ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA
ARCES - ADVANCED RESEARCH CENTER ON ELECTRONIC SYSTEMS

Low Correlated Sequences & Architectures and Algorithms for Analog-to-Information Converters: Theory, Design, Implementation and Applications

A THESIS PRESENTED BY

Salvador Javier Haboba

SUPERVISORS

Professor Riccardo Rovatti

Professor Gianluca Setti

Professor Sergio Callegari

COORDINATOR

Professor Claudio Fiegna

DOCTORATE ON INFORMATION TECHNOLOGY
JANUARY 2010 - DECEMBER 2012
XXV CYCLE - ING-INF/01

ABSTRACT

Most electronic systems can be described in a very simplified way as an assemblage of analog and digital components put all together in order to perform a certain function. Nowadays, there is an increasing tendency to reduce the analog components, and to replace them by operations performed in the digital domain. This tendency has led to the emergence of new electronic systems that are more flexible, cheaper and robust. However, no matter the amount of digital process implemented, there will be always an analog part to be sorted out and thus, the step of converting digital signals into analog signals and vice versa cannot be avoided. This conversion can be more or less complex depending on the characteristics of the signals. Thus, even if it is desirable to replace functions carried out by analog components by digital processes, it is equally important to do so in a way that simplifies the conversion from digital to analog signals and vice versa.

In the present thesis, we have study strategies based on increasing the amount of processing in the digital domain in such a way that the implementation of analog hardware stages can be simplified. To this aim, we have proposed the use of very low quantized signals, i.e. 1-bit, for the acquisition and for the generation of particular classes of signals.

More specifically, on one hand, we have proposed a method for the generation of sets of binary sequences to be used in multiple-input multiple-output active sensing applications, such as radar, sonar and medical imaging. The generated sets of sequences have very low auto- and cross-correlation sidelobes, a desired property for this kind of applications, providing performance metrics far better than those from other families of binary sequences, and a comparable performance

to that of multibit approaches. The advantage of using binary sequences is, for instance, the simplification of the implementation of the transmitters always present in these applications.

On the other hand, we have proposed a new architecture for an analog to digital converter. This architecture can be viewed as an extension of the functionalities of a classical Delta-Sigma converter which, by taking 1-bit measurements at a rate much bigger than that of the signal bandwidth, produces a signal estimate with an accuracy that depends on the ratio between the sampling rate and the signal bandwidth. In our case, relying on the structure of the signal of interest, and assuming that its information content is much smaller than its bandwidth, we are able to produce a signal estimate that depends on the ratio between the sampling rate and the information content of the signal.

ACKNOWLEDGMENTS

The possibility to perform this thesis was thanks to the Erasmus Mundus scholarship that financed my stage at the University of Bologna. I would like to thanks authorities from Erasmus for giving me this opportunity.

I would also want to thanks authorities and researchers from the Advance Research Center on Electronic Systems (ARCES) for accepting me as part of this big teem. Thanks to my supervisors Riccardo Rovatti, Gianluca Setti and Sergio Callegari for guiding me during these years, for sharing with me their knowledge, and for introducing me on the amazing field of compressing sensing and on many other topics. I would like to thanks my mates: Carlos, Mauro, Salvatore, Fabio and Valerio, with whom I shared most of the time, sometimes discussing about our researches and sometimes just chatting about life.

Coming from a different country, life in Italy and especially in Bologna was easier and even wonderful thanks to all the friends I made. They supported me and encouraged me during difficult periods, and I share with them special moments that made of Bologna one of the best cities to live. Thanks to Victor, Silvia, Martín, Fercha, Gonzalo, Carlos, Mariano Yanina, Sandra y Laura (las chirusas), Maria Jose y Sofía (las chirusiñas), the Celtic friends (Mark, Aaron, JB, Kyle, Valentina, Cristina Chapinara), Sahar and Nasrin.

Thanks also to my friend Demian for the ideas we shared through the distance, for understanding my absence and for some advices in this work.

I would like to thank my family: my parents and my sisters for their constant

support and encouragement.

Finally, I would like to specially thank Victoria, for her invaluable support during all the time we spent in Bologna, La Plata and everywhere, for her advices, encouragement, for being my partner and for making me enjoy life together.

