy* of projections" ($\kappa_\rho = 1$) to "average *power* of projections" (e.g., $\kappa_\rho = T^{-1}$ for signals observed in $[-T/2, T/2]$).

It is worthwhile to highlight that $\rho(a, \beta)$ depends on how the second-order features of the two processes combine. In fact, I may expand the definition as

$$\begin{aligned}
&\rho(a,\beta) = \\
&= \kappa_\rho \mathbf{E}_{a,\beta}\left[\int_{-T/2}^{T/2}\int_{-T/2}^{T/2} a^*(t)\beta(t)a(s)\beta^*(s)\mathrm{d}t\mathrm{d}s\right] \\
&= \kappa_\rho \int_{-T/2}^{T/2}\int_{-T/2}^{T/2} \mathbf{E}_a\left[a^*(t)a(s)\right]\mathbf{E}_\beta\left[\beta(t)\beta^*(s)\right]\mathrm{d}t\mathrm{d}s \\
&= \kappa_\rho \int_{-T/2}^{T/2}\int_{-T/2}^{T/2} C_a(t,s)C_\beta^*(t,s)\mathrm{d}t\mathrm{d}s \qquad (3.4)
\end{aligned}$$

where $\cdot^*$ stands for complex conjugation and the two correlation functions $C_a$ and $C_\beta$ have been implicitly defined.

From the toy example in Section 3.1, I know that choosing the process $b$ that maximizes the rakeness $\rho(a,b)$ leads to good average SNR of the projections, but may destroy the RI property making the system insensitive to some signal components.

To counter this *over-tuning* effect one may require that the process $b$ is "random enough" to assign a non-zero probability to realizations that, despite being sub-optimal from the point of view of energy collection, allow the detection of components of the original signal that would be overlooked otherwise. Actually, this intuitive approach is fully supported by the existing results on the RI property. In fact, it is known [17] that if the matrix $\underline{P}$ is made of random independent entries, its RI constant is small with a substantially large probability.

In general, enforcing the randomness of a process can be a subtle task since the very definition of what is random (entropic, algorithmically complex, etc.) can be extremely sophisticated and also dependent more on philosophical than technical consideration.

Here, for simplicity's sake, I limit ourselves to energy/power considerations and define a measure of the (non)randomness of a process as its self-rakeness, i.e., the average amount of energy/power of the projection of one of its realization onto another realization when the two are drawn independently. The rationale behind this quantification of randomness is that, if $\rho(b, b)$ is high, then independent realizations of the process tend to align and thus to be substantially the same, implying a low "randomness" of the process itself [32].

This definition nicely fits into a mathematical formulation of the design path that increases the rakeness $\rho(a, b)$ while leaving $b$ random enough. In fact, given a certain sparse stochastic process $a$, I determine the stochastic process $b$ generating the projecting waveforms to employ in an RMPI architecture by solving the following optimization problem

$$\begin{aligned} &\max_b \ \rho(a, b) \\ &\text{s.t.} \quad \begin{aligned} &\langle b, b \rangle = e \\ &\rho(b, b) \leq re^2 \end{aligned} \end{aligned} \tag{3.5}$$

where $e$ is the energy of the projection waveforms and $r$ is a randomness-enforcing design parameter.

Roughly speaking, solving (3.5) will ensure that the resulting waveforms will have constant energy (due to constraint $< b, b >= e$) paired with the ability of maximizing the average SNR of the projections (thanks to the capability of maximizing the energy of the acquired samples since I impose that $\max_b \rho(a, b)$) while maintaining the chance of detecting components of the original signal that carry smaller amounts of energy/power (thanks to the fact that

each realization of the process $b$ has $\rho(b, b) \leq re$, i.e., low autocorrelation and thus "large" randomness).

In (3.5), the parameter $e$ acts as a normalization factor, since if $b'$ is the solution for $e = e'$, then $b'' = \sqrt{e''/e'}b'$ is the solution of the same problem for $e = e''$.

On the contrary, $r$ is the parameter controlling the trade-off between the two design criteria I want to blend, i.e., RI and rakeness. Hence, different values of $r$ lead to waveform with different final performance.

Regrettably, though it is easy to accept that, thanks to their ability to maximize the energy of the samples, the resulting $b$ may be able to increase the performance of the overall sensing system, the latter may rely (especially in the reconstruction part) on heavily non-linear and iterative operations that are difficult to model. For this reason, though feasible bounds for the parameter $r$ can (and will) be derived theoretically in Section 3.4 and 3.5, the choice of its exact value is a matter of fine tuning of the global system, and it must be determined through numerical simulation.

## 3.4 Localization in the frequency domain

In this Section I specialize (3.5) to the case in which the statistical features of $a$ that cause the localization of its energy/power can be straightforwardly highlighted by Fourier analysis.

I will concentrate on the time interval $[-T/2, T/2]$ and set $\kappa_\rho = T^{-1}$ in (3.3).

To express $\rho$ in terms of the frequency-domain features of the processes $a$ and $\beta$ in it, let me assume that both of them are second-order stationary.

Leveraging on this, I may define the single-argument correlation functions $C_a(s-t) = C_a(t,s)$ and $C_\beta(s-t) = C_\beta(t,s)$ whose Fourier transforms are nothing but the power spectra $\hat{a}(f)$ and $\hat{\beta}(f)$ of the two processes.

For $\rho(a,\beta)$ I obtain

$$\begin{aligned}
&\rho(a,\beta) = \\
&= \frac{1}{2T}\int_{-T}^{T} C_a(p)C_\beta^*(p)\int_{-T+|p|}^{T-|p|}\mathrm{d}q\mathrm{d}p \\
&= \int_{-T}^{T} C_a(p)C_\beta^*(p)\left(1-\frac{|p|}{T}\right)\mathrm{d}p \qquad (3.6)\\
&= \int_{-T}^{T}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{a}(f)\hat{\beta}^*(g)e^{2\pi\mathbf{i}\,(f-g)p}\left(1-\frac{|p|}{T}\right)\mathrm{d}p\mathrm{d}f\mathrm{d}g \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{a}(f)\hat{\beta}^*(g)\int_{-T}^{T}e^{2\pi\mathbf{i}\,(f-g)p}\left(1-\frac{|p|}{T}\right)\mathrm{d}p\mathrm{d}f\mathrm{d}g \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{a}(f)\hat{\beta}^*(g)h_T(f-g)\mathrm{d}f\mathrm{d}g \qquad (3.7)
\end{aligned}$$

where

$$h_T(f) = \int_{-T}^{T} e^{2\pi\mathbf{i}\,fp}\left(1-\frac{|p|}{T}\right)\mathrm{d}p = \frac{\sin^2(\pi Tf)}{\pi^2 Tf^2}$$

For simplicity's sake I may focus on the antipodal case in which the projection waveforms have a constant-modulus amplitude ($\pm 1$), duration $T$, and thus automatically satisfy the constant energy constraint $< b, b >= e$ in (3.5) with $e = T$, needed to make projection tuning possible.

With this, the power spectrum of the projection waveforms can be designed by solving (3.5) re-expressed in the

frequency domain. To do so, use (3.7) to rewrite $\rho(a, b)$ and $\rho(b, b)$ in (3.5) and consider

$$\begin{aligned} \max_{\hat{b}} \quad & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{a}(f)\hat{b}(g)h_T(f-g)\mathrm{d}f\mathrm{d}g \\ & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{b}(f)\hat{b}(g)h_T(f-g)\mathrm{d}f\mathrm{d}g \le rT \\ & \hat{b}(f) \ge 0 \\ \text{s.}t. \quad & \int_{-\infty}^{\infty} \hat{b}(f)\mathrm{d}f = 1 \\ & \hat{b}(f) = \hat{b}(-f) \end{aligned} \tag{3.8}$$

where the last three constraints encode the fact that $\hat{b}$ must be a power spectrum of a unit-power, real signal.

Once that $r$ is fixed, (3.8) can be solved by assuming that $\hat{a}$ concentrates its power in the frequency interval $[-B, B]$ and applying some kind of finite-elements methods, i.e., approximating all the entailed functions with linear combinations of basic function elements on which the integrals can be computed at least numerically.

As an example, select a frequency interval $[-B, B]$ and partition it $2n + 1$ subintervals of equal length $\Delta f = 2B/(2n + 1)$ $F_j = [j - \Delta f/2, j + \Delta f/2]$ for $j = -n, \ldots, n$. Assume now that $\hat{b}(f)$ is constant in each $F_j$ and define $\chi_j(f)$ as the indicator function of $F_j$, i.e. $\chi_j(f) = 1$ if $f \in F_j$ and 0 otherwise. I have $\hat{b}(f) = \sum_{j=-n}^{n} b_j\chi_j(f)$ for certain coefficients $b_{-n}, \ldots, b_n$.

This can be substituted into (3.8) to obtain

$$\max_{b_{-n},\ldots,b_n} \sum_{j=-n}^{n} w_j b_j$$

$$\sum_{j=-n}^{n} \sum_{k=-n}^{n} b_j b_k W_{j,k} \leq rT$$

$$b_j \geq 0 \quad j = -n, \ldots, n \tag{3.9}$$

s.t.

$$\Delta f \sum_{j=-n}^{n} b_j = 1$$

$$b_j = b_{-j} \quad j = -n, \ldots, n$$

with

$$w_j = \int_{-\infty}^{\infty} \int_{F_j} \hat{a}(f) h_T(f-g) \mathrm{d}f \mathrm{d}g$$

$$W_{j,k} = \int_{F_j} \int_{F_k} h_T(f-g) \mathrm{d}f \mathrm{d}g$$

This leaves me with the vector of $2n+1$ unknown coefficients $b_{-n}, \ldots, b_n$ that must determined by solving an optimization problem characterized by a linear objective function and few linear and quadratic constraints. Plenty of numerical methods exist for solving such problems even for large number of basic-elements and thus for extremely effective approximations (commercial products such as MATLAB or CPLEX provide full support for large-scale version of these problems).

Once that the optimum $\hat{b}(f)$ has been computed, one may resort to known methods to generate an antipodal process with such a spectrum exploiting a linear probability feedback (LPF) [33] [34] [35] [36], some additional comments will be reported in the Section 4.1. Slices

of length $T$ of this process can be used as projection waveforms in an RMPI architecture for the CS of the original $a$.

Note that, even if this is needed to arrive at a final working system, the core of rakeness-based design concerns the solution of (3.5) for frequency-localized signals to obtain the best spectral profile of the projecting waveforms, independently of their physical realization. How such a spectral profile can be obtained using antipodal PAMs is an implementation-dependent choice, which allows to realize an hardware system for sparse and localized signal acquisition which does not require analog multipliers [15].

As far as the range in which $r$ should vary to administer the trade off between RI and rakeness, note that, since $\rho(b, b)$ is a measure of (non)randomness, it must be minimum when the process $b$ is white in its bandwidth, i.e., when $\hat{b}(f) = 1/(2B)$ for $f \in [-B, B]$ and 0 otherwise.

Plugging this into (3.7) and defining $c = BT$ one gets

$$r \geq r^{min} = \\ = \frac{\mathrm{C}i(4\pi c) + 4\pi c \mathrm{S}i(4\pi c) - \log(4\pi c) + \cos(4\pi c) - \gamma - 1}{4\pi^2 c^2}$$

where $\gamma$ is the Euler's constant and C$i$ and S$i$ are respectively the cos-integral and sin-integral functions.

The quantity $r^{min}c$ is a monotonically and rapidly increasing function of $c$ with $\lim_{c\to\infty} r^{min}c = 1/2$. Hence, I may safely use such an asymptotic value to set $1/(2c)$ as a suitable lower bound for $r$ in any practical conditions.

Again, from the meaning of $\rho(b, b)$ I got that it is maximum when the waveforms produced by the process

are constant. This implies $C_b(\tau) = 1$ that can be plugged into (3.6) to obtain

$$r \leq r^{max} = \frac{1}{T}\int_{-T}^{T}\left(1 - \frac{|p|}{T}\right)\mathrm{d}p = 1$$

Overall, the tuning of the overall system will optimize performance by choosing $r \in \left[\frac{1}{2c}, 1\right]$.

## 3.5 Localization in a generic domain

Slices of second-order stationary processes (that enjoy a simple and well-studied characterization in the frequency domain) do not exhaust the set of signals that I may want to acquire.

To cope with more general cases assume to work in normalized conditions such that both the waveforms to be acquired and the projection waveforms have unit energy, i.e., $\int_{-\frac{T}{2}}^{\frac{T}{2}} |a(t)|^2\mathrm{d}t = \int_{-\frac{T}{2}}^{\frac{T}{2}} |b(t)|^2\mathrm{d}t = 1$, where the latter constrain sets $e = 1$ in (3.5).

When I comply with this assumption (possibly by scaling the original signals), if $C_x$ represents either $C_a$ or $C_b$, I have that

- $C_x$ is Hermitian, i.e., $C_x(t, s) = C_x^*(s, t)$;
- $C_x$ is positive semidefinite, i.e., for any integrable function $\xi(t)$
  the quadratic form $\int_{-\frac{T}{2}}^{\frac{T}{2}}\int_{-\frac{T}{2}}^{\frac{T}{2}} \xi^*(t)C_x(t, s)\xi(s)\mathrm{d}t\mathrm{d}s = \mathbf{E}\left[\left|\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)\xi(t)\mathrm{d}t\right|^2\right]$ yields a non-negative result;

- $C_x$ has a unit trace, i.e.,

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} C_x(t,t)\mathrm{d}t = \\ = \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{E}[|x(t)|^2]\mathrm{d}t = \mathbf{E}\left[\int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 \mathrm{d}t\right] = 1.$$

From this, I know (see e.g. [37]) that two sequences of orthonormal functions $\vartheta_0(t), \vartheta_1(t), \ldots$ and $\varphi_0(t), \varphi_1(t), \ldots$ exist, along with the sequences of real non-negative numbers $\mu_0 \geq \mu_1 \geq \ldots$ and $\lambda_0 \geq \lambda_1 \geq \ldots$ such that $\sum_{j=0}^{\infty} \mu_j = \sum_{j=0}^{\infty} \lambda_j = 1$ and

$$C_a(t,s) = \sum_{j=0}^{\infty} \mu_j \vartheta_j^*(t)\vartheta_j(s) \tag{3.10}$$

$$C_b(t,s) = \sum_{j=0}^{\infty} \lambda_j \varphi_j^*(t)\varphi_j(s) \tag{3.11}$$

By substituting the generalized spectral expansions for the two correlation functions (3.10) and (3.11) into (3.4) one gets

$$\rho(a,b) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \lambda_j \mu_k \Xi_{j,k}$$

$$\rho(b,b) = \sum_{j=0}^{\infty} \lambda_j^2$$

where the real and nonnegative numbers

$$\Xi_{j,k} = \left|\int_{-\frac{T}{2}}^{\frac{T}{2}} \varphi_j(t)\vartheta_k^*(t)\mathrm{d}t\right|^2$$

are the squared modulus of the projections of each $\varphi_j$ on every $\vartheta_k$ (and vice versa).

The orthonormality of the $\vartheta_k$ guarantees that the sum of the squared modulus of the projections of $\varphi_j$ must equal

the squared length of $\phi_j$ itself and thus, since $\phi_j$ is normal, that $\sum_{j=0}^{\infty} \Xi_{j,k} = 1$. Conversely, from the fact that the $\phi_j$ are orthonormal I have also $\sum_{k=0}^{\infty} \Xi_{j,k} = 1$.

Hence, the optimization problem (3.5) can be rewritten in totally generic terms as

$$\max_{\hat{\eta}} \max_{\Xi} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \hat{\eta}_j \mu_k \Xi_{j,k}$$

$$\text{s.t.} \quad \begin{aligned} \hat{\eta}_j &\geq 0 \ \forall j \\ \sum_{j=0}^{\infty} \hat{\eta}_j &= 1 \\ \sum_{j=0}^{\infty} \hat{\eta}_j^2 &\leq r \\ \Xi_{j,k} &\geq 0 \ \forall j,k \\ \sum_{j=0}^{\infty} \Xi_{j,k} &= 1 \ \forall k \\ \sum_{k=0}^{\infty} \Xi_{j,k} &= 1 \ \forall j \end{aligned} \tag{3.12}$$

Note that the two max operators address separately the problem of finding an optimal basis ($\max_{\Xi}$) and then the optimal energy distribution over that basis ($\max_{\hat{\eta}}$).

As far as the range of $r$ is concerned, assume to know that $J$ is an integer such that $\hat{\eta}_j = 0$ for $j \geq J$. It can be easily seen that $\max \sum_{j=0}^{J-1} \hat{\eta}_j^2$ subject to the constraints $\hat{\eta}_j \geq 0$ and $\sum_{j=0}^{J-1} \hat{\eta}_j = 1$ is 1 and is attained when $\hat{\eta}_0 = 1$ and $\hat{\eta}_j = 0$ for $j > 0$. It is also easy to see that $\min \sum_{j=0}^{J-1} \hat{\eta}_j^2$ subject to the constraints $\hat{\eta}_j \geq 0$ and $\sum_{j=0}^{J-1} \hat{\eta}_j = 1$ is $1/J$ and is attained when $\hat{\eta}_j = 1/J$ for $j = 0, \ldots, J-1$. Hence, $r \in [1/J, 1]$.

In particular, the lower bound $r \geq 1/J$ rewritten as $rJ \geq 1$ can be read as a general rule of thumb, i.e., the

more random the process that generates the projection waveforms, the larger the number of non-zero eigenvalues in the spectral expansion of its correlation function.

The solution of (3.12) is derived in the Appendix and depends on the two partial sums

$$\Sigma_1(J) = \sum_{j=0}^{J-1} \mu_j \tag{3.13}$$

$$\Sigma_2(J) = \sum_{j=0}^{J-1} \mu_j^2 \tag{3.14}$$

to obtain

$$\phi_j = \partial_j \tag{3.15}$$

$$\lambda_j = \lambda_j(J) = \frac{1}{J}\left[1 + \frac{J\mu_j - \Sigma_1(J)}{\sqrt{\frac{\Sigma_2(J) - \frac{1}{J}\Sigma_1^2(J)}{r - \frac{1}{J}}}}\right] \tag{3.16}$$

which hold for $j = 0, 1, \ldots, J-1$ where $J$ is defined by

$$J = \max\left\{j \;\middle|\; \lambda_{j-1}(j) > 0\right\} \tag{3.17}$$

By definition, all the eigenvalues $\lambda_j$ for $j \geq J$ are null.

### 3.5.1 Finite dimensional signals

The special case in which the signal to be acquired can be written as a linear combination of known waveforms by means of random coefficients is, for us, extremely interesting and deserves some further discussion.

Let me assume that (3.1) holds for orthonormal $u_j$ ($j = 0, \ldots, N-1$) and let me compute

$$C_a(t,s) = \sum_{j=0}^{N-1}\sum_{k=0}^{N-1} \mathbf{E}[a_j^* a_k] u_j^*(t) u_k(s) \tag{3.18}$$

The correlation matrix $\underline{A} = [\underline{A}_{j,k}] = \left[\mathbf{E}[a_j^* a_k]\right]$ is Hermitian and positive semidefinite, hence it can be written as $\underline{A} = \underline{Q}\underline{M}\underline{Q}^\dagger$ where $\cdot^\dagger$ stands for transposition and conjugation, $\underline{M}$ is a diagonal matrix with real non-negative diagonal entries, and $\underline{Q}$ is an orthonormal matrix whose columns are the eigenvectors of $\underline{A}$.

With this, I may rewrite (3.18) as

$$\begin{aligned} C_a(t,s) &= \sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} \underline{Q}_{j,l}\underline{M}_{l,l}\underline{Q}^*_{k,l} u_j^*(s) u_k(t) \\ &= \sum_{l=0}^{N-1} \underline{M}_{l,l} \sum_{j=0}^{N-1} \underline{Q}_{j,l} u_j^*(t) \sum_{k=0}^{N-1} \underline{Q}^*_{k,l} u_k(s) \end{aligned}$$

Hence, I may express $C_a(t,s)$ in the form needed for writing (3.10) and thus the solution of (3.12) by simply setting $\vartheta_j = \sum_{k=0}^{N-1} \underline{Q}^*_{k,j} u_k$ and $\mu_j = \underline{M}_{j,j}$ for $j = 0, \ldots, N-1$.

This straightforward derivation clarifies that, when I have identified sparseness along a certain signal basis, the statistic of the coefficients gives me hints on the basis that may be used to highlight localization. Along this other basis, localization itself is nothing but the difference between the lower-index, largest eigenvalues $\mu_0, \mu_1, \ldots$ and the others.

A bridge is also built between the general treatment of rakeness in this Section and the frequency-domain analysis of the Section 3.4. In fact, if $a$ is substantially band-limited in the frequency interval $[-B, B]$ and is considered in the time interval $[-T/2, T/2]$ its realizations may be well expressed as a linear combination of waveforms that are the truncated version of prolate spheroidal wave

functions [38] [39]. It is known that, if $c = BT$ then $N = 2c$ functions are enough to achieve an approximation quality that dramatically increases as $c \to \infty$. Hence, the solution of (3.12) will feature $J = N = 2c$ for values of $r \in [1/J, 1] = [\frac{1}{2c}, 1]$.

From an operative point of view, whatever analysis allows me to obtain the generalized spectral expansion of $C_a$ as in (3.10), i may use (3.15), (3.16), (3.17) and (3.11) to compute the correlation function $C_b$ of the process generating the projection waveforms.

To fit this $C_b$ into an actual RMPI architecture, I must generate a binary or antipodal PAM signal with such a non-stationary correlation. The details of the mechanism allowing this are far beyond the scope of this paper and will be the topic of a future communication.

It is here enough to say that, if the number of symbols $S$ in each waveform is limited to few tens (say $S < 100$), I are able, depending on $C_b$, to automatically determine two sets of cardinality $s = S(S + 1)/2$: the first set $\{z_0, z_1, \ldots, z_{s-1}\}$ contains sequences of modulating symbols, while the second set $\{\zeta_0, \zeta_1, \ldots, \zeta_{s-1}\}$ contains probabilities, so that $\sum_{j=0}^{s-1} \zeta_j = 1$.

These two sets are such that, if each time a projection waveform is needed, the modulating symbols in $z_j$ are used with probability $\zeta_j$, then the resulting process has the desired correlation.

In any case, let me stress that, as noted before for frequency-localized signals, the core of rakeness-based design concerns the solution of (3.5), which is here described for generically localized signals. Once that the correlation of the best projection waveforms is determined, their actual realization depends on implementation as-

sumption that may vary from application to application.

## 3.6 Some Proofs

### 3.6.1 Solution of (3.12)

The first subproblem $\max_{\Xi}$ can be solved leveraging on the fact that it is a linear problems with linear constraints. Since, in principle, it may involve an infinite number of variables I should proceed by steps.

Let $P_n$ be the optimization problem

$$\begin{aligned} \max_{\Xi} & \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} \lambda_j \mu_k \Xi_{j,k} \\ & \Xi_{j,k} \geq 0 \ \forall j,k \\ s.t. & \sum_{j=0}^{\infty} \Xi_{j,k} = 1 \ \forall k \\ & \sum_{k=0}^{\infty} \Xi_{j,k} = 1 \ \forall j \end{aligned}$$

so that $P_\infty$ is the basis finding subproblem in (3.12).

Since all the series involved in the definition of $P_\infty$ are convergent, I have that, independently of $\Xi_{j,k}$,

$$\lim_{n\to\infty} \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} \lambda_j \mu_k \Xi_{j,k} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \lambda_j \mu_k \Xi_{j,k}$$

Moreover, since all the summands are positive, the limit is from below.

Let me now assume to have solved $P_\infty$ yielding a value $\sigma(P_\infty)$ corresponding to a certain optimal choice $\hat{\Xi}^{\infty}_{j,k}$.

Given any $\epsilon > 0$ there is a $\bar{n}$ such that for any $n \geq \bar{n}$

$$0 \leq \sigma(P_\infty) - \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} \lambda_j \mu_k \hat{\Xi}^{\infty}_{j,k} \leq \epsilon$$

Yet, by solving $P_n$ I get a solution $\sigma(P_n)$ such that

$$\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\hat{\eta}_j\mu_k\hat{\Xi}_{j,k}^{\infty} \leq \sigma(P_n) \leq \sigma(P_\infty)$$

where the last inequality holds since every feasible configuration for $P_n$ is also a feasible configuration for $P_\infty$.

Altogether I get that for any $n \geq \bar{n}$

$$0 \leq \sigma(P_\infty) - \sigma(P_n) \leq \epsilon$$

that is

$$\lim_{n\to\infty}\sigma(P_n) = \sigma(P_\infty)$$

from below.

From this I know that, if the solutions $\hat{\Xi}_{j,k}^{n}$ of $P_n$ have a limit, such a limit yields $\sigma(P_\infty)$.

To study the solutions of $P_n$ I may first recall that the polytope

$$\begin{array}{rcl}
\Xi_{j,k} & \geq & 0 \; j,k = 0,\ldots,n-1 \\
\sum_{j=0}^{n-1}\Xi_{j,k} & = & 1 \; k = 0,\ldots,n-1 \\
\sum_{k=0}^{n-1}\Xi_{j,k} & = & 1 \; j = 0,\ldots,n-1
\end{array}$$

is the one characterizing the so-called "assignment" problems [40] and is well known [41] to have vertices for $\Xi_{j,k}$ for $j,k = 0,\ldots,n-1$ equal to a permutation matrix. Hence, let $\xi : \{0,1,\ldots,n-1\} \mapsto \{0,1,\ldots,n-1\}$ be the bijection such that

$$\Xi_{j,k} = \begin{cases} 1 & \text{if } k = \xi(j) \\ 0 & \text{otherwise} \end{cases}$$

i have

$$\sigma(P_n) = \sum_{j=0}^{n-1}\hat{\eta}_j\mu_{\xi(j)}$$

for some optimally chosen $\xi$.

Actually, I may prove that such an optimal $\xi$ is the identity. I do it by induction.

For $n = 2$ there are only two permutations corresponding to the two candidate solutions $\sigma' = \lambda_0\mu_0 + \lambda_1\mu_1$ and $\sigma'' = \lambda_0\mu_1 + \lambda_1\mu_0$. Yet, from the sorting of the $\lambda_j$ and of the $\mu_j$ I have $\sigma' - \sigma'' = (\lambda_0 - \lambda_1)(\mu_0 - \mu_1) \geq 0$.

This confirms that the optimum solution is the one corresponding to $\xi(j) = j$ for $j = 0, 1$.

Assume now that this is true for $n$ up to a certain $\bar{n}$ and that I have solved $P_{\bar{n}+1}$ by means of a permutation $\xi$.

If $\xi(0) = \bar{\jmath} > 0$ then $\sigma(P_{\bar{n}+1}) = \lambda_0\mu_{\bar{\jmath}} + \sigma'$. Yet, $\sigma'$ must be the value of the solution of a problem with $\bar{n}$ terms $\lambda_1, \ldots, \lambda_{\bar{n}}$ and $\mu_1, \ldots, \mu_{\bar{\jmath}-1}, \mu_{\bar{\jmath}+1}, \ldots, \mu_{\bar{n}}$. Since i assumed to know how problems with $\bar{n}$ terms are solved i know that

$$\sigma' = \sum_{j=1}^{\bar{\jmath}} \lambda_j\mu_{j-1} + \sum_{j=\bar{\jmath}+1}^{\bar{n}} \lambda_j\mu_j$$

It is now easy to see that the value $\lambda_0\mu_{\bar{\jmath}} + \sigma'$ of the alleged solution is actually smaller than $\sum_{j=0}^{\bar{n}} \lambda_j\mu_j$.

In fact

$$\begin{aligned}
&\sum_{j=0}^{\bar{n}} \lambda_j\mu_j - \lambda_0\mu_{\bar{\jmath}} - \sum_{j=1}^{\bar{\jmath}} \lambda_j\mu_{j-1} - \sum_{j=\bar{\jmath}+1}^{\bar{n}} \lambda_j\mu_j = \\
&= \lambda_0(\mu_0 - \mu_{\bar{\jmath}}) - \sum_{j=1}^{\bar{\jmath}} \lambda_j(\mu_{j-1} - \mu_j) \\
&= \lambda_0 \sum_{j=1}^{\bar{\jmath}} (\mu_{j-1} - \mu_j) - \sum_{j=1}^{\bar{\jmath}} \lambda_j(\mu_j - \mu_{j-1}) \\
&= \sum_{j=1}^{\bar{\jmath}} (\lambda_0 - \lambda_j)(\mu_{j-1} - \mu_j) \geq 0
\end{aligned}$$

Hence, the optimal permutation must feature $\xi(0) = 0$. This reduces the solution of $P_{\bar{n}+1}$ to the solution of $P_{\bar{n}}$ that i already know to be $\xi(j) = j$ for $j = 1, \ldots, n-1$.

In the light of this, every $P_n$ has a solution corresponding to $\xi(j) = j$ for $j = 0, \ldots, n-1$ and the solution of $P_\infty$ is

$$\begin{aligned} \Xi_{j,k} &= \delta_{j,k} \\ \sigma(P_\infty) &= \sum_{j=0}^{\infty} \pi_j \mu_j \end{aligned}$$

This solves the basis-selection problem and yields (3.15).

The original (3.12) now becomes

$$\begin{aligned} \max_{\pi} & \sum_{j=0}^{\infty} \pi_j \mu_j \\ & \pi_j \geq 0 \ \forall j \\ \text{s.t.} & \textstyle\sum_{j=0}^{\infty} \pi_j = 1 \\ & \textstyle\sum_{j=0}^{\infty} \pi_j^2 \leq r \end{aligned} \tag{3.19}$$

Since the $\pi_j$ are non-negative and sorted in non-increasing order I have that the set of indexes such that $\pi_j > 0$ must be of the kind $\{0, 1, \ldots, J-1\}$ for some integer $J \geq 0$. I also know that, to allow $\sum_{j=0}^{J-1} \pi_j = 1$ and $\sum_{j=0}^{J-1} \pi_j^2 = r$ to hold simultaneously i must have $r \geq 1/J$ and thus $J \geq 1/r$.

Hence, for a given $J \geq 1/r$ my problem can be recast into

$$\begin{aligned} \max_{\pi} & \sum_{j=0}^{J-1} \pi_j \mu_j \\ & \pi_j > 0 \ j = 0, \ldots, J-1 \\ \text{s.t.} & \textstyle\sum_{j=0}^{J-1} \pi_j = 1 \\ & \textstyle\sum_{j=0}^{J-1} \pi_j^2 \leq r \end{aligned} \tag{3.20}$$

Note that the feasibility set of (3.20) for a certain $J = \bar{J}$ contains points that are arbitrarily close to those of the feasibility set of (3.20) for any $J < \bar{J}$. Hence, to maximize the rakeness I should try to have $J$ as large as possible.

To determine the $J$ leading to maximum rakeness note first that, if I drop the randomness constraint $\sum_{j=0}^{J-1} \hat{\pi}_j^2 \leq r$, the relaxed problem has the trivial solution $\hat{\pi}_0 = 1$ and $\hat{\pi}_j = 0$ for $j > 0$. Such a solution is not feasible for the original problem since $\sum_{j=0}^{J-1} \hat{\pi}_j^2 = 1 \geq r$, hence the corresponding optimum must be attained when the randomness constraint is active, i.e. for $\sum_{j=0}^{J-1} \hat{\pi}_j^2 = r$.

The Karush-Kuhn-Tucker conditions for (3.19) with the inequality constraint substituted by the equality constraint are

$$\begin{aligned}
\mu_j + \ell' + \ell'' \hat{\pi}_j + \ell_j''' &= 0 \ \forall j \\
\hat{\pi}_j &\geq 0 \ \forall j \\
\textstyle\sum_{j=0}^{\infty} \hat{\pi}_j &= 1 \\
\textstyle\sum_{j=0}^{\infty} \hat{\pi}_j^2 &= r \\
\ell_j''' &\geq 0 \ \forall j \\
\ell_j''' \hat{\pi}_j &= 0 \ \forall j
\end{aligned}$$

where $\ell'$ is the Lagrange multiplier corresponding to $\sum_{j=0}^{\infty} \hat{\pi}_j = 1$, $\ell''$ is the Lagrange multiplier corresponding to $\sum_{j=0}^{\infty} \hat{\pi}_j^2 = r$, and $\ell_j'''$ are the Lagrange multipliers corresponding to $\hat{\pi}_j \geq 0$ which must hold $\forall j$.

Since for $\hat{\pi}_j > 0$ the constraint $\ell_j''' \hat{\pi}_j = 0$ sets $\ell_j''' = 0$ I know that

$$\hat{\pi}_j = -\frac{\mu_j + \ell'}{\ell''} \tag{3.21}$$

for $j = 0, \ldots, J-1$.

Since the sequences $\hat{\pi}_j$ and $\mu_j$ are both decreasing, i must have $\ell'' < 0$ and thus $\ell' > -\mu_j$ for $j = 0, 1, \ldots, J-1$,

i.e., $\ell' > -\mu_{J-1}$. Hence,

$$J = \max\left\{j \mid \ \ell' > -\mu_{j-1}\right\}$$

To see how the two parameters $\ell'$ and $\ell''$ depend on $J$ note that they should satisfy the simultaneous equations

$$\begin{array}{rcl} \sum_{j=0}^{J-1} \frac{\mu_j+\ell'}{-\ell''} & = & 1 \\ \sum_{j=0}^{J-1} \left(\frac{\mu_j+\ell'}{\ell''}\right)^2 & = & r \end{array}$$

i.e., exploiting (3.13) and (3.14),

$$\begin{array}{rcl} \Sigma_1(J) + J\ell' & = & -\ell'' \\ \Sigma_2(J) + 2\Sigma_1(J)\ell' + J\ell'^2 & = & \ell''^2 r \end{array}$$

Such equations can be solved for $\ell'$ and $\ell''$ and the resulting values substituted in (3.21) to yield (3.16).

### 3.6.2 Real and nonnegative values in (3.16)

The denominator within the square root is positive whenever $r > 1/J$.

To show that the corresponding numerator is also nonnegative write

$$\begin{array}{l} J\Sigma_2(J) - \Sigma_1^2(J) = \\ = J\displaystyle\sum_{j=0}^{J-1} \mu_j^2 - \sum_{j=0}^{J-1}\sum_{k=0}^{J-1} \mu_j\mu_k \\ = J\displaystyle\sum_{j=0}^{J-1} \mu_j \left[\mu_j - \frac{1}{J}\sum_{k=0}^{J-1} \mu_k\right] \end{array}$$

Let now $\zeta_j = \mu_j - \frac{1}{J}\sum_{k=0}^{J-1}\mu_k$. I have that the $\zeta_j$ are decreasing and such that $\sum_{j=0}^{J-1}\zeta_j = 0$. Hence, there is a $j'$

such that $\sum_{j=0}^{j'-1} \zeta_j = \sum_{j=j'}^{J-1}(-\zeta_j) \geq 0$. Hence,

$$
\begin{aligned}
& J\Sigma_2(J) - \Sigma_1^2(J) = \\
& = J\left[\sum_{j=0}^{j'-1} \mu_j \zeta_j - \sum_{j=j'}^{J-1} \mu_j(-\zeta_j)\right] \\
& \geq J\left[\mu_{j'-1} \sum_{j=0}^{j'-1} \zeta_j - \mu_{j'} \sum_{j=j'}^{J-1}(-\zeta_j)\right] \\
& = J(\mu_{j'-1} - \mu_{j'}) \sum_{j=0}^{j'-1} \zeta_j \geq 0
\end{aligned}
$$

CHAPTER

# 4 Rakeness on ECGs and small images

In this Chapter I introduce rakeness as a design criterion to optimize the performance of two acquisition systems, one that deals with Electro Cardio Graphic (ECG) signals, which can be easily modeled in the frequency domain as I did in Section 3.4, and the other that must be described relying on the generalized spectral expansions in Section 3.5 since its target signals are images.

Despite the fact that the two scenarios are different, the path i follow in designing an acquisition system based on CS system is the same and can be summarized in few steps:

i. identify the basis with respect to which the signal to acquire is sparse;

ii. identify the basis with respect to which the signal is localized;

iii. solve (3.5) for a number of possible values $r$ in its range;

iv. for each value of $r$, implement an RMPI architecture exploiting the sparsity revealed in i) and in which the projection waveforms are as close as possible to the optimal ones;

v. perform Monte-Carlo simulations to evaluate the resulting systems and select the best performing one.

Note that, in the classical design flow of a CS system, i) is a prerequisite while iv) is tackled once assuming that the projection waveform are random PAM signals with independently and identically distributed ("i.i.d." from now on) symbols. This is what will be taken as the reference case to quantitatively assess the improvements due to rakeness-based design.

In all cases, the performance index is the average reconstruction SNR (ARSNR), i.e. the average ratio between the energy of the original signal over the energy of the difference between the original signal and the reconstructed one. ARSNR values are always plotted at the center of an interval accounting for the variances of the corresponding reconstruction SNRs.

Note also that the implementation constraints (e.g., the restriction of projection waveform to PAM profiles with antipodal symbols) come into play only in iv).

Finally, one may observe that steps iii-v are nothing but an elementary line-search for the best possible value of $r$. As a matter of fact, the values of $r$ for which a definite improvement can be obtained are easily identified by means of a very small numbers of trials.

To simulate these scenarios I need two different random symbols generators, one is needed to generate

antipodal random values with a prescribed power spectral density (it will be used to acquire ECGs). In these cases the LPF [35] can be used.

The second one is needed to generate random antipodal symbols with a prescribed second order characterization, this is the 2-Values Random Vectors Generator (2vRVG).

Both symbols generators will be briefly discussed in the first two section of this Chapter and then Rakeness will be applied to the acquisition of ECGs and small images in the last two Sections.

## 4.1 Linear Probability Feedback Process

As anticipated in the previous chapter, There is a well defined class of stochastic processes able to generate a stream of antipodal symbols with prescribed power spectral density (PSD), i.e.,linear probability feedback processes (LPF) [33] [34] [35] [36]. The scheme of an LPF generator is shown in Figure 4.1, where the PSD of the antipodal values $a_j$ is imposed by fixing the internal coefficients of the filter included in the feedback loop. Now I will briefly discuss this generator in order to have a more comprehensible reading about the mechanism used to generate the sampling sequence needed in ECG signals acquisition. A complete presentation of the result about CS based on rakeness on ECG signals is the topic of the Section 4.3.

The LPF relies on a causal time-invariant linear filter

with impulse response $h_k$ and transfer function

$$H_m(z) = \sum_{k=1}^{m} h_k z^{-k}$$

Though, in principle, there is no need for $m$ to be finite it is assumed so here. The output of the filter $-H_m(z)$ produces the process $s_j = -\sum_{k=1}^{m} h_k a_{j-k}$, assuming $\sum_{k=1}^{m} |h_k| \leq 1$ and since $a_j \in \{-1, 1\}$ is such that $s_j \in [-1, 1]$. The process $s_j$ is then fed into a comparator and matched with the process $z_j$ that is made of independent random thresholds uniformly distributed in $[-1, 1]$. The comparator yields antipodal values $a_j \in \{-1, 1\}$ that are fed back into the filter to continue generation.