"Y si trabajás al pedo
y estás haciendo algo nuevo, adelante ..."
Charly Garcia

CONTENTS

LIST OF FIGURES

1. INTRODUCTION

Although most modern electronics systems are composed by a combination of analog and digital components, the rapidly evolving capabilities of digital electronics are shifting every function (before) handled in the analog domain into the digital domain.

The advances in integrated circuits design have enabled the creation of digital processing systems that are more flexible, cheaper and robust than their analog counterparts. This has lead to one of the most significant development during the last decades of electronic systems design: replacing analog components to perform their operation in the digital domain.

However, for these systems to interface with the real world, conversions between analog signals and digital signals are required. Analog-to-Digital (AD) and Digital-to-Analog (DA) converters are the responsible of that conversion.

Most AD and DA converters rely on the Nyquist-Shannon sampling theorem that determines how any signal can be exactly recovered from a set of uniformly spaced samples taken at a rate of at least twice the highest frequency present in the signal of interest.

Nyquist-Shannon sampling theorem imposes a requirement on the time domain to the problem of how to represent an analog signal by a series of samples without any lost of information. However, in order to process and store samples in a digital system, we must be able to represent each sample using a finite number of bits, and hence the measurements will typically be subject to the unavoidable quantization error. By increasing the number of bits of the measurements, the

quantization process can be neglected, or hidden with respect to processes present in the system such as thermal noise. The main drawback of this approach, is that the cost of increasing the number of bits for Nyquist based AD and DA converters require a huge amount of analog hardware. As an example, "flash" AD converters exponentially increase its hardware complexity with the number of resolution bits, and becomes impractical at resolutions over 8 bits due to the large number of comparators required.

There is a different approach for AD and DA conversion that it is not based on the Nyquist-Shannon sampling theorem. Delta-Sigma converters rely on the utilization of a very small amount of bits to quantize signals (1-bit quantization is the most typical value used). Delta-Sigma converters achieve this by trading-off resolution with the sampling frequency. These converters oversample the signal by a large factor with respect to its bandwidth and, by a filtering processing (analog or digital) they are able to obtain a final signal represented with an accuracy much bigger than the one used in the sampling process.

Among other advantages (low power, low cost) with respect to other converter architectures, at the heart of the Delta-Sigma is the simplification on the quantization stage that allows the converter to operate with no linearity degradation. However, Delta-Sigma converters only allows to efficiently operate with signals with a reduced spectra occupancy, due to the high oversampling ratio needed to obtain the desired precision.

The main motivation of this dissertation is to study simplification strategies for the implementation of analog hardware stages present in most mixed systems, by increasing the amount of processing in the digital domain. We accomplished this by proposing the utilization of very low quantized signals, i.e. 1-bit, for the acquisition and for the generation of particular classes of signals. Following this approach we have archived performances similar (or even better) than those obtained through multibit approaches.

We first present the use of Legendre sequences (1-bit sequences) for the gener-

ation of sets of sequences with good auto- and cross-correlation properties. These **Sets of Low Correlated Sequences** can be used in MIMO (Multiple Input Multiple Output) active sensing systems achieving a significant improvement with respect to other sets of binary sequences, and a similar performance to the one achieved by multibit approaches.

Secondly, we present a new architecture for an Analog to Digital converter (or more precisely, an **Analog to Information Converter**) that, based on a Delta-Sigma converter, produces a stream of 1-bit measurements, and achieves a reconstruction performance proportional to the signal information content instead of that of the signal bandwidth.

1.1 Sets of Low Correlated Sequences

The design of sequences sets with low aperiodic auto- and cross-correlations is present in many fields of engineering and plays an important role in many applications such as radar, sonar, communications, medical imaging and other active sensing applications.

The task of designing sets of sequences with prescribed correlation properties is a particular case of the general problem of waveform synthesis that is often a key point in establishing the performance of transmission, synchronization, or active sensing systems [1, 2].

Good auto-correlation properties means that any sequence in the set is nearly uncorrelated with its own shifted version while good cross-correlation means that any member of the sequences set is nearly uncorrelated with any other members at any shift. A commonly used metric of the goodness of the correlation is the *Integrated Sidelobe Level* (ISL), being good set of sequences those having a low ISL value.

Although many state-of-the art algorithms were proposed for the minimization of the ISL [3, 4, 2, 5, 6], their performance is largely impaired when quantization

is taken into consideration. However, implementation constraints strongly favors discrete-valued signals, possibly enforcing quantization to an extremely limited number of levels.

What we propose here is a procedure to construct sets of antipodal sequences with extremely low ISL. The resulting performance largely exceeds that of classical methods for the direct generation of low-ISL sets of sequences.

1.2 Analog to Information Converters

Analog to Digital conversion is one of the most important operations in signal processing. It maps a continuous-time and real-value signal into a discrete sequence of discrete values. Classical sampling methods rely on the hypothesis that the analog signal to be acquired is band-limited, and the Nyquist-Shannon theorem states the minimum distance between samples (or Nyquist rate) needed to uniquely describe the analog signal by its samples.

While the assumption of bandlimited signals is of broad application, many natural signals when represented in a proper basis, correspond to vectors in which many components have a small value, or represent a small fraction of the total energy. This characteristic called "sparsity" is usually exploited to represent the signal with a much smaller amount of data, and closer to the signal information content.

A novel sampling paradigm that goes against the common approach in data acquisition has emerged in the last years and is called Compressive Sensing (CS) [7, 8, 9]. CS theory asserts that one can recover certain signals and images from far fewer samples or measurements than those used by traditional methods. This is possible due to the fact that many natural signals are sparse or compressible, and, by measuring in a particular way, it is possible to acquire the complete information content of those kind signals.

Analog-to-Information converters relies on this idea, to measure the informa-

tion content of the signal instead of measuring the complete redundant data available for a particular measurement domain.

Following this approach, we have proposed a novel architecture for an Analog-to-Information converter that allows a simple hardware implementation for the acquisition of large bandwidth signals that are sparse over a variety of supports.

1.3 Overview and Main Contributions

This thesis is mainly concerned on how signal processing techniques can be applied to real hardware applications and help to reduce the complexity of its implementation.

This work is divided into two parts, the first part tackle the problem of sequences synthesis, and how a proper design of simple antipodal signals can achieve a performance similar to that obtained using multibit sequences for active sensing applications. The main contributions of this part are:

- an analysis of the degradation of state-of-the-art algorithms for sequences synthesis when quantization is imposed;
- a method based on generating functions for the calculation of the cross-correlation components of the ISL of a set of sequences;
- a procedure to construct sets of antipodal sequences with extremely low ISL;
- an analytical expression for the asymptotic ISL of sets of rotated Legendre sequences.

The second part of the thesis concerns about the implementation of a Analog-to-Information converter that produces a stream of 1-bit measurements and a final resolution after reconstruction that is proportional to the information content of the signal. The main contribution of this part are:

- a new architecture for compressive sensing that produces 1-bit measurements;
- a new reconstruction algorithm for the proposed architecture that exploits not only the sparsity hypothesis but also the hardware architecture of the acquisition system;
- a theoretical analysis of the capabilities of a Delta-Sigma modulator to extract the information content of a signal, that is later extended for the analysis of the proposed architecture;
- a hardware implementation of the proposed architecture, and a measurement setup to validate the theoretical analysis.

We conclude this thesis work with a summary of our findings in Chapter 7.

Part I

LOW CORRELATED SEQUENCES

2. INTEGRATED SIDELOBE LEVEL PROBLEM

2.1 Introduction

The design of sequences sets with good correlation properties is present in many fields of engineering such as radar, sonar, communications, medical imaging and so on. Active sensing applications, have been greatly beneficed by the use of multiple-input multiple-output (MIMO) systems. This kind of systems, transmit orthogonal waveforms via its antennas allowing to achieve a great increase virtual aperture.

As an example, traditional phased-array radar system only transmits a single waveform trough its antennas. However, by the use of MIMO radar system a large increase in parameter identifiability [10], detection performance [11], and resolution [12] can be achieved.

Besides orthogonality, good auto- and cross-correlation properties of the transmitted waveforms are also often required [13, 14, 15].

Good auto-correlation properties means that any sequence in the set is nearly uncorrelated with its own shifted version while good cross-correlation means that any member of the sequences set is nearly uncorrelated with any other members at any shift.

The design of a set of signals with small auto-correlation sidelobes and small cross-correlation between sequences at any time delay ensure that the receiver matched filter can extract the desired information while attenuating undesired signals.