As reported in [36], the PSD $\Psi_a(f)$ of $a_j$ can be expressed as:

$$\Psi_a(f) = \frac{\left|1 + H_m(e^{2\pi i f})\right|^{-2}}{\int_{-1/2}^{1/2} \left|1 + H_m(e^{2\pi i f})\right|^{-2} df} \tag{4.1}$$

In principle one could use (4.1) to derive $h_k$ once $\Psi_a(f)$ has been set to a desired PSD profile similar to the one imposed by the solution of the optimization problem 3.9 for a specific class of input signals. Regrettably, since inverting (4.1) is a prohibitive task, one can rely on numerical methods such as the modified gradient descent [35] or more sophisticated heuristic techniques [36].

When ECG signals are acquired by CS based on rakeness, $m = 10$ is sufficient to reproduce the smooth PSD imposed by the output of the optimization problem 3.9.

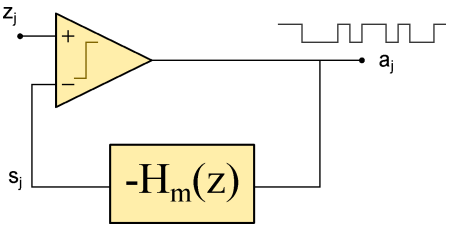



**Figure 4.1:** the block scheme of the LPF process.

## 4.2 2-Value Random Vector Generator

The aim of this section is to derive a stochastic generator of arrays $x = (x_0, \ldots, x_{n-1})^\top$ with $x_j \in \{-1, +1\}$ and a prescribed cross-correlation. This corresponds to a set of given numbers $|\gamma_{j,k}| \leq 1$ with $0 \leq j < k < n$ such that

$$\mathbf{E}[x_j x_k] = \gamma_{j,k}$$

In general, all the instances of the $n$-dimensional real random vector $x$ are such that $x \in X^n$ for a certain finite set $X$ with $|X|$ elements, in this case $X = \{-1, +1\}$.

The system to design is a generator of instances of $x$ guaranteeing that the 2-nd order statistical correlations between their components $\mathbf{E}[x_j x_k]$ are as close as possible to a prescribed correlation $\gamma_{j,k}$.

This is a variant of the antipodal spectrum synthesis proposed in [35] and discussed in the previous section, in that it is equivalent of considering only process values collected in $n$ subsequent time steps and arranged in the

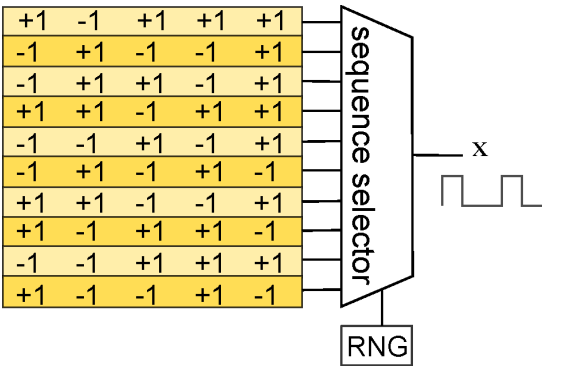



**Figure 4.2:** Block diagram of the 2-value Random Vector Generator with $n = 5$, where RNG is a random number generator uniformly distributed in [0, 1].

vector $x$ for which a possibly non-stationary correlation (i.e., such that, in general $\gamma_{j,k} \neq \gamma_{j+l,k+l}$) is specified.

The general structure of the generator propose here is that of a simple lookup-table whose entries are read and output according to a certain probability distribution.

In particular, once the joint PDF of the random vector $x$ is given as a function $p : X^n \mapsto [0, 1]$, I may concentrate on its support;

$$P = \text{supp } p = \{x \in X^n \mid p(x) > 0\}$$

storing it along with the corresponding values of $p$.

With this data structure, the availability of a simple source of independent real random variables with a known PDF (typically uniform in [0, 1]) is enough to select which vector in $P$ is output each time an instance of $x$ is needed, as it is reported on Figure 4.2.

With such a trivial scheme in mind, the problem reduces to the determination of $p$ starting from the

correlation values in such a way that $P$ has the smallest possible cardinality, i.e., as far as possible from the obvious upper bound $|X|^n$.

As a first step note that, due to the discrete nature of the vector $x$, the correlation between its $j$-th and $k$-th component can be written as $\mathbf{E}[x_j x_k] = \sum_{x \in X^n} p(x) x_j x_k$ for $0 \leq j < k < n$, where the cases $j = k$, related to the auto-correlation, have the trivial solution $\mathbf{E}[x_j^2] = \sum_{x \in X^n} p(x) x_j^2 = 1$ with $x_j \in \{-1, +1\}$.

With this, I may define $\Delta_{j,k} = \gamma_{j,k} - \sum_{x \in X^n} p(x) x_j x_k$ as the individual deviation between the cross-correlation due to the choice of $p$ and the desired cross-correlations.

I choose to measure the difference between the actual and desired correlation profiles with:

$$\Delta_1 = \sum_{0 \leq j < k < n} w_{j,k} \left| \Delta_{j,k} \right|$$

for some set of positive weigths $w_{j,k}$ with $0 \leq j < k < n$.

Such a difference can be used as a guiding criterion in the design procedure by solving the following optimization problem

$$\begin{aligned} \min \quad & \Delta_1 \\ \text{s.t.} \quad & \sum_{x \in X^n} p(x) = 1 \\ & p(x) \geq 0 \qquad \text{for } x \in X^n \end{aligned} \tag{4.2}$$

whose variables are nothing but the $|X|^n$ values $p(x)$ assumed by the PDF $p$ at the points in $X^n$.

Techniques for solving (4.2) will be discussed in the following subsections.

By now, let me analyze the support of the resulting $p$, by recognizing that (4.2) is a linear programming (LP)

problem. In fact, define the variables $\delta^+_{j,k}$ and $\delta^-_{j,k}$ for $0 \leq j < k < n$ to lay down the following minimization problem

$$\begin{aligned}
\min \quad & \sum_{0 \leq j < k < n} w_{j,k} \left( \delta^+_{j,k} + \delta^-_{j,k} \right) \\
& \sum_{x \in X^n} p(x) = 1 \\
& \delta^+_{j,k} - \delta^-_{j,k} + \sum_{x \in X^n} p(x) x_j x_k = \gamma_{j,k} \quad 0 \leq j < k < n \\
\text{s.t.} \quad & \delta^+_{j,k} \geq 0 \qquad\qquad 0 \leq j < k < n \\
& \delta^-_{j,k} \geq 0 \qquad\qquad 0 \leq j < k < n \\
& p(x) \geq 0 \qquad\qquad \text{for } x \in X^n
\end{aligned} \tag{4.3}$$

This LP problem has $|X|^n + 2\binom{n}{2}$ variables that must be non-negative and must satisfy $\binom{n}{2} + 1$ additional equalities constraints.

The solution to (4.3) is the same of (4.2). In fact, for any PDF $p$, the constraints force $\Delta_{j,k} = \delta^+_{j,k} - \delta^-_{j,k}$ with $\delta^+_{j,k}, \delta^-_{j,k} \geq 0$ for every $0 \leq j < k < n$. Once $\Delta_{j,k}$ is known, to minimize the contribution of $\delta^+_{j,k} + \delta^-_{j,k}$ to the objective function it is convenient to set either $\delta^+_{j,k}$ or $\delta^-_{j,k}$ to zero (depending on the sign of $\Delta_{j,k}$) and let the other be the absolute value of $\Delta_{j,k}$. Hence, what is actually minimized is the weighted sum of the absolute values of the $\Delta_{j,k}$ and thus $\Delta_1$.

The set $\mathcal{F}$ of points satisfying the constraints (the so-called *feasibility* space) is non-empty since it contains at

least the point

$$
\begin{aligned}
p(x) &= 0 && \text{for } x \in X^n \\
\delta^+_{j,k} &= \begin{cases} \gamma_{j,k} & \text{if } \gamma_{j,k} \geq 0 \\ 0 & \text{otherwise} \end{cases} && \text{for } 0 \leq j \leq k < n \\
\delta^-_{j,k} &= \begin{cases} -\gamma_{j,k} & \text{if } \gamma_{j,k} < 0 \\ 0 & \text{otherwise} \end{cases} && \text{for } 0 \leq j \leq k < n
\end{aligned}
$$

and it is bounded since the probabilities and correlations are bounded, i.e., it is a polytope in $\mathbb{R}^{|X|^n+2\binom{n}{2}}$.

In this case, a basic result of Operation Research, ensures that a solution of (4.3) exists in one of the vertices of $\mathcal{F}$.

Since vertices are 0-dimensional subspaces and equalities in (4.3) provide only $\binom{n}{2}+1$ linear constraints, in each vertex also $|X|^n + \binom{n}{2} - 1$ non-negativity constraints must be "active", i.e. they should hold as equalities zeroing the corresponding variables. Hence, each vertex and thus also at least one solution to (4.3) has at least $|X|^n - \binom{n}{2} - 1$ vanishing entries and a support of not more than $\binom{n}{2} + 1$ entries. The same holds for the support of $p$ that is a subset of the support of the solution of (4.3).

This is actually extremely good news since $\binom{n}{2}+1$ grows quadratically and not exponentially with $n$ thus making the stochastic lookup-table approach sensible for non-negligible values of $n$.

In this case, once I know that small-support PDFs exist I may build PDFs with a larger support.

Assume, in fact, that two different PDFs $p'(x)$ and $p''(x)$ exist for which the difference between the actual correlations and the desired ones are $\Delta'_1$ and $\Delta''_1$ respectively.

c=

$$\left(w_{0,1}\dots w_{n-2,n-1}\,\middle|\,w_{0,1}\dots w_{n-2,n-1}\,\middle|\,0\dots 0\right)$$

A=

$$\left(\begin{array}{c|c|c} 0\dots 0 & 0\dots 0 & 1\dots 1 \\ I_{\binom{n}{2}} & -I_{\binom{n}{2}} & \xi_0\dots\xi_{|X|^n-1} \end{array}\right)$$

b=

$$\begin{pmatrix} 1 \\ \gamma_{0,0} \\ \vdots \\ \gamma_{n-2,n-1} \end{pmatrix}$$

**TABLE 4.1:** The vectors and the matrix describing the LP implied by 4.3

Choose an $\epsilon \in ]0,1[$ and set $p(x) = \epsilon p'(x) + (1-\epsilon)p''(x)$. Clearly, $p(x)$ is a PDF, its support is the union of $P'$ and $P''$ and, for $0 \le j < k < n$ I may write

$$\begin{aligned} \left|\Delta_{j,k}\right| &= \left|\gamma_{j,k} - \sum_{x\in X^n} p(x)x_j x_k\right| \\ &= \Big|\epsilon\gamma_{j,k} + (1-\epsilon)\gamma_{j,k} \\ &\qquad -\epsilon\sum_{x\in X^n} p'(x)x_j x_k - (1-\epsilon)\sum_{x\in X^n} p''(x)x_j x_k\Big| \\ &\le \epsilon\left|\Delta'_{j,k}\right| + (1-\epsilon)\left|\Delta''_{j,k}\right| \end{aligned}$$

confirming that the "quality" of the new PDF is comparable with that of the two generating PDFs.

### 4.2.1 Efficient synthesis procedures

The general standard formulation for a LP problem can be written in matrix terms as

$$\begin{aligned} \min \quad & z = cq \\ \text{s.t.} \quad & Aq = b \\ & q \geq 0 \end{aligned} \tag{4.4}$$

where $z$ is the real number to minimize that depends linearly on $V$ non-negative variables collected in the column vector $q$ through the coefficients contained in the $V$-dimensional row vector $c$. The same variables are also subject to $C$ equality constraints that are described by the $C \times V$ matrix $A$ and by the the $C$-dimensional column vector $b$.

Now, (4.3) can be mapped into(4.4). To do so in a systematic way, agree to sort two-index quantities like $\gamma_{j,k}$ with $0 \leq j < k < n$ as $\gamma_{0,1}, \gamma_{0,2}, \ldots, \gamma_{1,2}, \gamma_{1,3}, \ldots, \gamma_{n-2,n-1}$ when they have to be aligned in a single array. According to this, for any $x \in X^n$ define the vector $\xi$ aligning the values $x_j x_k$ for $0 \leq j < k < n$ and enumerate all such vectors as $\xi_0, \ldots, \xi_{|X|^n-1}$ in an arbitrary order.

Relying on these definitions, I may address (4.3) and note that, in this case, $V = |X|^n + 2\binom{n}{2}$ while $C = \binom{n}{2} + 1$.

The vectors and matrix defining the LP problem equivalent to the before mentioned criterion are reported in Table 4.1 where $I_m$ is the $m \times m$ identity matrix.

Though extremely efficient tools exist to tackle the LP problems defined by Table 4.1, it is evident that, when $n$ increases, the number of columns in the $A$ matrices increases exponentially. This is a clear mark of the

combinatorial nature of the problem itself, that ultimately stems from the fact that I am dealing with discrete-valued random variables.

Luckily enough, the mapping of combinatorial selection problems into very huge LP problems is a largely explored field with established techniques that can be applied to my cases.

In particular, since I am dealing with matrices with a huge number of columns, I should look into *column generation* methods that directly follow from the properties of the most celebrated algorithm for solving LP problems, i.e., the simplex method.

Seen from my point of view (and overlooking all assumptions and numerical machinery needed to make it work) the simplex method leverages on the knowledge that at least a solution of (4.4) exists for which only $C$ of the $V$ entries of $q$ are non-zero. This means that only $C$ of the columns of the matrix $A$ are involved in the linear combination yielding $b$.

Hence, solving the LP problem amounts to selecting the *right* set of columns. Once a set of $C$ columns is selected (usually called a *basis*), under suitable non-degeneracy assumptions, the solution of a linear system of equalities allows to derive the value of the non-zero entries of $q$ and thus of the value of $z$ that must be minimized.

The second key point of the method is that, thanks to the linearity of the problem, optimality can be pursued in an iterative way by taking any candidate basis and possibly substituting its elements with new columns one at a time so that the corresponding value of $z$ decreases at each substitution.

Hence, the elementary step is made of two actions:

i. find a column (if any) that can enter the basis to reduce the value of $z$;

ii. find the column that can exit from the basis without impairing the reduction of $z$.

Assume that the $C$ non-zero entires of $q$ are selected and that we collect in $A'$ the corresponding columns of $A$ and in $c'$ the corresponding entries of $c$.

Moreover, indicate with $a_j$ the $j$-th column of $A$ and with $c_j$ the corresponding coefficients in the row vector $c$ to define the so-called *reduced cost* of the $j$-th column $\rho_j = c_j - c'(A')^{-1}A_j$.

The reduced cost $\rho_j$ is the increase in the objective function per unit of increase of the variable $q_j$ that is currently set to 0.

From this I get that if $\rho_j \geq 0$ there is no advantage in setting $q_j > 0$ and thus in introducing $A_j$ in the basis.

Actually, if $\rho_j \geq 0$ for every $j$ such that $q_j = 0$ then the current solution is the optimal one since no new column can be substituted into the current basis to reduce $z$.

On the contrary, any column $A_j$ corresponding to $\rho_j < 0$ is a legitimate candidate entry in the basis. A common approach to maximize the gain in introducing a new column, and thus to minimize the number of steps taken to update any starting basis to the optimal one, is to select the column $A_j$ that is the solution of the optimization problem

$$\begin{aligned} \min \quad & \rho = c_j - c'(A')^{-1}A_j \\ \text{s.t.} \quad & \begin{array}{l} A_j \text{ is a column of } A \\ A_j \text{ is not a column } A' \end{array} \end{aligned} \tag{4.5}$$

Regrettably, though the objective function of this subordinate problem is linear, its constraints need not be so.

In my case, for example, the vast majority of valid columns are of the kind $\xi = (x_0x_1, x_0x_2, \ldots, x_{n-2}x_{n-1})^\top$ whose corresponding entries in $c$ are zero. Hence, the optimization problem of choosing the best among these columns to be substituted into the current basis $A'$ becomes

$$\begin{aligned} \min \ \rho = -c'(A')^{-1} & \begin{pmatrix} x_0x_0 \\ x_0x_1 \\ \vdots \\ x_{n-1}x_{n-1} \end{pmatrix} \\ & x_0, \ldots, x_{n-1} \in X \\ \text{s.t.} & \begin{pmatrix} x_0x_0 \\ x_0x_1 \\ \vdots \\ x_{n-1}x_{n-1} \end{pmatrix} \notin A' \end{aligned} \tag{4.6}$$

that is a quadratic problem (QP) with discrete constraints $x_j \in X$.

Regrettably, even the simplest version of such optimization problems, i.e., the binary QP (BQP), is known to be NP-complete and thus poses serious problems when $n$ increases. A common way to reduce the complexity of this kind of problems is a translation into a linearized form.

To do so, set $x_j = 2s_j - 1$ for $j = 0, \ldots, n-1$ and $s_j \in \{0, 1\}$. Moreover, consider the additional variables $y_{j,k} \in [0, 1]$ to write

$$
\begin{aligned}
\min -c'(A')^{-1} & \begin{pmatrix} 1 \\ 4y_{0,1} - 2s_0 - 2s_1 + 1 \\ 4y_{0,2} - 2s_0 - 2s_2 + 1 \\ \vdots \\ 4y_{n-2,n-1} - 2s_{n-2} - 2s_{n-1} + 1 \end{pmatrix} \\
& s_j \in \{0, 1\} \ \text{ for } j = 0, \ldots, n-1 \\
& y_{j,k} \geq 0 \ \text{ for } 0 \leq j < k < n \\
\text{s.t.} \quad & y_{j,k} \geq s_j + s_k - 1 \ \text{ for } 0 \leq j < k < n \\
& y_{j,k} \leq s_j \ \text{ for } 0 \leq j < k < n \\
& y_{j,k} \leq s_k \ \text{ for } 0 \leq j < k < n
\end{aligned}
$$

Actually, it is not surprising that somewhere in the solution of the problem I must face the problem of selecting the good choice out of $2^n$ possible choices. In any case even if the "hard" part of the problem is here I may probably address up to few tens of bits instead of limiting to 16. Actually a set of heuristic algorithms are under investigation and now it is possible to arrive at sequence with $n = 100$.

The description of these algorithms is outside the scope of this discussion and so it will be omitted. Some results about a real application of this generator can be found in Section 4.4, where it is used to generate antipodal random vectors in according to the output of 3.15 - 3.16 when simple images are acquired.

A full description of 2vRVG will be reported in future communication; if the reader is interested in I invite him or her by email at *mmangia@arces.unibo.it*.

## 4.3 Acquisition of ECGs

The ECG time shape represents the voltage between two different electrodes placed on the body at two specific positions. It records the electrical field produced by the myocardium, i.e., for each heart beat it cyclically reports the successive atrial depolarization/repolarization and ventricular depolarization/repolarization.

The application of CS techniques to ECG acquisition has been the topic of recent contributions [42] [43] [44] [25] aimed to either reducing the amount of data needed to represent the signal in mobile applications or to achieve an high compression ratio in data storage systems.

In order to demonstrate the effectiveness of rakeness-based design for CS of ECGs, I need a broad collection of realistic realizations. To achieve this goal, I used a synthetic generator of ECGs, thoroughly discussed in [20] that provides signals not corrupted by noise to which I add white Gaussian noise with suitable power. The amount of noise is chosen so that the considered environment is realistic, but it can be arbitrary from the point of view of rakeness-based design that is independent of the noise level.

The generator core is expressed by the following set of three coupled ordinary differentially equations [20]

$$\begin{cases} \dot{x}_1 = \omega_1 x_1 - \omega_2 x_2 \\ \dot{x}_2 = \omega_1 x_1 + \omega_2 x_2 \\ \dot{x}_3 = -\sum_{i \in \{P,Q,R,S,T\}} y_i \Theta_i \exp\left(-\frac{\Theta_i^2}{2\upsilon_i^2}\right) - (x_3 - \bar{x}_3) \end{cases} \tag{4.7}$$

Each heart beat is represented by a complete revolution on an attracting limit cycle in the $(x_1, x_2)$ plane. The

**TABLE 4.2:** Parameters bounds used in the ECG generator.

| Index | P | Q | R | S | T |
|---|---|---|---|---|---|
| $\Theta_i$ | -75;-65 | -20;-5 | -5;5 | 10;20 | 95;105 |
| $\gamma_i$ | 1;1.4 | -5.2;-4.8 | 27;33 | -7.7;-7.3 | 0.5;1 |
| $\upsilon_i$ | 0.05;0.45 | -0.1;0.3 | -0.1;0.3 | -0.1;0.3 | 0.2;0.6 |

shape of ECG signal is obtained introducing five attractors/repellors points in the $x_3$ direction in correspondence to the peaks and valleys that characterize the time shape of the signal and which are conventionally labeled by P, Q, R, S and T; furthermore $\bar{x}_3$ in (4.7) represents the mean value of the generated ECG.

In order to mimic the behavior of ECGs in patients affected by the most studied cardiac illness, the parameters $\gamma_i$, $\upsilon_i$ and $\Theta_i$, $i \in \{P,Q,R,S,T\}$, characterizing each considered signal are taken from a set of random variables uniformly distributed within the bounds reported in Table 4.2. In addition, I randomly set the heart rate between 50Hz and 100Hz by property adjusting $\omega_1$ and $\omega_2$.

Though CS methods are classically developed for sparse representation with respect to signal bases, they have a straightforward generalizzation to sparse representation with respect to dictionaries, i.e., redundant collections of non-indipendent waveforms [45] [46].

This is, in fact, the case of ECGs, for which a dictionary made of Gabor atoms

$$g_{s,u,v,w}(t) = \frac{1}{\sqrt{s}} e^{-\pi\left(\frac{t-u}{s}\right)^2} \cos(vt + w)$$

can be used [47] [26].

In my experiment, a total of 507 atoms are used corresponding to different quadruple of parameters $(s, u, v, w)$.

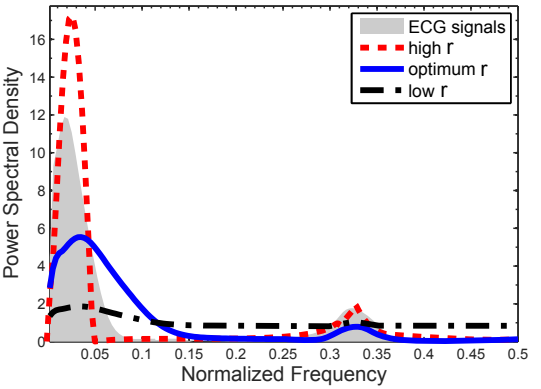



**Figure 4.3:** Average spectra of real ECG signals (gray area) and of the sampling PAM sequences corresponding to the optimum (solid line) as well as an high (dashed line) and a low value (dash-dotted line) of $r$.

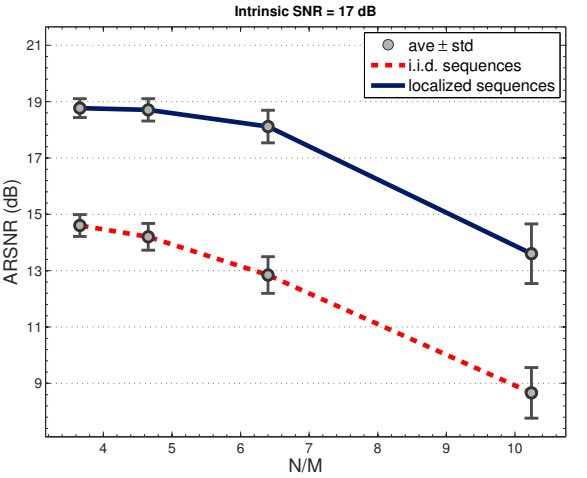



**Figure 4.4:** Average value of the reconstructed SNR (ARSNR) as a function of the signal compression ratio N/M between the number of Nyquist samples and of CS measures. The dashed line refers to i.i.d. sampling waveforms and the solid line to rakness-optimized ones.

This collection of Gabor atoms is obtained using a greedy algorithm able to extract a limited number of functions from a broader set [26]. With respect to this dictionary the sparse representation of a typical ECG heartbeat waveform requires about 14 non-zero coefficients. Furthermore, in my simulations, $T = 1$s within which the signal is sampled $N = 256$ times (a common choice for ECG equipments).

To apply the results in Section 3.4, I first compute the average of the power spectral densities on a huge amount of ECGs obtained by a famous database of biosignals [48]. By this procedure I obtain $\hat{a}(f)$, the input of the optimization problem (3.8). The shape of $\hat{a}(f)$ is the gray profile shown in Figure 4.3. Next, I find the optimum $r$ as described at the beginning of this Chapter. Figure 4.3 shows the optimum profile for the best value $r = 0.038$ (solid line) as well as for a smaller value (dash-dotted line) and for a larger value (dashed line).

Finally the LPF generator mentioned in Section 4.1 is used to produce the antipodal sequences with the optimized spectral profiles. These sequences are used to take $M$ measurements in a time window of length T, to which I add white Gaussian noise to construct the measurement vector $\underline{m}$ according to (3.2).

To determine the performance of rakeness based design, I consider a test set of 2000 synthetic ECG signals. These signals are acquired by projecting them both on localized antipodal sequences and on i.i.d. antipodal sequences (classically employed in CS-based methods and my reference case). The resulting ARSNRs are shown in Figure 4.4 as a function of the ratio between the intrinsic dimension of the signal and the number of CS measures.

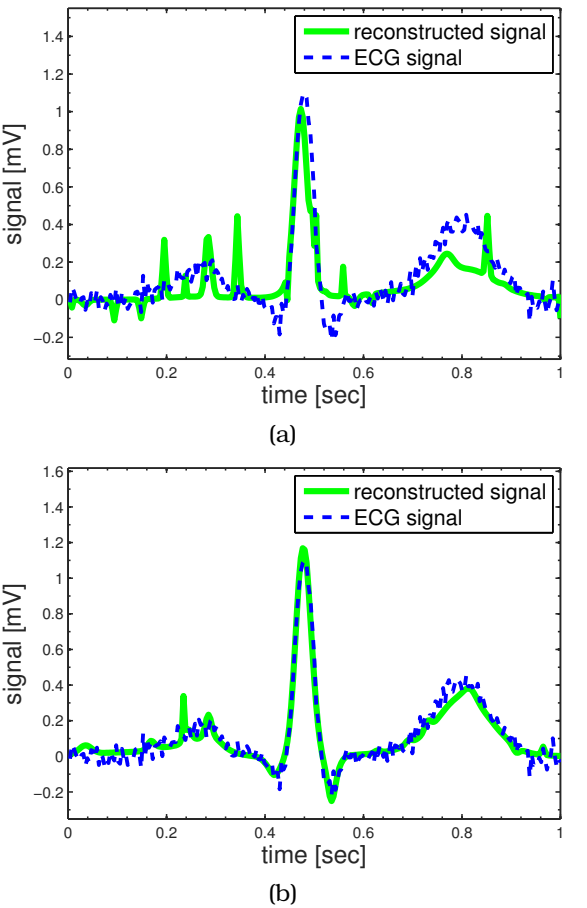



**Figure 4.5:** Original (solid line) and reconstructed (dashed line) ECG when i.i.d sampling waveforms are used (plot (a)) and when rakeness-optimized sequence are exploited (plot (b)). In both cases N=256 and M=32 and the intrinsic SNR=17dB.

In both cases the intrinsic SNR is equal to 17dB.

As it can be noticed, rakeness-based design allows to achieve an improvement of at least 3.5dB in ARSNR with respect to the i.i.d. case, and even yields denoising (i.e. and ARSNR larger than the intrinsic SNR) for small compression ratio values. To give a visual representation of the improvement, Figure 4.5 reports, for $M = 32$ (compression ratio equal to 8 respect to a classic ADC at Nyquist rate), a comparison between an ECG signal and the reconstructed one for the i.i.d. (a) and rakeness-based (b) case. Direct visual inspection is enough to confirm the superiority of my approach.

## 4.4 Acquisition of small images

In this second case the signal to acquire is a $24 \times 24$-pixel image, each pixel value ranging from 0 (black) to 1 (white), which represents a small white printed number or letter on a black background with a gray-level dithering to make the curves smoother to the human eye. Number and letters are randomly rotated and offset from the center of the image but never clipped.

Although due to random rotations and offsets almost all pixels have a non-vanishing probability of being non-zero, a typical image contains only about 85 bright pixels, so that can be considered sparse in the base of 2-dim discrete delta functions that evaluate to 1 at a single pixel position and zero elsewhere. I may thus think of acquiring them using a RMPI architecture that projects along $24 \times 24$ antipodal random grids to obtain measurements that are enough to reconstruct the image but whose number $M$ is

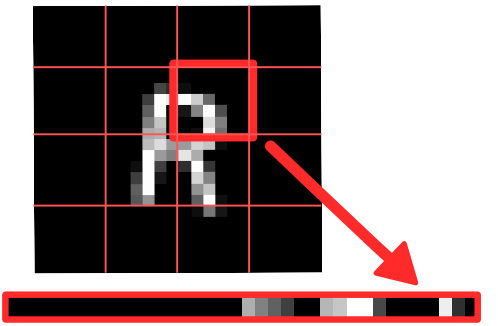


**Figure 4.6:** A sample image, its partition and rearrangement into a vector containing the value of each pixel.

much less than the number $N = 24 \times 24 = 576$ of the original pixels.

To simplify the design phase, the generation of the random grids is done by adjoining $4 \times 4$ subgrids each with $6 \times 6$ antipodal values whose statistic is optimized by solving (3.12).

To allow calculations, the values in each subgrid are rearranged into a 36-dimensional vector as schematically reported in Figure 4.6, that also highlights the subgrid on which I will focus in the following.

In that region, and due to the vector rearrangement, I may list the modulating symbols of the projection waveform $b$ with $b_j$ for $j = 0, \ldots, 35$. The same can be done for the incoming signal $a$ when it is expressed along the basis of 2-dim discrete delta's with coefficients that I may indicate with $a_j$ for $j = 0, \ldots, 35$.

If $\underline{a} = (a_0, \ldots, a_{35})^\top$, I may follow the development of subsection 3.5.1 to estimate the $36 \times 36$ matrix $\underline{A} = \mathbf{E}[\underline{a}\underline{a}^\top]$ by empirical averaging over a training set of 2080

randomly generated images.

The resulting matrix is reported in graphic form in Figure 4.7-(a) where, for each pair of indexes $j, k = 0, \ldots, 35$, a point is laid down whose brightness is proportional to the values of $\underline{A}_{j,k} = \mathbf{E}[a_j a_k]$.

The eigenvalues $\mu_0, \ldots, \mu_{35}$ of that $\underline{A}$ are reported as the light bars in Figure 4.7-(c). By exploiting (3.16) and (3.17) for $r = 0.047$ I get $J = 36$ and the eigenvalues $\hat{\mu}_0, \ldots, \hat{\mu}_{35}$ reported as dark bars in figure 4.7-(c).

From the eigenvectors of $\underline{A}$ and these new eigenvalues, I may construct the correlation matrix of the values in this projection subgrid. Since the subgrid contains $6 \times 6 = 36$ values, its correlation matrix $\underline{B}$ has dimensions $36 \times 36$ and is reported in Figure 4.7-(b) in a graphical form adopting the same convention used to represent the values of $\underline{A}$.

Once that $\underline{B}$ is known, I can use the 2vRNG to generate the antipodal grids needed in the acquisition procedure.

The same design process is repeated for each of the 4 central $6 \times 6$ regions in the image while the 12 outer subgrids are built from independent and uniformly distributed antipodal symbols. All the subgrids are finally compounded in a complete $24 \times 24$ projection grid.

As a comparison case, projections are also taken by using i.i.d. symbols for all the elements of the projection grid.

In both cases, noise is added to the projections before they take their place in the vector $\underline{m}$ of $M$ measurements according to (3.2), and reconstruction is performed using the algorithm reported in [12].

Figure 4.8 reports the ARSNR (over 3000 trials) of the reconstructed images and compares the performance of

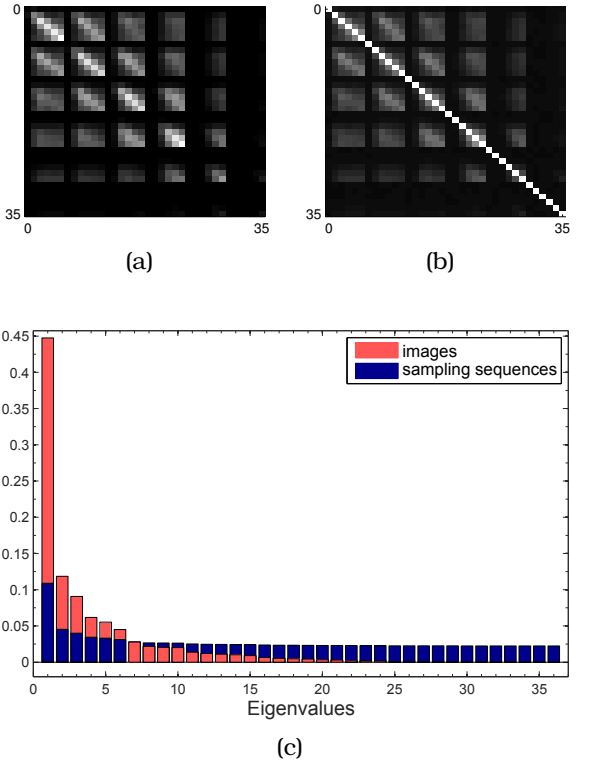



**Figure 4.7:** Correlation matrix of the pixels in one of the four central regions (a) Correlation matrix of the optimal projection process (b) Eigenvalues of the above correlation matrices (c).

an RMPI based on rakeness-optimized projections and on i.i.d. projections for different values of the compression ratio $N/M$. In both cases the intrinsic SNR is 17dB.

It is evident from Figure 4.8 that, even if it is exploited only in the central portion of the images, rakeness-based design leads to non-negligible improvement of at least 1dB.

A qualitative appreciation of such an improvement can be obtained from Figure 4.9 in which 5 images (a) are

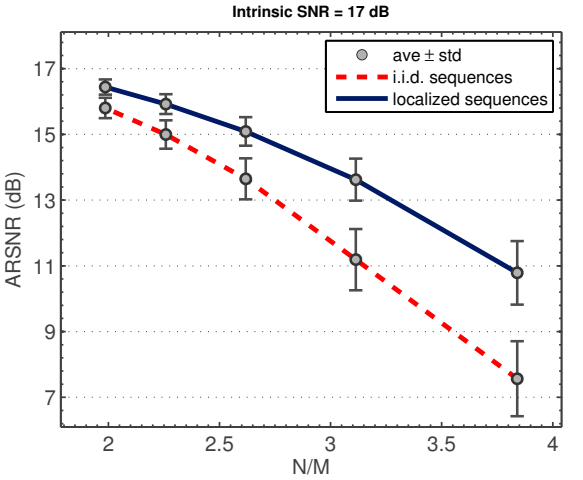



**Figure 4.8:** Quality of the reconstructed images when rakeness-optimized or i.i.d. projection grids are used in an RMPI architecture for different compression ratios.

acquired and reconstructed by means of $M = 115$ over $N = 576$ rakeness-optimized projections (b) or by the same number of i.i.d. projections (c). Reconstruction artifacts are visibly reduced by adopting rakeness-based design.

## 4.5 Conclusion

Compressive sensing exploits the fact that, when looked at in the right domain, the information content of a signal can be much less than what appears when I look at it in time or frequency (i.e., the signal is sparse).

Acquisition schemes that exploit sparsity may lead to considerable advantages in terms of sensing system design since, for example, if the information content is much less than the signal bandwidth, sub-Nyquist

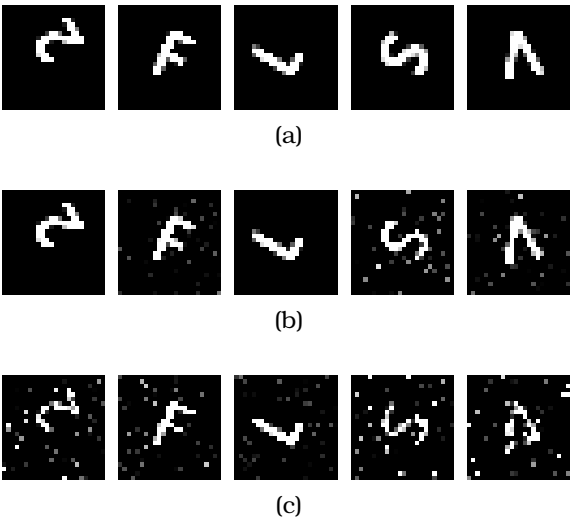


**Figure 4.9:** Sample images (a) and their reconstruction based on rakeness-optimized projection grids (b) or on i.i.d. projection grids (c).

sampling can be employed.

To all this I add the consideration that, in a possibly different domain, the energy of the signals may be not uniformly distributed (i.e., the signal is localized) and when noise is present, it is convenient to adapt the system to "rake" as much signal energy as possible.

By itself, this is not a novel concept since it appears, for example, in matched filters and rake receivers used in telecommunications. Yet, in my context, the efforts to collect the energy of the signal must be balanced with the guarantee that all details of its underlying structure can be captured when immersed in noise. This brings my to a trade-off that I propose to address in statistical terms by means of an optimization problem: maximize the

"rakeness" while obeying to a constraint ensuring that the measurements are random enough to capture all signal details.

The topic presented in both this chapter and previous one develops the formal definition of such problem as well as its solution for stationary signals whose localization can be highlighted in the frequency domain, and for more generic non-stationary signals whose localization is more evident in suitably defined domains.

The applicability of both techniques is demonstrated by sample applications to the acquisition of ECG tracks and small letter images.

CHAPTER

# 5 Rakeness and new architectures

REGARDING news in CS, an innovative approach was introduced by Mamaghanian *et al* [49]; The main novelty is a new encoding strategy: the authors propose a modified version of the well known RMPI architecture, discussed in Chapter 2, named Spread Spectrum RMPI (S-RMPI), which is characterized by a pre-modulation of the input signal before the classical sensing stage. This step was introduced to reduce the power consumption by enabling low frequency switching in the sensing stage [1].

In this chapter, I combine this innovative encoding procedure with rakeness to obtain a further switching frequency reduction and, at same time, to lower the amount of projections needed to reach a properly reconstructed signal. As a test signal for this approach I use electrocardiograms (ECG), i.e., the same used when S-RMPI was presented [49], since they represent a good test class for possible applications in the area of wireless body

[1] it is possible to halve the internal switching frequency.

sensor networks [25].

It is important to highlight that this approach can be very useful in all applications that involve fast switching, such as RF applications. In fact, it yields a lower switching rate with respect to classical RMPI.