A commonly used metric of the goodness of the correlation is the *Integrated Sidelobe Level* (ISL). The ISL of a set of M sequences each of N (possibly complex) symbols that we will indicate with $x_j^{(p)}$ with $j = 0, \ldots, N-1$ and $p = 0, \ldots, M-1$ is defined as

$$\mathrm{ISL} = \sum_{p=0}^{M-1} \sum_{\substack{k=-N+1 \\ k\neq 0}}^{N-1} \left| X_{x^{(p)}x^{(p)}}(k) \right|^2 + \sum_{p=0}^{M-1} \sum_{\substack{q=0 \\ p\neq q}}^{M-1} \sum_{k=-N+1}^{N-1} \left| X_{x^{(p)}x^{(q)}}(k) \right|^2 \qquad (2.1)$$

where

$$X_{x^{(p)}x^{(p)}}(k) = \sum_{j=\max\{0,-k\}}^{\min\{N-k,N\}-1} x_j^{(p)} x_{j+k}^{*(p)} \qquad k = -N+1 \ldots N-1$$

is is the auto-correlation of the sequence $\boldsymbol{x}^{(p)}$, and

$$X_{x^{(p)}x^{(q)}}(k) = \sum_{j=\max\{0,-k\}}^{\min\{N-k,N\}-1} x_j^{(p)} x_{j+k}^{*(q)} \qquad k = -N+1 \ldots N-1$$

is the cross-correlation between the sequences $\boldsymbol{x}^{(p)}$ and $\boldsymbol{x}^{(q)}$.

Good set of sequences are those having a low ISL value.

Due to the strong interest in the design of sequences with low ISL value, many algorithms have been suggested for its minimization [4, 2, 5, 6, 1]. Such a problem may be far from trivial when constraints are introduced. For example, reception may have to be stopped after a certain time thus spoiling the adoption of periodic signals and leading to the consideration of clipped or aperiodic correlations. Further to that, implementation strongly favors discrete-valued signals, possibly

enforcing quantization to an extremely limited number of levels.

This latter constraint, in particular, is known to make optimization-based methods hard to apply since continuous-optimization must either undergo quantization or be simply discarded in favor of almost exhaustive scans.

Within this scenario, starting from the classical problem of designing an antipodal sequence with a low *Integrated Sidelobe Level* (ISL) [16] we address its generalization to sequence sets, for which "lobes" are considered both for autocorrelation and for cross-correlations.

2.2 Problem Formulation

Given M and N, and based on (2.1) the general problem is that of finding the sequence set minimizing the ISL.

Commonly, a further unimodularity constraint is put on the sequences thus requiring that $|x_j^{(p)}| = 1$ for $p = 0, \ldots, M-1$ and $j = 0, \ldots, N-1$. Such a constraint is application-driven in that it eases the implementation of the transmitters managing the electrical signals corresponding to the sequence symbols. In fact, in this case one may set $x_j^{(p)} = e^{\mathbf{i}\theta_j^{(p)}}$, where $\mathbf{i}$ is the imaginary unit, with $\theta_j^{(p)} \in (-\pi, \pi]$ and design the set of phase sequences $\{\theta_j^{(p)}\}_{j=0}^{N-1}$ for $p = 0, \ldots, M-1$ that can be simply transmitted by constant-envelope modulations.

Given this constraint, it is known that the ISL cannot be decreased below its lower bound [17]

$$\mathrm{ISL}^{\min} = N^2 M(M-1)$$

so that the effectiveness of any approach can be measured in normalized terms by

$$\epsilon = \frac{\mathrm{ISL}^{\min}}{\mathrm{ISL}}$$

better approaches featuring an ϵ closer to 1.

It is well known that sets of unimodular sequences with extremely high effectiveness can be obtained by the application of algorithms [3] that are extensions of those devised to minimize ISL in the single sequence case ($M = 1$) [16].

Yet, when those algorithms meet the even more implementation–friendly constraint of antipodal sequences, i.e. $x_j^{(p)} = \pm 1$ for $p = 0, \ldots, M-1$ and $j = 0, \ldots, N-1$, their effectiveness is largely impaired.

Actually, the antipodal problem is recognized as being much more difficult: a known effect of the impossibility of addressing it with the tools of continuous optimization and the need of resorting to enumeration-based discrete optimization techniques.