The Chapter is organized as follow: Sec. 5.1 reports the mathematical model and presents method I propose to combine S-RMPI with rakeness. Sec 5.2 describes the considered simulation setting, while Sec. 5.3 reports results obtained by Montecarlo simulations. After that, some conclusions end the Chapter.

## 5.1 Mathematical Models

The encoding strategy of an AIC based on CS is related to a particular chosen architecture. The one considered here is the S-RMPI, where all involved signals are referred to a generic time window of length $T$. The block diagram is shown in Figure 5.1-(a) where the input signal $x$ is represented by a set of samples $x \in \mathbb{R}^N$ acquired at Nyquist rate which is modulated by a spreading sequence ($SS$) expressed as a set of $n$ random antipodal symbols ($n \geq N$). The obtained signal is input to $M$ different branches, all composed by random demodulators (RD) [4, 49].

Each RD is characterized by a specific sensing waveform $\phi_j$ ($j = 1, \ldots, M$) made by a sequence of $m$ antipodal random values, with $m \leq N \leq n$. The output of the S-RMPI is the measurements vector $y \in \mathbb{R}^M$ sampled at frequency $1/T$. Remember that $N/T$ represents the sampling rate of the input signal, acquired at Nyquist rate, while SS and $\phi_j$ are characterized by chipping frequencies $n/T$ and $m/T$

respectively.

From a mathematical point of view, if I impose $n = m = N$ then I can write $y = \Phi S x = F x$, where $\Phi \in \mathbb{R}^{M,N}$ represents the sensing matrix composed row by row with the sensing sequences $\phi_j$ and $S \in \mathbb{C}^{N,N}$ is a diagonal matrix whose non-null elements are the $SS$ coefficients. The matrix $F = \Phi S$ represents the linear operator that links the input signal to the measurements vector.

In the more general setting when $m \neq N \neq n$ I can rearrange the model differently. I will do so with the final goal of reducing the operating frequency of all switches used in a possible implementation of this architecture. Therefore, I need to reduce both $m$ and $n$ as much as possible, noting that $n \geq N$ to maintain the signal information before the sensing stage and $m < n$, which corresponds to an undersampling factor $r = \frac{n}{m}$ w.r.t. the Nyquist frequency. With this, mathematical formulation for the measurements is rearranged as follows:

$$y_i = \sum_{j=1}^{n} S_{j,j} \Phi_{i,\lceil \frac{j}{r} \rceil} x_j$$

In this case I can write again $y = F \cdot x$ where $F \in \mathbb{R}^{M,n}$ and the generic element of $F$ is equal to

$$F_{i,j} = S_{j,j} \Phi_{i,\lceil \frac{j}{r} \rceil}$$

This approach is equivalent to the classical RMPI architecture where the sensing matrix is $F$ instead of $\Phi$. For this kind of model the rakeness optimization problem can be solved as discussed in Chapter 3. The result of this procedure, in the more generic case, is the correlation matrix $C_F$ of the stochastic process used to generate the rows of $F$. Now, the aim will be to impose $C_F$ by fixing

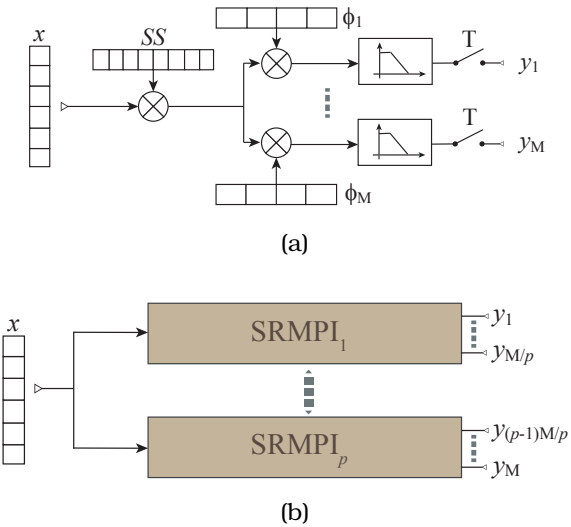



**Figure 5.1:** block diagram of spread spectrum random modulation pre-integration (a) and multi spread spectrum random modulation pre-integration (b).

the correlation matrix of the process used to generate the rows of $\Phi$ and $SS$.

Forcing the correlation of the product of instances of two stochastic processes is a very hard topic and it is outside the scope of this thesis, so I considered a more pragmatic approach. In the following $\Phi$ will be taken as a collection of i.i.d. antipodal random variables and the attention will focus on the the statistic of $SS$.

Let me indicate with $C_{F_{\{j,k\}}}$ the correlation between two elements of $F$ in the same row and at columns $j$ and $k$;

this is equal to:

$$C_{F_{\{j,k\}}} = \mathbf{E}\left[F_{\cdot,j}F_{\cdot,k}\right] = \mathbf{E}\left[\Phi_{\cdot,\lceil\frac{j}{r}\rceil}\Phi_{\cdot,\lceil\frac{k}{r}\rceil}S_{j,j}S_{k,k}\right] =$$

$$= \begin{cases} \mathbf{E}\left[S_{j,j}S_{k,k}\right] & \lceil\frac{j}{r}\rceil = \lceil\frac{k}{r}\rceil \\ \mathbf{E}\left[\Phi_{\cdot,\lceil\frac{j}{r}\rceil}\Phi_{\cdot,\lceil\frac{k}{r}\rceil}S_{j,j}S_{k,k}\right] & \lceil\frac{j}{r}\rceil \neq \lceil\frac{k}{r}\rceil \end{cases} \tag{5.1}$$

since $\Phi_{\cdot,j}\Phi_{\cdot,j} = 1$ due to the antipodality of the $\Phi_j$.

In the next subsection, I investigate a possible way to design $F$ with statistical properties which well match those required by the solution of the rakeness optimization problem. This will be obtained by tuning the statistic of the $SS$ symbols and fixing the statistic of $\Phi$ as mentioned above.

### 5.1.1 S-RMPI based on Rakeness

Considering $\Phi$ composed by i.i.d antipodal sequences, (5.1) can be rewritten as follows:

$$C_{F_{\{j,k\}}} = \begin{cases} \mathbf{E}\left[S_{j,j}S_{k,k}\right] & \text{with:} \quad \lceil\frac{j}{r}\rceil = \lceil\frac{k}{r}\rceil \\ 0 & \text{with:} \quad \lceil\frac{j}{r}\rceil \neq \lceil\frac{k}{r}\rceil \end{cases} \tag{5.2}$$

where the last equality holds since $F_{i,j} = S_{j,j}\Phi_{i,\lceil\frac{j}{r}\rceil}$ is an i.i.d random variable for all the $j,k$ couples that verify $\lceil\frac{j}{r}\rceil \neq \lceil\frac{k}{r}\rceil$. Following my aim, I fix the $SS$ correlation equal to the one imposed by the solution of rakeness optimization problem and as final result the elements on each row of $F$ have the desired correlation only in the cases where $\lceil\frac{j}{r}\rceil = \lceil\frac{k}{r}\rceil$.

To preserve most of correlation profile it is necessary to work with high $r$ values, but this can compromise the *diversity* in the sensing, i.e., it may increase the probability that two or more quite similar row instances in

$F$ exist. Hence, increasing $r$ means destroying the ability of sensing sequences to collect information in whole signal domain. On the other hand, with low $r$ values the lost spanning ability of the sensing sequences is compensated by the spreading sequence effect but this results to be insufficient for $r$ greater than 2 (see [49] for more details).

A toy case is reported in Figure 5.2. The left image represents the estimated correlation matrix of the process used to generate the spreading sequences ($n = 16$). As sensing sequences I considered i.i.d sequences with $r = 4$, which corresponds to a $4 \times 4$ correlation matrix (the identity matrix, as shown in the center image). Here I also report the estimated correlation of the product between sensing and spreading sequences (the right image) which confirms the presented formulation of (5.2). The correlation of $F$ is a block diagonal matrix with $m$ different $r \times r$ blocks in the diagonal, which contain the same values of the correlation matrix of the stochastic process used in the spreading sequences generation.

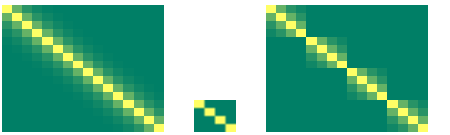



**Figure 5.2:** $16 \times 16$ spreading sequences correlation matrix (a); $4 \times 4$ i.i.d. sensing sequences correlation matrix (b); $16 \times 16$ $F$ correlation matrix (c). Note that dark means zero and bright mean one.

### 5.1.2 Multi S-RMPI based on Rakeness

To cope with the *diversity* reduction imposed by S-RMPI based on high $r$ values ($r > 2$) I propose a Multi Spread Spectrum RMPI (MS-RMPI) approach. The idea is to use more than one spreading sequence to preserve the diversity during acquisition. In this way, the probability that similar rows of $F$ occur is decreased and thus the MS-RMPI is able to collect more signal information in the spreading stage. The block diagram of MS-RMPI is reported in Figure 5.1-(b). It is composed by $p$ different S-RMPI blocks, where: *i*) the $S$s are generated with the same statistic, *ii*) each block produces $M/p$ measurements which can be collected to yield the measurement vector $y$.

In this paper the improvement associated to MS-RMPI is shown only by numerical simulations, while a deeper theoretical analysis on the impact of diversity in the presented architectures is left for future communications.

## 5.2 Simulation Setting

As mentioned in before, I use ECGs as the designed class of signals to test the presented method. It is known that ECG are sparse both in the discrete wavelet basis (DWB) [49] and on a dictionary of Gabor Functions (GF) [16, 26]. When S-RMPI and MS-RMPI are used for encoding, the reconstruction based on GF presents a tiny performance improvement with respect to DWB, so the following results will be related to the GF case only.

The used signal instances are synthetic ECGs generated by [20] according to the setting described in Section 4.3, sampled at 256 Hz with a time window equal to 1

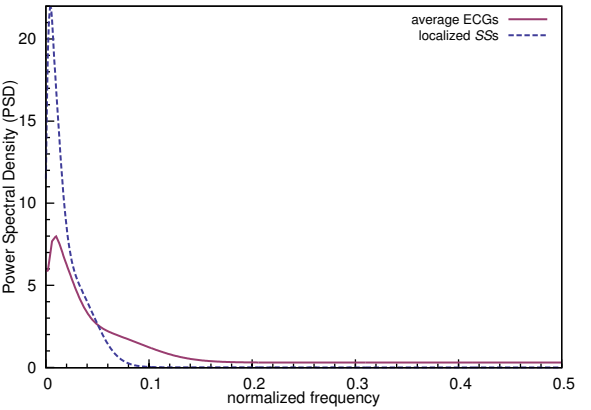



**Figure 5.3:** average of normalized power spectral density of ECG signals (dashed curve) and power spectral densities obtained by solution of rakeness optimization problem (solid curve).

second. The rakeness optimization problem is solved in the frequency domain by assuming that the ECG process is cyclostationary. Following the procedure in Section 3.4 one is able to match the ECG spectral profile with that of the sequence that maximizes the information obtained by CS.

These two spectral profiles are shown in Figure 5.3: the dashed line represents the average PSD of the synthetic ECGs and the solid one is the output of the rakeness optimization problem. It is important to note that the average PSD of ECGs is close to zero in the high frequency range, but this does not mean that instances do not present high pass components, but only that the probability of such an event is low. Rakeness copes with this feature by imposing a non-zero PSD of the spreading sequences also at high frequencies.

In all cases I focus the attention both on i.i.d SSs [49]

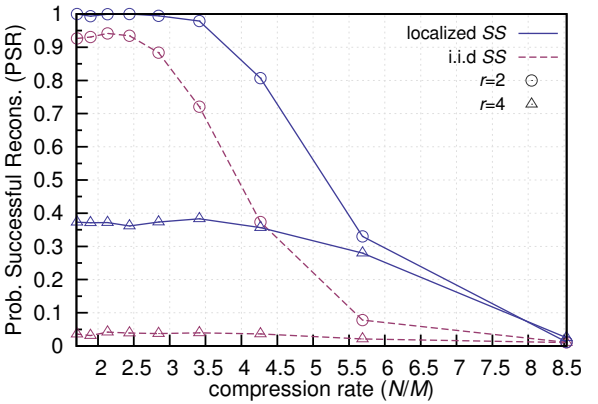



**Figure 5.4:** PSR as function of compression rate ($N/M$) where the encoding is done by S-RMPI with i.i.d. $SS$s (dashed lines) and localized $SS$s (solid lines) for $r = \{2, 4\}$.

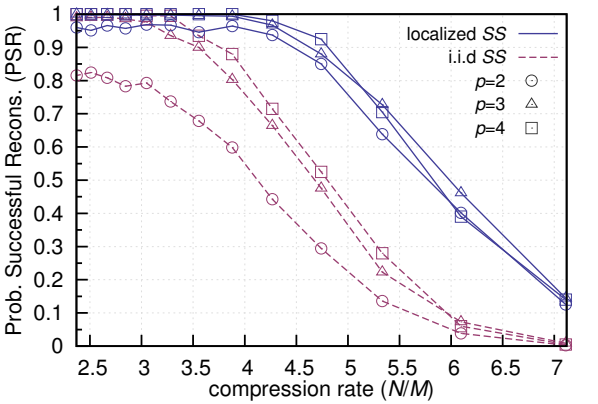



**Figure 5.5:** PSR as function of compression rate ($N/M$) where the encoding is done by MS-RMPI with i.i.d. $SS$s (dashed lines) and localized $SS$s (solid lines) for $r = 4$ and $p = \{2, 3, 4\}$.

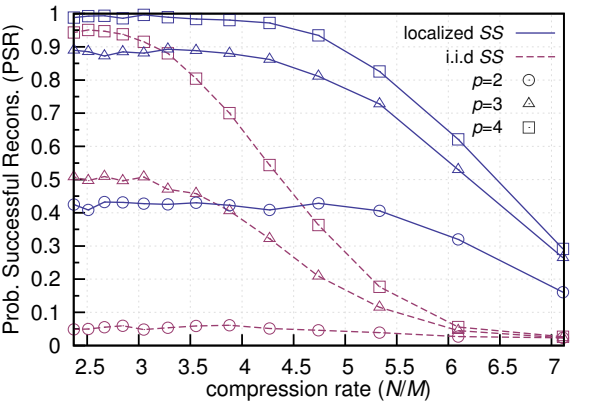



**Figure 5.6:** PSR as function of compression rate ($N/M$) where the encoding is done by MS-RMPI with i.i.d. *SS*s (dashed lines) and localized *SS*s (solid lines) for $r = 8$ and $p = \{2, 3, 4\}$.

(my reference case) and on localized *SS*s generated by a LPF [33–35] where the PSD is the one imposed by the rakeness optimization problem as mentioned before [16]. The aim is to numerically verify improvements in term of chipping frequency reduction, i.e., either a reduction of $m$ or an increase of $r$, introduced by the combination of spread spectrum CS and rakeness. To do this I performed Montecarlo simulations over 2000 trials.

Coherently with the introduced setting, in all considered instances the original signal was reconstructed by solving the optimization problem

$$a_r = \arg\left\{\max_a \|a\|_{l_1}\right\} \quad s.t \quad y = \Phi\Psi a$$

so that reconstructed signal is $x_r = \Psi \cdot a_r$, where $\Psi$ is the GF set. All optimization problems were solved by using the Cplex optimization toolbox [50].

The chosen figure of merit is the probability of successful reconstruction (PSR), defined as PSR = $Pr(\|a\| \geq$

$20\|a - a_r\|$).

## 5.3 Results

First the S-RMPI was simulated with $r = \{2, 4\}$. The related results are shown in Figure 5.4. When $r = 2$ the system is close to completely destroying the imposed correlation and so the improvement introduced by rakeness w.r.t. the reference architecture is limited. Furthermore to reach PSR=1 rakeness is needed. On the other hand, when $r = 4$ I note a higher improvement introduced by rakeness w.r.t. independent spreading sequences but PSR=1 is never reached.

As discussed in Section 5.1.2, MS-RMPI was introduced to guarantee a correct reconstruction associated to values of $r$ greater than 2. To verify this assumption, at first the MS-RMPI was tested by simulation for $r = 4$.

The PSR trends are shown in Figure 5.5 which clarifies the advantages introduces by this new approach. Here I am able to reach PSR=1 using $p = 3$, i.e., three different spreading sequences, in both considered environments either with i.i.d or localized sequences, but localized sequences reach PSR=1 with a compression rate $\leq 4$ while i.i.d. sequences reach PSR=1 only for compression rate $\leq 2.5$.

The case $r = 8$ was also analyzed, which corresponds to $m = 32$ and implies a strong reduction of the chipping frequency in the sensing. The results in terms of PSR are shown in Figure 5.6 which shows similar improvement as Figure 5.5 for $r = 4$.

## 5.4 Conclusions

I demonstrated that combining rakeness with S-RMPI produces a strong reduction of the internal chipping frequency in the sensing if a multi spread architecture is used. In the last presented group of results MS-RMPI reaches PSR=1 by means of 4 SS at Nyquist rate and 85 projections done with an internal frequency reduced by a factor 8 with respect to Nyquist. In the same setting, classical RMPI arrives to PSR=1 only by more than 100 projections, all of them computed at Nyquist rate[2].

[2] classical RMPI needs a minimum amount of projections related to the sparse level of the input signal $K$ by the low $M > CK\log_{10}(N/K)$, where $C$ is a constant usually equal to 4 or 5. For ECG sparse on GF and in the considered simulation setting I have $K$ in average greater then 20, so $M$ is around 100.

PART

# II

# UWB System based on DS-CDMA, Narrowband Interference Reduction

CHAPTER

# 6 Introduction to DS-CDMA

THIS chapter, coupled with the next one, presents a way to cope with the need of simultaneously rejecting narrowband interference and multi-access interference in a UWB system based on direct-sequence CDMA. With this aim in mind, I rely on a closed-form expression of the system bit error probability in presence of both effects. By means of such a formula, I evaluate the effect of spectrum shaping techniques applied to the spreading sequences. The availability of a certain number of degrees of freedom in deciding the spectral profile allows us to cope with different configurations depending on the relative interfering power but also on the relative position of the signal center frequency and the narrowband interferer.

## 6.1 Introduction

Ultra Wide Band (UWB) systems play a key role in the efforts devoted to the design of the next generation

communication infrastructure since they may help in addressing the basic need of a connectivity that should be, in principle, independent on the environment and as little supervised as possible.

UWB systems employ signals that feature a very low power density spectrum in order to appear almost equivalent to channel natural disturbances. To meet such a low power density request, and at the same time, to deliver enough energy to the receivers, UWB systems fill large portion of the spectrum, thus almost surely overlapping with other systems or services, being them either wideband or narrowband.

Coexistence in such shared, unsupervised environment depends on the ability of narrowband systems to tolerate an increased noise floor (possibly increasing their power budget) and on the capability of the wideband systems to reject narrowband interference (possibly adapting their transmission scheme to time-varying scenarios).

The point addressed in this part of thesis is the second one, i.e., when the UWB signals are involved in a Direct-Sequence spread spectrum transmission with Multiple Access based on Code Division (DS-CDMA). Such a classical technique [1] may allow arbitrary spreading and is suitable for extremely simple implementations of both transmitter and receiver.

In our scenario, the narrowband interferer may be either an intentionally emitted jammer or the abstract subsumption of the effects of a traditional non-UWB service and, for the sake of simplicity, it is modeled as a single sinusoidal tone.

Among all the possible figures of merit that quantify the performance of communication (e.g., multipath

robustness, system capacity in terms of users number [51]) I concentrate on the joint effect of Multiple Access Interference (MAI) and NarrowBand Interference (NBI) on a simple Matched Filter (MF) receiver.