In the following we concentrate on antipodal sequences.

Under such an assumption, the particular case $M = 1$ in which only auto-correlation terms appear, has attracted a lot of attention by itself. This led to a conspicuous literature analyzing more than a family of sequences for which ISL or the equivalent *Merit Factor* $\mathrm{MF} = N^2/\mathrm{ISL}$ can be computed analytically at least in the asymptotic case $N \to \infty$ (see, e.g., [18, 19, 20, 21, 22]). Beyond that a list of best known sequences [23] is available for N up to 304.

Our purpose is to develop an analytical expression that may drive optimization in some particular difficult cases, most notably when the antipodal constrain ($x_j^p = \pm 1$) is imposed.

To facilitate the discussion, denote the sum of squares corresponding to the auto-correlation terms as

$$\mathbb{X}_{x^{(p)}\,x^{(p)}} = \sum_{\substack{k=-N+1 \\ k\neq 0}}^{N-1} \left| X_{x^{(p)}\,x^{(p)}}(k) \right|^2 \tag{2.2}$$

and the sum of squares corresponding to the cross-correlation terms as

$$\mathbb{X}_{x^{(p)}\,x^{(q)}} = \sum_{k=-N+1}^{N-1} \left|X_{x^{(p)}\,x^{(q)}}(k)\right|^2 \qquad p \neq q \tag{2.3}$$

so that

$$\mathrm{ISL} = \sum_{p=0}^{M-1} \mathbb{X}_{x^{(p)}\,x^{(p)}} + \sum_{p=0}^{M-1} \sum_{\substack{q=0 \\ p\neq q}}^{M-1} \mathbb{X}_{x^{(p)}\,x^{(q)}} \tag{2.4}$$

A general method for the calculation of $\mathbb{X}_{x^{(p)}\,x^{(p)}}$ of any sequences of odd length is presented in [19, 24]. This method hinges on generating functions and writes correlations as proper sums of their values on the unit circle in the complex plane. The method works well when we have analytical insights on the generating functions.

Extending the ideas of [19], in section 2.3 we devise a general method for the calculation of $\mathbb{X}_{x^{(p)}\,x^{(q)}}$ in (2.3) of any pair of real sequences of odd length and thus, together with the result in [19, 24], the ISL for a set of sequences. In section 3.1 we use this method to obtain an asymptotic expression for the ISL value of a set formed by different rotations of Legendre sequences. Finally, in section 3.2 we propose an optimization procedure based on the latter expression where we find the optimal rotations that minimize the ISL for any sequences length N.

Throughout this chapter we use the following asymptotic notation:

We say that

- two sequences a_N and b_N are asymptotically equivalent, $a_N \sim b_N$ iff

$$\lim_{N\to\infty} \frac{a_N}{b_N} = 1$$

- a_N is asymptotically bounded by b_N, $a_N = O(b_N)$ iff

$$\exists M > 0 \text{ and } \exists N_o \qquad \Big| \qquad |a_N| \leq M\,|b_N| \qquad \forall N > N_o$$

2.3 Calculation of the cross-correlation terms in the ISL

Let $a_0, a_1, \ldots, a_{N-1}$ and $b_0, b_1, \ldots, b_{N-1}$ be two real sequences of length N, we want to obtain an expression for $\mathbb{X}_{ab}$.

If we define the generating functions of the two sequences as

$$Q_a(z) = \sum_{j=0}^{N-1} a_j z^j \qquad\qquad Q_b(z) = \sum_{j=0}^{N-1} b_j z^j$$

we have that

$$Q_a(z)Q_b^*(z) = \sum_{k=-N+1}^{N-1} X_{ab}(k) z^{-k}$$

and thus

$$|Q_a(z)Q_b^*(z)|^2 = \sum_{k=-N+1}^{N-1} \sum_{l=-N+1}^{N-1} X_{ab}(k) X_{ab}(l) z^{-k+l}$$

Now, set $\epsilon_j = e^{\frac{2\pi \mathbf{i}}{N} j}$ and note that for $k, l = -N+1, \ldots, N-1$,

$$\sum_{j=0}^{N-1} \epsilon_j^{-k+l} = \begin{cases} N & \text{if } -l+k = -N, 0, N \\ 0 & \text{otherwise} \end{cases}$$