Our starting point is the result reported in [52] that expresses the performance of a conventional DS-CDMA system with no multi-access when it is disturbed by a jamming tone. Such a result is extended to take into account the presence of a noise-like disturbance due to other asynchronous users whose spreading codes cannot be perfectly orthogonal to the useful receiver one for all the possible time shifts. This disturbance has been also extensively investigated in classical DS-CDMA and UWB systems relying on possible spreading codes optimizations based on chaotic systems (piece-wise affine Markov maps; see [33, 34] for a survey on chaos-based sequences generation for various applications), both in Additive White Gaussian Noise (AWGN) channels [53, 54] and in presence of multipath [55,56]. The influence of real pulse shapes has been also taken into account [57–60], while the ultimate limit of these systems in terms of Shannon capacity is studied in [61, 62].

Here, I extend the results presented in [63] and show how the rejection of both narrowband interference and multi-access disturbance can be achieved by the proper spectrum shaping of the spreading codes towards different optimal profiles. Furthermore, we here also address the trade-off between the two optimal designs for different scenarios as far as the power of the narrowband jammer and the number of competing users are concerned. Different optimal solutions will be found depending on the relative interfering power but also on the relative positioning of

the signal center frequency and the jammed frequency.

## 6.2 System Model

The starting point of our analysis is the model of a standard asynchronous DS-CDMA system [1, 64] including transmission module, channel and receiver. Figure 6.1 shows a simplified baseband equivalent scheme including the effects of thermal channel noise (modeled as AWGN) and NBI, in addition to MAI. To evaluate system performance, let us assume the presence of a common carrier with frequency $f_c = \omega_c/(2\pi)$, let $U$ be the total number of users, and express the generic $u$-th information signal as

$$S^u(t) = V \sum_{s=-\infty}^{\infty} S_s^u g_T(t - sT) \tag{6.1}$$

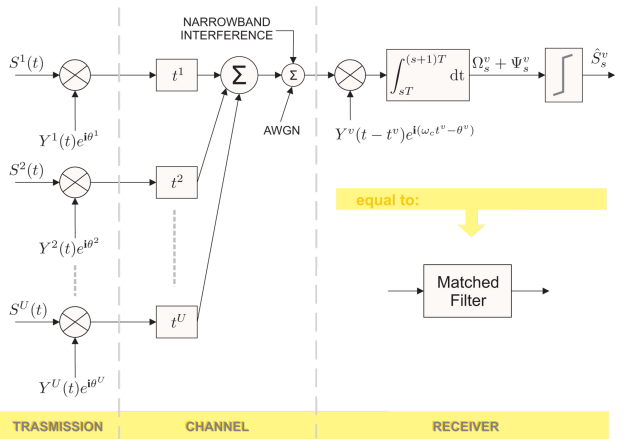



**Figure 6.1:** Block diagram of the baseband equivalent scheme of a UWB asynchronous DS-CDMA system.

where $T$ is the bit duration, $S_s^u \in \{-1; +1\}$ is the $s$-th information symbol for the $u$-th user, $g_T$ is a rectangular pulse which is 1 within $[0; T]$ and zero otherwise, and where the peak amplitude $V$ is assumed equal for all users. The signal $S^u(t)$ is then multiplied by the spreading signal

$$Y^u(t) = \sum_{s=-\infty}^{\infty} y_s^u g_{T/N}\left(t - s\frac{T}{N}\right) \tag{6.2}$$

where $N$ is the spreading factor, $y_s^u$ are the antipodal spreading symbols of the $u$-th user and $g_{\frac{T}{N}}$ is a rectangular pulse which is 1 within $[0; \frac{T}{N}]$ and 0 outside, and where, following the approach in [1] [64], I will assume that the corresponding Spreading Sequence (SS) $\mathrm{y}^u = \{y_s^u\}$ is periodic, with period equal to $N$.

The resulting signal is then transmitted along the (baseband equivalent) channel together with the spread-spectrum signals from the other users. Each transmitter adopts a different spreading code, assigned at the connection start-up. The receiver is a simple MF which is made of a multiplier combining the incoming signal with a synchronized replica of the spreading sequence of the $v$-th user and of an integrate-and-dump stage in charge of extracting the information symbol by correlation.

For each user I consider a different delay $t^u$ and a different phase $\partial^u$.These quantities are assumed as uniformly distributed random variables to model transmission from mobile terminals to a fixed base-station. If I take the $v$-th receiver as a reference, I can define the relative delays $\Delta t^{uv} = t^u - t^v$ and the relative phase $\Delta\partial^{uv} = \partial^u - \partial^u$.

We do this to arrive at the expression of the contribution of the $u$-th user to the signal produced by the $v$-th MF,

given by

$$\Upsilon_s^{uv} = \frac{1}{2T}\int_{sT}^{(s+1)T} S^u(t-\Delta t^{uv})Y^u(t-\Delta t^{uv})Y^v(t)e^{\mathbf{i}\Delta\vartheta^{uv}}dt \tag{6.3}$$

where $\mathbf{i}$ is the imaginary unit.

In the following I will assume that the system performance is mainly limited by MAI, thermal noise and NBI, namely I will suppose that spreading-despreading sequence synchronization has been achieved, and that multipath effects are negligible. Despite these simplifying assumptions, performance computation is quite difficult in this case, and, in order to obtain a suitable model, I rely on two previous results corresponding to "corner" cases where either NBI and thermal noise or MAI are present alone, as presented in Subsection 6.2.1 and Subsection 6.2.2, respectively.

### 6.2.1 Multiple Access Interference

We want to evaluate in terms of bit error probability the impact of MAI, and to this purpose, I define the *partial cross-correlation* function between two generic spreading sequences $\mathrm{y}^u = \{y_k^u\}$ and $\mathrm{y}^v = \{y_k^v\}$ as

$$\Gamma_{N,\tau}(\mathrm{y}^u,\mathrm{y}^v) = \begin{cases} \sum_{k=0}^{N-\tau-1} y_k^u y_{k+\tau}^v & \text{if } \tau = 0,1,\ldots,N-1 \\ \Gamma_{N,-\tau}(\mathrm{y}^v,\mathrm{y}^u) & \text{if } \tau = -N+1,\ldots,-2,-1 \\ 0 & \text{if } |\tau| \geq N \end{cases}$$

Exploiting the previous expression, I can use (6.3) to write the useful signal component as

$$\Omega_s^v = \Upsilon_s^{vv} = \frac{VS_s^v}{2T}\int_{sT}^{(s+1)T}[Y^v(t)]^2dt = \frac{VS_s^v}{2N}\Gamma_{N,0}(\mathrm{y}^v,\mathrm{y}^v) = \frac{VS_s^v}{2}$$

and the interfering signal component (which depends on MAI)

$$\Psi_s^v = \sum_{\substack{u=1 \\ u\neq v}}^{U} \Upsilon_s^{uv} =$$

$$= \sum_{\substack{u=1 \\ u\neq v}}^{U} \frac{1}{2T} \int_{sT}^{(s+1)T} S^u(t-\Delta t^{uv}) Y^u(t-\Delta t^{uv}) Y^v(t) e^{\mathbf{i}\Delta\partial^{uv}} dt \tag{6.4}$$

We can assume that the information symbols are independent identically distributed random variable. This brings us to the following expression [64] [65]

$$\Psi_s^v =$$

$$\frac{1}{2N} \sum_{\substack{u=1 \\ u\neq v}}^{U} V e^{\mathbf{i}\Delta\partial^{uv}} \left\{ S_{s-1}^u \left[ \left( \frac{\Delta t^{uv}}{T} - \left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor \right) \Gamma_{N,\left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor - N+1}(\mathrm{y}^u, \mathrm{y}^v) + \right.\right.$$

$$\left. \left( 1 - \frac{\Delta t^{uv}}{T} + \left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor \right) \Gam_{N,\left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor - N}(\mathrm{y}^u, \mathrm{y}^v) \right] +$$

$$+ S_s^u \left[ \left( \frac{\Delta t^{uv}}{T} - \left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor \right) \Gamma_{N,\left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor + 1}(\mathrm{y}^u, \mathrm{y}^v) + \right.$$

$$\left.\left. \left( 1 - \frac{\Delta t^{uv}}{T} + \left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor \right) \Gamma_{N,\left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor}(\mathrm{y}^u, \mathrm{y}^v) \right] \right\}$$

where, $\Delta\partial^{uv} \in [-\pi; \pi]$ and the normalized delay $\frac{\Delta t^{uv}}{T}$ has been split into a integer part $\left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor$ and a fractional part $\frac{\Delta t^{uv}}{T} - \left\lfloor \frac{\Delta t^{uv}}{T} \right\rfloor$. Since the interference term is a sum of many zero-mean independent random variables, I can model it as zero-mean Gaussian random variable, whose variance

is equal to [64] [65]

$$(\sigma^{v})^2 = \mathbf{E}[(\Psi_s^{v})^2] =$$

$$= \sum_{\substack{u=1 \\ u\neq v}}^{U} \frac{(V)^2}{24N^3} \sum_{j=-N+1}^{N-1} \left( 2\Gamma_{N,j}^2(\mathrm{y}^{u},\mathrm{y}^{v}) + \Gamma_{N,j}(\mathrm{y}^{u},\mathrm{y}^{v})\Gamma_{N,j+1}(\mathrm{y}^{u},\mathrm{y}^{v}) \right)$$

where the expectation $\mathbf{E}[\cdot]$ is taken over the phase $\Delta\vartheta^{uv}$, the delay $\Delta t^{uv}$ (that is assumed uniformly distributed in $[-T;T]$) and the information symbols $S_s^u$ (that are assumed independent and equally distributed, so that $S_s^u = -1$ and $S_s^u = +1$ have the same probability).

MAI for the $v$-th link is proportional to $(\sigma^{v})^2$, so that I can evaluate the performance in terms of bit-error probability for the $s$-th transmitted symbol as $P_{err}^{v} = \frac{1}{2}\mathrm{erfc}\sqrt{(\Omega_s^{v})^2/[2(\sigma^{v})^2]} = \frac{1}{2}\mathrm{erfc}\sqrt{V^2/[8(\sigma^{v})^2]}$. The overall system performance can be obtained by averaging over all possible useful users $v$ by defining

$$\mathrm{BE}P_{MAI} = \mathbf{E}_{\mathrm{y}^{v}}[P_{err}^{v}] = \frac{1}{2}\,\mathbf{E}_{\mathrm{y}^{v}}\left[\mathrm{erfc}\sqrt{\frac{V^2}{8(\sigma^{v})^2}}\right]$$

where $\mathbf{E}_{\mathrm{y}^{v}}[\cdot]$ is the expectation over the set of sequences. Analytical handing is difficult due the presence of the erfc function. A common workaround to this is to use the Standard Gaussian Approximation (SGA) which considers the spreading codes values as random variables (see [1] and [66] for an in-depth analysis of this approximation) to simplify the above expression in

$$\mathrm{BE}P_{MAI} \approx \frac{1}{2}\mathrm{erfc}\sqrt{\frac{V^2}{8\,\mathbf{E}_{\mathrm{y}^{v}}[(\sigma^{v})^2]}}$$

Working on the inner expectation I get [64] [65]

$$\mathbf{E}_{\mathrm{y}^v}[(\sigma^v)^2] =$$

$$= \frac{U-1}{3N^3} \sum_{j=-N+1}^{N-1} \mathbf{E}_{\mathrm{y}^u,\mathrm{y}^v}^{u\neq v} \left[ 2\Gamma_{N,j}^2(\mathrm{y}^u,\mathrm{y}^v) + \Gamma_{N,j}(\mathrm{y}^u,\mathrm{y}^v)\Gamma_{N,j+1}(\mathrm{y}^u,\mathrm{y}^v) \right]$$

so that, by defining R,

$$\mathrm{R} = \frac{1}{3N^3} \sum_{j=-N+1}^{N-1} \mathbf{E}_{\mathrm{y}^u,\mathrm{y}^v}^{u\neq v} \left[ 2\Gamma_{N,j}^2(\mathrm{y}^u,\mathrm{y}^v) + \Gamma_{N,j}(\mathrm{y}^u,\mathrm{y}^v)\Gamma_{N,j+1}(\mathrm{y}^u,\mathrm{y}^v) \right]$$

we come to the final expression

$$\mathrm{B}EP_{\mathrm{MAI}} \approx \frac{1}{2}\mathrm{erfc}\sqrt{\frac{V^2}{(U-1)\mathrm{R}}} \tag{6.5}$$

which allows us to interpret R as an *e*xpected interference-to-signal ratio per interfering user, i.e., an expected degradation in system performance when a new user is added.

The expression of R can be further simplified taking into account that I are using antipodal symbol for the spreading sequences. More specifically, assuming second-order stationary sequences, i.e. such that $\mathbf{E}_{\mathrm{y}^v}[y_m^v y_n^v] = \mathbf{E}_{\mathrm{y}^v}[y_0^v y_{|m-n|}^v]$, I may use the auto-correlation function $A_k = \mathbf{E}_{\mathrm{y}^v}[y_0^v y_k^v]$ to write, with few algebraic manipulation [65], an alternative expression of R depending only on $A_k$ and $N$, namely

$$\mathrm{R} = \frac{2}{3N} + \frac{4}{3N^3} \sum_{k=1}^{N-1} \left[ (N-k)^2 A_k^2 + \frac{(N-k+1)(N-k)}{2} A_k A_{k-1} \right] \tag{6.6}$$

In [67] R is minimized and minimally interfering sequences are approximated by those with $A_k \approx r^k$ where $r = -2+\sqrt{3}$, which can easily be generated by suitable piecewise affine Markov maps [33]. In other terms, chaos-based

spreading sequences can be generated, whose adoption allows to optimize performance in asynchronous (UWB) DS-CDMA systems when MAI is the main cause of non-ideality and I will use this case as a benchmark in our present study.

### 6.2.2 Narrowband Interference and Thermal Noise

For the present study, symbol recovery is hindered not only by MAI, but also by the presence of thermal noise and, most important, by an NBI, which models either an intentional jamming or a conventional non spread-spectrum transmissions, or both.

For this scenario, I consider the impact of NBI and thermal noise on $\upsilon$-th user, where the former is expressed as a single sinusoidal jamming signal $\sqrt{2I}\cos(2\pi f_0 t + \phi_0)$, of (normalized baseband equivalent) frequency $f_0$, initial phase $\phi_0$ and transmitted power $I$. Hence, relying on [52] one can compute performance in terms of bit error probability as

$$\mathrm{BEP_I} = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} J_0\left(\omega\sqrt{\frac{I}{C}\frac{|H^{\upsilon}(f_0)|^2}{2T}}\right)\frac{\sin\omega}{\omega}e^{\left(-\frac{\omega^2}{4}\frac{N_0}{E_b}\right)}d\omega \tag{6.7}$$

where $J_0(\omega)$ is the 0-th order Bessel function of the first kind, $E_b$ is the transmitted energy per information symbol, $C = E_b/T$ denotes the corresponding useful received power, and thermal noise is modeled as a two-sided power spectral density equal to $N_0/2$. Furthermore, $H^{\upsilon}(f)$ is the transfer function of the MF for $\upsilon$-th link and represents the Fourier transform of $(b(t;-1)-b(t;+1))$, where $b(t;\pm 1)$ is equal to a unit-energy waveform used to transmit the

information symbol $-1$ or $+1$. To find its expression, let us express the unit-energy waveform of the $v$-th user as

$$b^v(t) = \sqrt{\frac{2}{NE_g}} \sum_{j=0}^{N-1} y_j^v g_{\frac{T}{N}}\left(t - j\frac{T}{N}\right)$$

so that, exploiting the fact that the a chip waveform energy is $E_g = T/N$, the desired transfer function can be written as

$$|H^v(f_0)| = N\sqrt{2T}\left|sinc\left(f_0\frac{T}{N}\right)\right|\left|\sum_{s=0}^{N-1} y_s^v e^{-2\pi i f_0 k\frac{T}{N}}\right|$$

### 6.2.3 All causes of error

The aim of this section is to obtain a final expression of the system performance in terms of bit error probability, similar to (6.5) and (6.7), and which includes all considered causes of error at the same time. To do this, I may first note that, within the limit of validity of SGA, MAI is a Gaussian random variable whose effect on system performance is similar to thermal noise. Consequently, I may equivalently model the effect of both MAI and thermal noise exploiting (6.7) if I suitably increase the noise power spectral density from $N_0$ to $\tilde{N}_0$. The same effect of MAI at the output of the MF is obtained if, taking into account that $\int_{-1/T}^{1/T} |H^v(f)|^2 df = C$ for each SS, we impose

$$\frac{\tilde{N}_0}{2} = \frac{N_0}{2} + \frac{(U-1)\mathrm{R}}{C}$$

With this, I finally arrive to express the system performance in presence of NBI, MAI and thermal noise as

$$\begin{aligned} \mathrm{B}EP^v = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty J_0\left(\omega\sqrt{\frac{I}{\mathrm{C}}\frac{|H^v(f_0)|^2}{2T}}\right) sinc(\omega)\cdot \\ \cdot \exp\left(-\frac{\omega^2}{4}\cdot\frac{N_0 + 2(U-1)\mathrm{R}/C}{E_b}\right)d\omega \end{aligned} \tag{6.8}$$

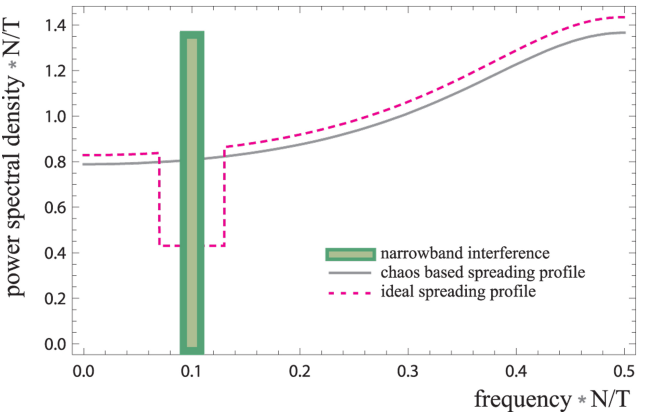



**Figure 6.2:** Normalized PSD of the chaos-based spreading sequence minimizing (6.5) (solid line), jamming signal with $f_0 = 0.1N/T$ (thick solid line) and ideal PSD of spreading sequences capable of minimizing the effect of the NBI at the $\upsilon$-th useful user receiver (dashed line).

As a final remark, I need to note that (6.8) depends on the particular choice of the spreading sequence of the $\upsilon$-th useful user. Regrettably the analytic evaluation of $\mu(\mathrm{BEP}^{\upsilon}) = \mathbf{E}_{\mathbf{y}^{\upsilon}}[\mathrm{B}EP^{\upsilon}]$ is prohibitive from an analytic point of view and I will therefore rely, in Section 7.3, on its numerical evaluation based on a Montecarlo approach.

CHAPTER

# 7 Narrowband Interference Reduction

The preaviously chapter ends with an expression of bit error probability related to a generic user that is able to take into account all considered causes of error, thermal noise, NBI and MAI. This is very useful to test possible strategies that try to produce a performance increasing in terms of BEP.

In this chapter I will describe you my proposal in terms of addressing a trade-off between the reduction of the effects of both NBI and MAI by the shaping of SSs.

## 7.1 The idea

The main idea to reduce the effect of NBI in a DS-CDMA UWB communication relies on the dependence of $BEP^{v}$ in (6.8) on the SS, whose statistical features need thus to be appropriately chosen to achieve such a goal.