Hence, if we define

$$S' = \sum_{j=0}^{N-1} |Q_a(\epsilon_j)Q_b^*(\epsilon_j)|^2$$
$$= N \sum_{k=-N+1}^{N-1} X_{ab}^2(k) + N \sum_{k=1}^{2N-1} X_{ab}(k)X_{ab}(k-N) + N \sum_{k=-N+1}^{-1} X_{ab}(k)X_{ab}(k+N)$$

and (for N odd)

$$S'' = \sum_{j=0}^{N-1} |Q_a(-\epsilon_j)Q_b^*(-\epsilon_j)|^2$$
$$= N \sum_{k=-N+1}^{N-1} X_{ab}^2(k) + -N \sum_{k=1}^{2N-1} X_{ab}(k)X_{ab}(k-N) - N \sum_{k=-N+1}^{-1} X_{ab}(k)X_{ab}(k+N)$$

we can express $\mathbb{X}_{ab}$ (i.e. the sum of squares of cross-correlations as in (2.3)) as

$$\mathbb{X}_{ab} = \sum_{k=-N+1}^{N-1} X_{ab}^2(k) = \frac{S' + S''}{2N}$$

To compute S'' we use the Lagrange interpolation polynomials to calculate the values of $Q_a(-\epsilon_j)$ from $Q_a(\epsilon_k)$ for $j, k = 0, \ldots, N-1$. In this special case the data points (ϵ_k) coincide with the complex roots of unity and, for N odd, the Lagrange base polynomials simply reduce to $\frac{2}{N}\frac{\epsilon_k}{\epsilon_j+\epsilon_k}$ [25, p. 89]. Then

$$Q_a(-\epsilon_j) = \frac{2}{N} \sum_{k=0}^{N-1} \frac{\epsilon_k}{\epsilon_j + \epsilon_k} Q_a(\epsilon_k) \tag{2.5}$$

By substituting (2.5) into S'' and developing the product $|Q_a(-\epsilon_j)Q_b^*(-\epsilon_j)|^2$ we get

$$
\begin{aligned}
S'' = & \frac{16}{N^4}\sum_{j=0}^{N-1}\left[\sum_{k_1=0}^{N-1}\frac{\epsilon_{k_1}}{\epsilon_j+\epsilon_{k_1}}Q_a(\epsilon_{k_1})\sum_{l_1=0}^{N-1}\frac{\epsilon_{l_1}^*}{\epsilon_j^*+\epsilon_{l_1}^*}Q_a^*(\epsilon_{l_1})\right. \\
& \left.\sum_{k_2=0}^{N-1}\frac{\epsilon_{k_2}}{\epsilon_j+\epsilon_{k_2}}Q_b(\epsilon_{k_2})\sum_{l_2=0}^{N-1}\frac{\epsilon_{l_2}^*}{\epsilon_j^*+\epsilon_{l_2}^*}Q_b^*(\epsilon_{l_2})\right] \\
= & \frac{16}{N^4}\sum_{k_1=0}^{N-1}\sum_{l_1=0}^{N-1}\sum_{k_2=0}^{N-1}\sum_{l_2=0}^{N-1}Q_a(\epsilon_{k_1})Q_a^*(\epsilon_{l_1})Q_b(\epsilon_{k_2})Q_b^*(\epsilon_{l_2}) \\
& \sum_{j=0}^{N-1}\frac{\epsilon_{k_1}}{\epsilon_j+\epsilon_{k_1}}\frac{\epsilon_{l_1}^*}{\epsilon_j^*+\epsilon_{l_1}^*}\frac{\epsilon_{k_2}}{\epsilon_j+\epsilon_{k_2}}\frac{\epsilon_{l_2}^*}{\epsilon_j^*+\epsilon_{l_2}^*}
\end{aligned}
$$

in which we may exploit the fact that $\epsilon_j^* = 1/\epsilon_j$ to write

$$
\begin{aligned}
S'' = & \frac{16}{N^4}\sum_{k_1=0}^{N-1}\sum_{l_1=0}^{N-1}\sum_{k_2=0}^{N-1}\sum_{l_2=0}^{N-1}\epsilon_{k_1}\epsilon_{k_2}Q_a(\epsilon_{k_1})Q_a^*(\epsilon_{l_1})Q_b(\epsilon_{k_2})Q_b^*(\epsilon_{l_2}) \\
& \times\left\{\sum_{j=0}^{N-1}\frac{1}{\epsilon_j+\epsilon_{k_1}}\frac{\epsilon_j}{\epsilon_j+\epsilon_{l_1}}\frac{1}{\epsilon_j+\epsilon_{k_2}}\frac{\epsilon_j}{\epsilon_j+\epsilon_{l_2}}\right\} \qquad (2.6)
\end{aligned}
$$