The starting point of our procedure, that is described in [67] [54], is a chaos-based technique able to generate

antipodal SSs with autocorrelation profile $A_k \approx (-2+\sqrt{3})^k$, whose adoption leads to a minimum level of MAI and therefore of R. Figure 6.2 shows the normalized Power Spectral Density (PDS) of such sequences (solid line), which therefore gives an indication of the shape of the MF transfer function of the receiver. If the vertical thick line represents the NBI, it is therefore evident that a consistent part of its power is transferred at the output of the MF, i.e., at the input of the decision block, thus increasing the error probability. To cope with this, the most intuitive solution is to use SSs whose PSD features a stop-band in a (large-enough) neighborhood of the NBI frequency, as it is also shown in Figure 6.2 (dashed line).

Note that I have also assumed that the NBI frequency is fixed and known. This is not a critical hypothesis since the aim of this work is to minimize the possible interference of existing narrowband communication systems on a UWB system working in the same band.

Two issues are worth mentioning. On the one hand, independently of the method used to generate spreading sequences, those satisfying the above PSD condition will certainly not be the same that minimize R. This will result in an increased MAI with respect to optimal chaos-based spreading [67] so that the possibility of achieving an improvement in performance will critically depend on the effectiveness in suppressing NBI and on its intensity. On the other hand, as shown in [33], the (binary) quantized output of a one-dimensional chaotic map can exhibit only a low-pass, high-pass or flat PSD, so that SSs with a PSD similar to the one represented by the dashed-line of Figure 6.2 can be only obtained through a much more complex generator. The structure of one of such

generator, the same mentioned in Section 4.1, will be well describe in the next Subsection.

## 7.2 Sequences Generator

To obtain spreading sequences with assigned spectral profile I will rely on a Linear Probability Feedback Process (LPF) recently introduced in [33–35], whose scheme is shown in Figure 7.1 (a). It is based on a causal time-invariant linear filter with impulse response $h_k$ such that $h_k = 0$ for $k \leq 0$ and transfer function

$$H_m(z) = \sum_{k=1}^{m} h_k z^{-k}$$

Though, in principle, there is no need for $m$ to be a finite number, I will make this assumption here. The output of the filter $-H_m(z)$ produces the process $x_t = -\sum_{k=1}^{m} h_k y_{t-k}$. I will assume that $-1 \leq x_t \leq 1$, since, being the $y_t$ antipodal symbols, this is equivalent to impose the constraint

$$\sum_{k=1}^{m} |h_k| \leq 1 \tag{7.1}$$

The process $x_t$ is then fed into a comparator and matched with the process $a_t$ that is made of independent random thresholds uniformly distributed in $[-1, 1]$. The comparator yields the antipodal values $y_t \in \{-1, 1\}$ that are fed back into the filter to allow a continuous generation of symbols. Assuming the uniform cumulative distribution function

$$F(a) = \begin{cases} 0 & \text{if } a < -1 \\ \frac{1+a}{2} & \text{if } -1 \leq a \leq 1 \\ 1 & \text{if } a > 1 \end{cases}$$

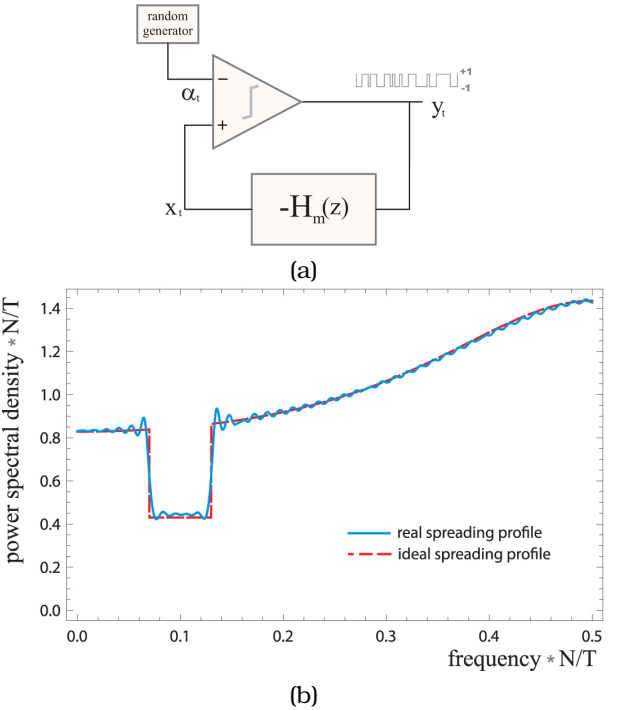



**Figure 7.1:** (a) Structure of a memory-$m$ antipodal linear probability feedback process generator, including a finite memory filter $-H_m(z)$, a random generator and a comparator. (b) Normalized ideal PSD compared with the actual one obtained using the memory-$m$ antipodal linear probability feedback process generator.

for each random variables $a_t$ in the scheme, I can write the probability of the current generated symbol $y_t$ given the previous sequence with memory $m$ as

$$\begin{aligned} &Pr\{y_t = +1 | y_{t-1}, y_{t-2}, \ldots, y_{t-m}\} = \\ &= F\left(-\sum_{k=1}^{m} h_k y_{t-k}\right) = \frac{1}{2}\left[1 - \sum_{k=1}^{m} h_k y_{t-k}\right] \end{aligned} \tag{7.2}$$

As discussed in details in [35], starting from (7.2) a thorough analysis leads to an expression of the normalized power spectral density of the generated antipodal symbols $y_t$, namely

$$\Lambda_y(f) = \frac{\left|1 + H_m(e^{2\pi \mathbf{i} f})\right|^{-2}}{\int_{-1/2}^{1/2} \left|1 + H_m(e^{2\pi \mathbf{i} f})\right|^{-2} df} \tag{7.3}$$

In principle one could use (7.3) to derive $h_k$ once $\Lambda_y(f)$ has been set to a desired PSD profile similar to the one shown in Figure 6.2. Regrettably, inverting (7.3) is a prohibitive task. To cope with this I note that if the process $y_t$ is fed into a filter with transfer function $1 + H_m(f)$, the output is a white process with power spectral density equal to

$$P = \frac{1}{\int_{-1/2}^{1/2} \left|1 + H_m(e^{2\pi \mathbf{i} f})\right|^{-2} df}$$

so the whitening filter of $y_t$ is $1 + H_m(z)$. From this I get that $-H_m(z)$ must be the optimum linear predictor of $y_t$ that can be derived minimizing

$$\epsilon^2 = \mathbf{E}\left[\left(y_t + \sum_{k=1}^{m} h_k y_{t-k}\right)^2\right] =$$
$$= C_{0,0} + 2\sum_{k=1}^{m} C_{k-1,0} h_k + \sum_{k=1}^{m}\sum_{j=1}^{m} C_{k-1,j-1} h_k h_j$$

where $C_{k,j}$ is the correlation matrix of the process $y_t$, defined as

$$C_{k,j} = \mathbf{E}[y_{t-k} y_{t-j}] = \int_{-1/2}^{1/2} \Lambda_y(f) e^{2\pi \mathbf{i}(k-j)f} \mathrm{d}f =$$
$$= \int_{-1/2}^{1/2} \Lambda_y(f) \cos(2\pi(k-j)f) \mathrm{d}f$$

To solve this problem, I can use the classic approach based on Yule-Walker equations. Regrettably the solution, in general, do not satisfy (7.1). A numerical method such as a modified gradient descent was presented in [35] and more sophisticated heuristic techniques are used in [36].

As an example, exploiting the first technique with $m = 80$ I have been able to obtain spreading sequences whose normalized PSD is shown in Figure 7.1 (b) (continuous line). As it can be seen, the shape is almost superimposing with the target one (dashed line).

Note also that the presented method generates sequences with stationary statistical features, but this is not a limit in our case since I are supposing that the NBI frequency is fixed and know.

## 7.3 Numerical Results

As discussed at the end of Section 6.2, Equation (6.8) shows that the performance of the $v$-th user in terms of bit error probability $BEP^v$ is actually a random variable that depends in a non-linear way on the spreading sequence $\mathbf{y}^v$. Consequently, to evaluate the effectiveness of the NBI reduction methodology proposed in this paper, I must rely on a statistical characterization of $BEP^v$ by means of Montecarlo simulations.

To collect statistics for the generic $v$-th user and for a given system configuration, I first generate $2 \times 10^5$ spreading sequences using the LPF discussed in Section 7.2 for a fixed spectral shape. Then I compute the corresponding instance of $BEP^v$ exploiting (6.8), where R is

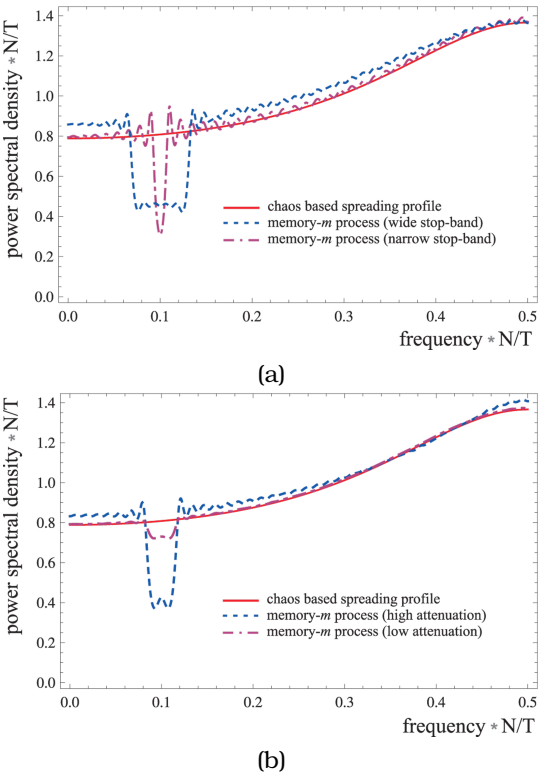



**Figure 7.2:** Plot (a) and (b) show the PSD associated to chaos-based spreading sequences minimizing MAI (continous-line) and those generated by an LPF with $m = 80$ targeting stop-bands with different widths and depths to reduce NBI (dotted and dash-dotted lines).

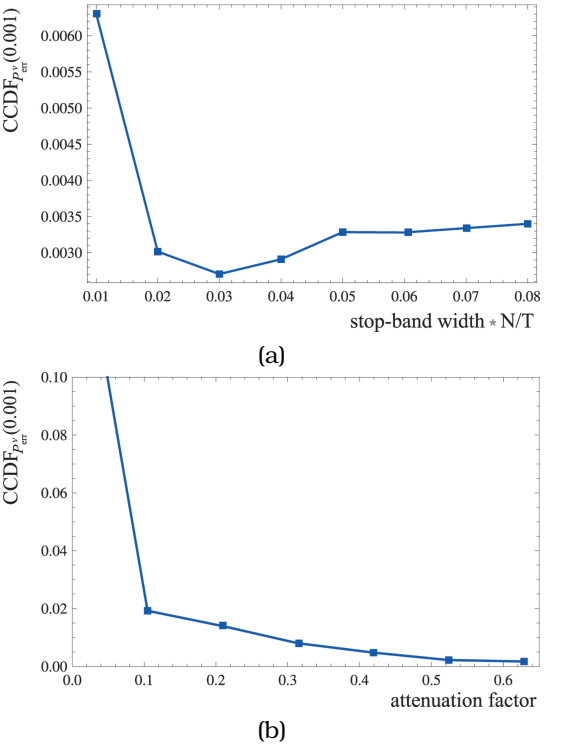



**Figure 7.3:** Plot (c) and (d) show $CCDF_{P_{\text{err}}^{v}}(10^{-3})$ as a function of the stop-band width and the attenuation factor $\rho$ (related to the stop-band depth), considering SIR= $-12$ dB, SNR= 20 dB, $f_0 = 0.1N/T$, $U = 12$.

evaluated through (6.6) using the autocorrelation function corresponding to the chosen spectral profile. As a realistic setting, following [52], I refer to an asynchronous DS-CDMA system with $T = 100$ ns, $N = 128$ (so that the spreading bandwidth is $N/T = 1.28$ GHz), and where the thermal noise power is relatively low (SNR= $E_b/N_0 = 20$ dB).

As a first step, I consider the relationship between

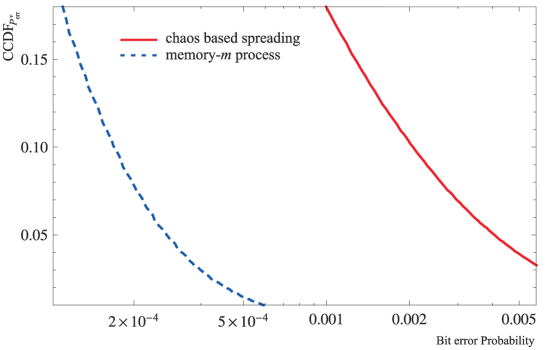



**Figure 7.4:** The C$CDF_{P_{\text{err}}^{U}}$ as a function of B$EP$, with: stop-band width = $0.03N/T$, attenuation factor $\rho = 0.62$, SIR= $-12$ dB, SNR= 20 dB, $f_0 = 0.1N/T$, $U = 12$.

the bit error probability and the width and depth of the introduced stop-band in the spectral profile. Figure 7.2 (a) shows the spectral profile of the generated sequences with maximum (dashed line) and minimum (dash-dotted line) width, while Figure 7.2 (b) shows similar figures used to investigate the role of stop band depth. In both figures the solid line is the considered benchmark, namely the PSD profile of chaos-based spreading sequences minimizing MAI only.

We refer to a system where $U = 12$ users are present, with a signal-to-interference-ratio SIR= $C/I = -12$ dB and where the NBI is centered in a neighborhood of $f_0 = 0.1N/T$. In this setting, the NBI plays an important role and its reduction is therefore expected to show appreciable effects, making it easier to find an optimum value of the stop-band width and of the stop-band depth.

Focusing on the stop-band depth, let us define the

attenuation factor $\rho$ for a fixed NBI frequency, namely

$$\rho = \frac{\mathrm{PSD}^{chaos-based}(f_0) - \mathrm{PSD}^{memory-m}(f_0)}{\mathrm{PSD}^{chaos-based}(f_0)}$$

The figure of merit I chose is the Complementary Cumulative Distribution Function (CCDF) of $\mathrm{BEP}^{v}$, defined as $\mathrm{CCDF}_{P_{\mathrm{err}}^{v}}(\delta) = \Pr\{\mathrm{BEP}^{v} > \delta\}$, which I will computed for the typical case of $\delta = 10^{-3}$. Let us refer to the results shown in Figure 7.3 (a) and (b). Plot (a) represents $\mathrm{CCDF}_{P_{\mathrm{err}}^{v}}(10^{-3})$ as a function of the normalized stop-band width, clearly showing a minimum for the optimal width value equal to $0.03N/T$ (corresponding to 39.322 MHz). It is worth noting that I investigate only a limited range for the stop-band width, as shown in Figure 7.2 (a). In fact, on the one end, for larger stop-bands the resulting SS profiles would be too much close to uniform and therefore too different with respect to the one minimizing the effect of MAI. On the other end, a narrower stop-bands would bring negligible advantage in NBI reduction.

Figure 7.3 (b) shows that $\mathrm{CCDF}_{P_{\mathrm{err}}^{v}}(10^{-3})$ is always decreasing with the attenuation factor $\rho$. Hence the optimal value would apparently correspond to $\rho = \rho_{max} = 1$, for which $\mathrm{PSD}^{memory-m}(f_0) = 0$. Yet, such a lower value of $\mathrm{PSD}^{memory-m}(f_0)$ cannot be achieved with LPFs, since they are characterized by an almost-everywhere non-null PSD [35]. Consequently, for practical purposes, I will use an upper bound for $\rho$ equal to the maximum value which guarantees correct generation of the SS via an LPF ($\rho = 0.62$ in this setting).

Figure 7.4 represents $\mathrm{CCDF}_{P_{\mathrm{err}}^{v}}(\mathrm{BEP})$ computed for the optimal values of the stop-band width and of the attenuator factor $\rho$ mentioned above. It can be clearly seen that memory-$m$ sequences always offer an advantage

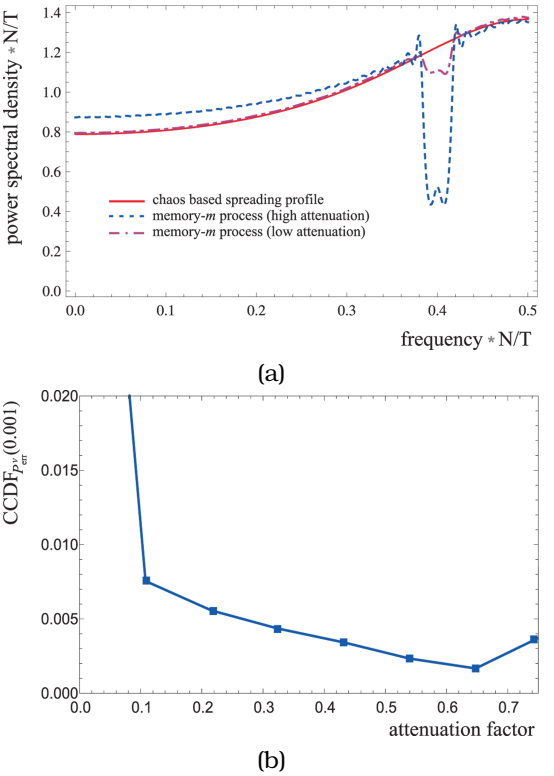



**Figure 7.5:** (a) PSD associated to chaos-based spreading sequences minimizing MAI (continous-line), and generated by LPF with $m = 80$ targeting stop-band with different depth to reduce NBI (dashed and dot-dashed lines. (b) $CCDF_{P_{\text{err}}^{v}}(10^{-3})$ as a function of the attenuation factor, considering: SIR= $-3$ dB, SNR= 20 dB, $f_0 = 0.4N/T$, $U = 20$.

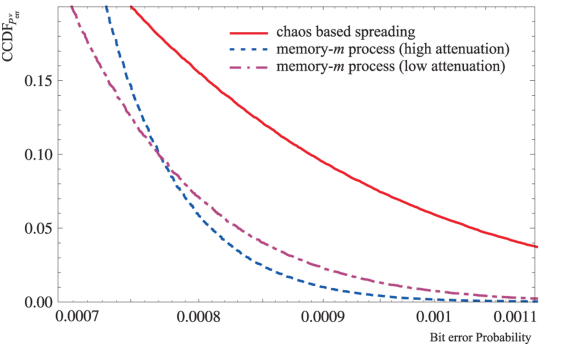



**Figure 7.6:** The C$CDF_{P_{\text{err}}^v}$ as a function of B$EP$, with: stop-band width = $0.03N/T$, SIR= $-3$ dB, SNR= 20 dB, $f_0 = 0.1N/T$, $U = 20$.

with respect to chaos-based ones, with an improvement in reducing C$CDF_{P_{\text{err}}^v}$ of at least 30%.

As a further step, I analyze the influence on system performance of the number of users $U$ (on which MAI directly depends, see (6.5)), of SIR (whose increment reduces NBI) and of $f_0$. The latter parameter plays a particularly important role. In fact, since the PSD of optimal chaos-based spreading is high-pass (see the continuous-line in Figure 7.2 (a) and (b)), one may intuitively accept that when $f_0 < N/2T$, the presence of a stop-band introduces a negligible perturbation of the profile with respect to the one minimizing MAI. Consequently, as one may expect, results of all performed numerical simulations confirm that the optimal width for the stop band is always close to the one corresponding to the minimum value in Figure 7.3 (a). On the contrary, the influence of $\rho$ is different, as reported in Figure 7.5 (a)-(b) where, with respect to the previous case, $U = 20$, SIR= $-3$ dB and $f_0 = 0.4N/T$. As it can be noticed, the optimal value of the attenuation

factor is $\rho = 0.65$, i.e., it is no more coincident with the maximum value compatible with the LPF generation method.

The associated results in terms of $\mathrm{CCDF}_{P_{\mathrm{err}}^{\upsilon}}$ are shown in Figure 7.6, where the dashed and the dot-dashed lines correspond respectively to $\rho = 0.65$ (optimum value for $\mathrm{BEP}^{\upsilon} = 10^{-3}$) and $\rho = 0.12$. It is interesting noticing that the proposed method always improves performance with respect to chaos-based spreading (continuous-line), but the actual improvement depends on the choice of $\rho$.

Given the above results, one may therefore conclude that, in any given scenario scenario, the presented method for SS generation offers an advantage with respect to the classical chaos-based solution, provided that a preliminary study on the influence of the stop-band width and depth in the SS profile is performed numerically to evaluate their optimal value.

Finally, Figures 7.7-7.8 represents the expected performance of $\mathrm{BEP}^{\upsilon}$ as a function of SIR, showing its mean value and its variance for SNR= 20 dB and stop-band width equal to $0.03N/T$, and where $\rho = 0.62$, $f_0 = 0.1N/T$ and $U = 12$ in plots 7.7 (a)-(b), while $\rho = 0.65$, $f_0 = 0.4N/T$ and $U = 20$ in plots 7.8 (a)-(b). Overall, memory-$m$ sequences offer an uniform variance reduction and a general performance improvement in most operating conditions.

## 7.4 Conclusion

In this part of the thesis, I presented a simple method to reduce the impact of a narrowband interference in

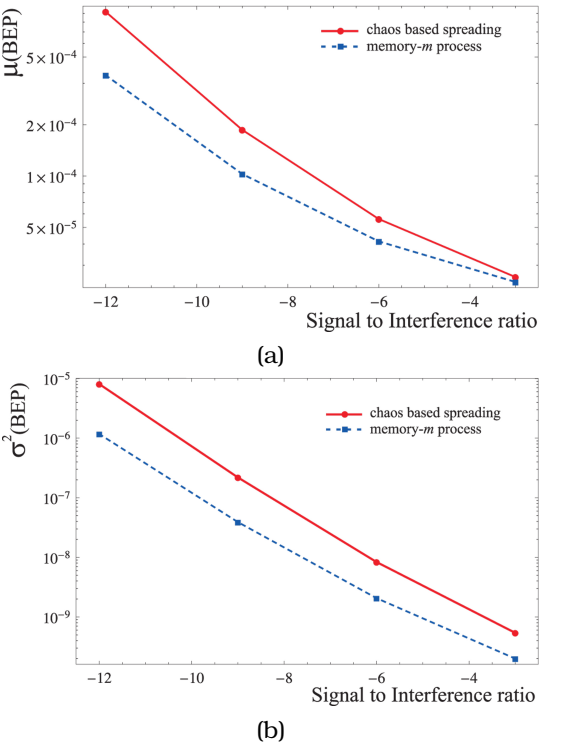



**Figure 7.7:** Plot (a) and (b) show mean and variance of B*EP* with: stop-band width = $0.03N/T$, $\rho = 0.62$, SNR= 20 dB, $f_0 = 0.1N/T$, $U = 12$.

UWB systems based on asynchronous DS-CDMA. The starting point is given by existing techniques reducing multi-access disturbances by appropriate chaos-based spreading codes. Then, the method is focused on shaping power spectral density of spreading sequences, made by symbols generated by LPFs.

Power spectral shaping reduces the UWB information signal in the NBI frequency band, yielding a heavy reduction of the interference due to overlapping narrowband

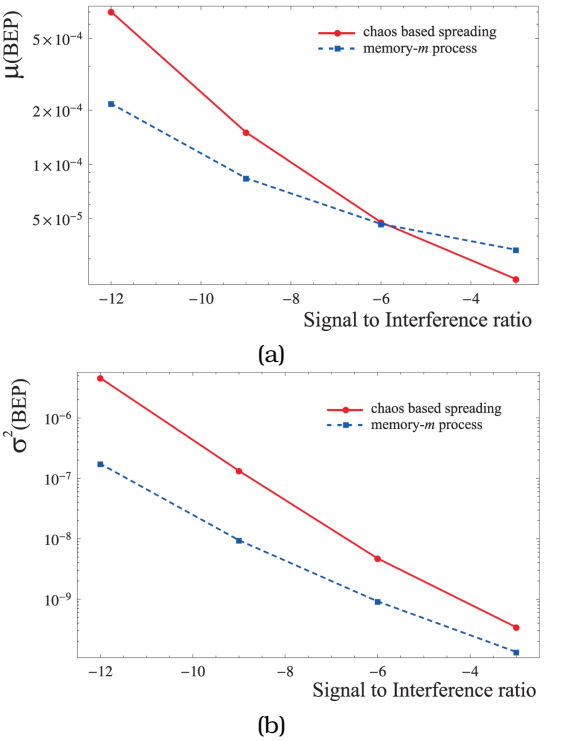



**Figure 7.8:** Plot (a) and (b) show mean and variance of B*EP* with: stop-band width = $0.03N/T$, $\rho = 0.65$, SNR= 20 dB, $f_0 = 0.4N/T$, $U = 20$.

transmission. This however causes an increment in multi-access disturbance with respect to chaos-based spreading that minimizes MAI, thus implying a trade-off between rejection of narrowband interference and tolerance of multi-access interference.

Such a trade-off can be effectively address for different system configurations depending on the number of simultaneous users on the channel, the relative position of the UWB and narrowband spectra, and on the ratio between

the power of the localized interferer and the power of UWB signal.

Numerical results reveal non-negligible improvements since, whenever SIR≥ 5dB, $U \geq 20$ and $N = 128$, I found that memory-$m$ sequences guarantee a lower average BEP$^v$, as well as an higher probability to reach any given quality link.

Note that in this paper I have assumed that the NBI frequency is fixed and know, as well as its power and the number of users in the system. As long as these parameters are unchanged, I have shown how to generate spreading sequences in order to maximize system performance; this solution however does not ensure anymore the optimum working point when one of the above parameter changes.

In order to ensure the performance optimization also in the case of a change in one (or more) of the above parameters, spreading sequences statistical features should be recomputed as soon as environmental conditions change. This is certainly not a problem for new users added to the system, that may immediately work with optimized sequences. However, old users (i.e. users that use spreading sequences optimized for a different environmental setting) may work with performance that may be lower with respect to the reference case, i.e. when sequences are generated with an autocorrelation profile $A_k \approx (-2 + \sqrt{3})^k$ without taking into account any NBI. To cope with this, a sequence renegotiation policy has to be implemented.

# Conclusions

In this thesis I reported and discussed the advantage introduced by designing the statistical properties of the antipodal sequences involved in both presented scenarios. Before a brief summary of the obtained results, I would highlight that the use of antipodal sequences with prescribed second order characterization does not correspond to a complete upsetting of the presented systems. In all cases where the involved sequence are not generated by an LPF process or a 2vRNG, it can be done be replacing the initial sequence generator one of them.

## On Compressive Sensing

After an ample discussion of both CS theory and its related architectures (Chapters 1 and 2), I proposed a different approach to the sampling sequence generation when the RMPI-based AIC is taken into account. This new

methodology is based on an additional assumption on the acquired class of signals, that is energy localization.

It is clear that working with localized signals is not a strong restriction on the signal classes being acquired, i.e., with the exception of noise, many classes of signals do not present a uniformly spread in information/energy content over the entire signal domain.

Working with this assumption, a new guideline was introduced for the sampling sequences characterization,i.e., the rakeness. Using rakeness it is possible to determinate the second order statistic of the spreading sequences that increase the amount of information/energy collected in the acquisition stage with the aim of reducing the minimum amount of projections needed to achieve signal reconstruction.

After that, I proposed innovative architectures for AIC, namely the S-RMPI and the MS-RMPI and their derivations based on rakeness. These new architectures guarantee comparable performances with respect to RMPI and at the same time it was obtained a strong reduction in terms of power consumption by slowing down the sampling process.

Numerical results were shown at the end of Chapter 4 and 5 ,while a full discussion of rakeness was reported in Chapter 3.

## On UWB system based on DS-CDMA

In the second part of this thesis a relatively simple technique for shaping the power spectrum of antipodal sequences was shown effective when used to design the

frequency profile of the signals transmitted by an UWB system based on DS-CDMA.

This profile can be adopted to at least partially move transmitted power out of frequency ranges that are subject to heavy interference due to overlapping narrowband transmissions.

In multi-user environments, a trade-off must be administrated between rejection of such narrowband interference and tolerance to multi-access disturbances which would be minimized by suitable chaos-based spreading codes.

The effectiveness of the proposed technique depends on the relative position of the UWB and narrowband spectra and on the power of the localized interferer.

In general, the analysis reveals non-negligible improvements. All results are reported in Chapter 7 and a related description of both the considered environment and all studied causes of error at the receiver are illustrated in Chapter 6.

# Bibliography

[1] M. B. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication–part1: System analysis," *IEEE Trans. Commun.*, vol. 25, pp. 795 –799, 1970.

[2] H. Nyquist, "Certain topics in telegraph transmission theory," *American Institute of Electrical Engineers, Transactions of the*, vol. 47, no. 2, pp. 617 –644, april 1928.

[3] C. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10 – 21, jan. 1949.

[4] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and Implementation of an Analog-to-Information Converter using Random Demodulation," in *Proceedings of 2007 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2007, pp. 1959–1962.

[5] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[6] E. J. Candes and M. B. Wakin, "An Introduction to Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[7] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 6, pp. 797–829, 2006. [Online]. Available: http://dx.doi.org/10.1002/cpa.20132

[8] A. Maleki and D. Donoho, "Optimally tuned iterative reconstruction algorithms for compressed sensing," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 330–341, 2010.

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

[10] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[11] E. J. Candes, J. K. Romberg, K. Justin, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

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

[13] C. Luo and J. H. McClellan, "Compressive sampling with a successive approximation adc architecture," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, may 2011, pp. 3920 –3923.

[14] P. Yenduri, A. Gilbert, M. Flynn, and S. Naraghi, "Rand ppm: A lowpower compressive sampling analog to digital converter," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, may 2011, pp. 5980 –5983.

[15] M. Mangia, R. Rovatti, and G. Setti, "Analog-to-information conversion of sparse and non-white signals: Statistical design of sensing waveforms," in *Proceedings of 2011 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2011, pp. 2129–2132.

[16] ——, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 5, pp. 1001 –1014, May 2012.

[17] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[18] S. Mendelson, A. Pajor, and N. Tomczac-Jagermann, "Uniform uncertainty principle for Bernoulli and sugaussian ensambles," *Constructive approximation*, vol. 28, Dec. 2008.

[19] S. Haykin, *Adaptive Filter Theory*. Prentice Hill, 2002.

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

[21] J. N. Laska, P. T. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in Action: Quantization, Saturation, and Compressive Sensing," *Applied and Computational Harmonic Analysis*, vol. 31, pp. 429–443, Nov. 2011.

[22] M. Mangia, F. Pareschi, R. Rovatti, G. Setti, and G. Frattini, "Coping with saturating projection stages in rmpi-based compressive sensing," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, may 2012, pp. 2805 –2808.

[23] E. J. Candes and J. Rosemberg, "Sparsity and incoherence in compressive sensing," *Inverse problems*, vol. 23, pp. 969–985, 2007.

[24] E. Allstot, A. Chen, A. Dixon, D. Gangopadhyay, and D. Allstot, "Compressive sampling of ecg biosignals: Quantization noise and sparsity considerations," in *Biomedical Circuits and Systems Conference (BioCAS), 2010 IEEE*, nov. 2010, pp. 41 –44.

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

[26] M. Mangia, J. Haboba, R. Rovatti, and G. Setti, "Rakeness-based approach to compressed sensing of ecgs," in *Biomedical Circuits and Systems Conference (BioCAS), 2011 IEEE*, Nov. 2011, pp. 424 –427.

[27] J. Haboba, M. Mangia, R. Rovatti, and G. Setti, "An architecture for 1-bit localized compressive sensing with application to eeg," in *Proceedings of 2011 IEEE International Symposium on Biomedical Circuits and Systems (BIOCAS)*, Nov. 2011, accepted for publication.

[28] R. Baker, H. Li, and R. De Vore, *CMOS Circuit Design, Layout and Simulation*. New York: IEEE Press Series on Microelectronic System, 1998.

[29] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed norm," *Signal Processing, IEEE Transactions on*, vol. 57, no. 1, pp. 289 –301, jan. 2009.

[30] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale l1-regularized least squares," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 606 –617, dec. 2007.

[31] M. Marcus and H. Minc, *A survay of Matrix theory and matrix inequality*. New York: Dover Publishing, 1992.

[32] J. Ranieri, R. Rovatti, and G. Setti, "Compressive sensing of localized signals: Application to analog-to-information conversion," in *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, June 2010, pp. 3513–3516.

[33] G. Setti, G. Mazzini, R. Rovatti, and S. Callegari, "Statistical modeling of discrete-time chaotic processes-basic finite-dimensional tools and applications," *Proceedings of the IEEE*, vol. 90, no. 5, pp. 662 –690, may 2002.

[34] R. Rovatti, G. Mazzini, G. Setti, and A. Giovanardi, "Statistical modeling and design of discrete-time chaotic processes: advanced finite-dimensional tools and applications," *Proceedings of the IEEE*, vol. 90, no. 5, pp. 820 –841, may 2002.

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

[36] R. Rovatti, G. Mazzini, G. Setti, and S. Vitali, "Linear probability feedback processes," in *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on*, may 2008, pp. 548 –551.

[37] A. Kolmogorov and S. Formin, *Introductory Real Analysis*. New York: Dover Publishing, 1975.

[38] D. Slepian and H. Pollak, "Prolate spheroidal wave functions, fourier analysis, and uncertainty-i," *Bell Syst. Tech J.*, vol. 40, pp. 43 –64, 1961.

[39] H. Landau and H. Pollak, "Prolate spheroidal wave functions, fourier analysis, and uncertainty-ii and iii," *Bell Syst. Tech J.*, vol. 40, pp. 65 –84, 1961.

[40] R. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems*. SIAM, 2009.

[41] R. Birkov, "Tres observationes sobre algebra linear," *Revista Facultad de Ciencas Exactas, Puras y Aplicadas Universidad Nacional d Tacuman, Aerie A (Matematicas y Fisica Teorica*, vol. 5, pp. 147 –151, 1946.

[42] R. Agarwal and S. Sonkusale, "Direct analog-to-qrs detection front-end architecture for wearable ecg applications," *Proc. on EMBC 2010*, pp. 6527 –6530, 2010.

[43] L. Polania, R. Carrillo, M. Blanco-Velasco, and K. E. Barner, "Compressed sensing based method for ecg compression," *IEEE International Conference on Acoustic Speech and Signal Processing*, pp. 761 –764, 2011.

[44] T. Zhang, W. andDu and H. Tang, "Energy-efficient ecg acquisition in body sensor network based on compressive sensing," *International J. of Digital Content Tecnology and its Application*, vol. 5, no. 4, pp. 18 –25, 2011.

[45] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries,"

*Information Theory, IEEE Transactions on*, vol. 54, no. 5, pp. 2210 –2219, may 2008.

[46] E. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, vol. 31, pp. 59 –73, 2011.

[47] Y. Wu, H. Zhang, H. Wang, and Y. Lu, "The sparse decomposition and compression of ecg and eeg based on matching pursuits," in *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, vol. 3, oct. 2010, pp. 1094 –1097.

[48] A.L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.

[49] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Design and exploration of low-power analog to information conversion based on compressed sensing," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, oct. 2012.

[50] [Online]. Available: http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/

[51] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*. New York: McGraw-Hill, 1994.

[52] A. Giorgetti, M. Chiani, and M. Z. Win, "The effect of narrowband interference on wideband wireless communication systems," *Communications, IEEE Transactions on*, vol. 53, no. 10, p. 1773, oct. 2005.

[53] T. Kohda and A. Tsuneda, "Even- and odd-correlation functions of chaotic chebyshev bit sequences for cdma," in *Spread Spectrum Techniques and Applications, 1994. IEEE ISSSTA '94., IEEE Third International Symposium on*, jul 1994, pp. 391 –395 vol.2.

[54] G. Cimatti, R. Rovatti, and G. Setti, "Chaos-based spreading in ds-uwb sensor networks increases available bit rate," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 54, no. 6, pp. 1327 –1339, june 2007.

[55] R. Rovatti, G. Mazzini, and G. Setti, "A tensor approach to higher order expectations of quantized chaotic trajectories. i. general theory and specialization to piecewise affine markov systems," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 47, no. 11, pp. 1571 – 1583, nov 2000.

[56] G. Mazzini, R. Rovatti, and G. Setti, "Chaos-based asynchronous ds-cdma systems and enhanced rake receivers: measuring the improvements," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 48, no. 12, pp. 1445 – 1453, dec. 2001.

[57] G. Setti, R. Rovatti, and G. Mazzini, "Performance of chaos-based asynchronous ds-cdma with different pulse shapes," *Communications Letters, IEEE*, vol. 8, no. 7, pp. 416 – 418, july 2004.

[58] G. Mazzini, G. Setti, and R. Rovatti, "Chip pulse shaping in asynchronous chaos-based ds-cdma,"

*Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 54, no. 10, pp. 2299 –2314, oct. 2007.

[59] Y. Jitsumatsu and T. Kohda, "Gaussian chip shaping enhances the superiority of markovian codes in ds/cdma systems," in *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on*, may 2006, p. 4 pp.

[60] Y. Jitsumatsu, M. Ogata, and T. Kohda, "A comparison between prolate spheroidal and gaussian fir pulse shaping filters," in *Signals and Electronic Systems, 2008. ICSES '08. International Conference on*, sept. 2008, pp. 343 –346.

[61] R. Rovatti, G. Mazzini, and G. Setti, "On the ultimate limits of chaos-based asynchronous ds-cdma-i: basic definitions and results," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 51, no. 7, pp. 1336 – 1347, july 2004.

[62] ——, "On the ultimate limits of chaos-based asynchronous ds-cdma-ii: analytical results and asymptotics," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 51, no. 7, pp. 1348 – 1364, july 2004.

[63] M. Mangia, R. Rovatti, and G. Setti, "Narrowband interference reduction in uwb systems based on spreading sequence spectrum shaping," in *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, 30 2010-june 2 2010, pp. 1799 –1802.

[64] G. Mazzini, G. Setti, and R. Rovatti, "Chaotic complex spreading sequences for asynchronous ds-cdma-part i: System modeling and results," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 44, pp. 937 –947, 1997.

[65] G. Mazzini, R. Rovatti, and S. G., *Chaos-Based Asynchronous DS-CDMA," Chap 3 in Chaotic Electronics in Telecommunications*. Boca Raton: CRC Press, 2000.

[66] J. Lehnert and M. Pursley, "Error probabilities for binary direct-sequence spread-spectrum communications with random signature sequences," *Communications, IEEE Transactions on*, vol. 35, no. 1, pp. 87 – 98, jan 1987.

[67] G. Mazzini, R. Rovatti, and G. Setti, "Interference minimisation by autocorrelation shaping in asynchronous ds-cdma systems: chaos-based spreading is nearly optimal," *Electronics Letters*, vol. 35, no. 13, pp. 1054 –1055, june 1999.