Let us define now the innermost sum of (2.6) as

$$
W(k_1,l_1,k_2,l_2) = \sum_{j=0}^{N-1}\frac{1}{\epsilon_j+\epsilon_{k_1}}\frac{\epsilon_j}{\epsilon_j+\epsilon_{l_1}}\frac{1}{\epsilon_j+\epsilon_{k_2}}\frac{\epsilon_j}{\epsilon_j+\epsilon_{l_2}} = \sum_{j=0}^{N-1} f_{k_1,l_1,k_2,l_2}(\epsilon_j)
$$

with

$$
f_{p,q,r,s}(z) = \frac{z^2}{(z+\epsilon_p)(z+\epsilon_q)(z+\epsilon_r)(z+\epsilon_s)}
$$

Depending on p, q, r, s, the rational function $f_{p,q,r,s}(z)$ can be transformed into a specific sum of simple rational parts. Each of these rational parts can be summed

separately. This path is fully developed in [19] and we here exploit the results therein.

In particular we have that

$A)$ for $0 \leq p < N$

$$W(p,p,p,p) = \frac{1}{16}\left(\frac{1}{3}N^4 + \frac{2}{3}N^2\right)\frac{1}{\epsilon_p^2}$$

$B)$ for $0 \leq p \neq q < N$

$$\begin{aligned} &W(p,p,p,q) = W(p,p,q,p) = W(p,q,p,p) = \\ &W(q,p,p,p) = \frac{1}{8}N^2\left(\frac{\epsilon_q + \epsilon_p}{\epsilon_p(\epsilon_q - \epsilon_p)^2}\right) \end{aligned}$$

$C)$ for $0 \leq p \neq q \neq r < N$

$$\begin{aligned} &W(p,p,q,r) = W(p,p,r,q) = W(p,q,p,r) = \\ &W(p,r,p,q) = W(p,q,r,p) = W(p,r,q,p) = \\ &W(q,p,r,p) = W(r,p,q,p) = W(q,r,p,p) = \\ &W(r,q,p,p) = -\frac{1}{4}N^2\frac{1}{\epsilon_q - \epsilon_p}\frac{1}{\epsilon_r - \epsilon_p} \end{aligned}$$

$D)$ for $0 \leq p \neq q < N$

$$W(p,p,q,q) = W(p,q,p,q) = W(p,q,q,p) = -\frac{1}{2}N^2\frac{1}{(\epsilon_p - \epsilon_q)^2}$$

$E)$ for $0 \leq p \neq q \neq r \neq s < N$

$$W(p,q,r,s) = 0$$

Taking into account all the above cases we may write

$$S'' = \frac{16}{N^4}\left(\alpha + \beta + \gamma + \delta\right)$$

, where the terms α, β, γ, and δ correspond to the contributions of the cases A, B, C and D respectively.

For the cases included in $A)$ we have that

$$\alpha = \frac{1}{16}\left(\frac{1}{3}N^4 + \frac{2}{3}N^2\right)\sum_{p=0}^{N-1}\left|Q_a(\epsilon_p)Q_b(\epsilon_p)\right|^2 \tag{2.7}$$

for the cases in $B)$ we have

$$\begin{aligned}\beta = \frac{1}{8}N^2\sum_{\substack{p,q=0\\p\neq q}}^{N-1}\Bigg\{ & \left(\frac{\epsilon_q+\epsilon_p}{\epsilon_p(\epsilon_q-\epsilon_p)^2}\right)\times \\ & \Big[\epsilon_p^2\left|Q_a(\epsilon_p)\right|^2 Q_b(\epsilon_p)Q_b^*(\epsilon_q)+ \\ & \epsilon_p\epsilon_q\left|Q_a(\epsilon_p)\right|^2 Q_b(\epsilon_q)Q_b^*(\epsilon_p)+ \\ & \epsilon_p^2 Q_a(\epsilon_p)Q_a^*(\epsilon_q)\left|Q_b(\epsilon_p)\right|^2+ \\ & \epsilon_q\epsilon_p Q_a(\epsilon_q)Q_a^*(\epsilon_p)\left|Q_b(\epsilon_p)\right|^2\Big]\Bigg\}\end{aligned} \tag{2.8}$$

for $C)$ we have

$$\gamma = \frac{1}{4}N^2 \sum_{\substack{p,q,r=0 \\ p\neq q\neq r}}^{N-1} \left\{ \frac{-1}{(\epsilon_q - \epsilon_p)(\epsilon_r - \epsilon_p)} \times \right. \tag{2.9}$$

$$\Big[\epsilon_p\epsilon_q \left|Q_a(\epsilon_p)\right|^2 Q_b(\epsilon_q)Q_b^*(\epsilon_r) +$$
$$\epsilon_p\epsilon_r \left|Q_a(\epsilon_p)\right|^2 Q_b(\epsilon_r)Q_b^*(\epsilon_q) +$$
$$\epsilon_p^2 Q_a(\epsilon_p)Q_a^*(\epsilon_q)Q_b(\epsilon_p)Q_b^*(\epsilon_r) +$$
$$\epsilon_p^2 Q_a(\epsilon_p)Q_a^*(\epsilon_r)Q_b(\epsilon_p)Q_b^*(\epsilon_q) +$$
$$\epsilon_p\epsilon_r Q_a(\epsilon_p)Q_a^*(\epsilon_q)Q_b(\epsilon_r)Q_b^*(\epsilon_p) +$$
$$\epsilon_p\epsilon_q Q_a(\epsilon_p)Q_a^*(\epsilon_r)Q_b(\epsilon_q)Q_b^*(\epsilon_p) +$$
$$\epsilon_q\epsilon_r Q_a(\epsilon_q)Q_a^*(\epsilon_p)Q_b(\epsilon_r)Q_b^*(\epsilon_p) +$$
$$\epsilon_r\epsilon_q Q_a(\epsilon_r)Q_a^*(\epsilon_p)Q_b(\epsilon_q)Q_b^*(\epsilon_p) +$$
$$\epsilon_q\epsilon_p Q_a(\epsilon_q)Q_a^*(\epsilon_r)\left|Q_b(\epsilon_p)\right|^2 +$$
$$\left.\left.\epsilon_r\epsilon_p Q_a(\epsilon_r)Q_a^*(\epsilon_q)\left|Q_b(\epsilon_p)\right|^2\Big]\right\}$$

and for D)

$$\delta = \frac{1}{2}N^2 \sum_{\substack{p,q=0 \\ p\neq q}}^{N-1} \left\{ \frac{-1}{(\epsilon_p - \epsilon_q)^2} \times \Big[\epsilon_p\epsilon_q \left|Q_a(\epsilon_p)Q_b(\epsilon_q)\right|^2 + \right. \tag{2.10}$$

$$\left.\epsilon_p^2 Q_a(\epsilon_p)Q_a^*(\epsilon_q)Q_b(\epsilon_p)Q_b^*(\epsilon_q) + \epsilon_p\epsilon_q Q_a(\epsilon_p)Q_a^*(\epsilon_q)Q_b(\epsilon_q)Q_b^*(\epsilon_p)\Big]\right\}$$

Summarizing, we can write the sum of squares corresponding to cross–correlations terms of the ISL as

$$\mathbb{X}_{ab} = \frac{1}{2N} \sum_{j=0}^{N-1} \left| Q_a(\epsilon_j) Q_b^*(\epsilon_j) \right|^2 + \frac{16}{N^4} \left(\alpha + \beta + \gamma + \delta \right)$$

where the quantities α, β, γ, δ are defined in (2.7), (2.8), (2.9), (2.10).

With the method presented above in conjunction with the method presented in [19], we can have an analytical expression for the ISL for any set of real sequences of odd length. The computation of the above equations seems to be hard at a first look, but in a number of cases, in particular for sequences from difference sets [24] may lead to significant results.

In the following, we use this method to evaluate the asymptotic trend of the ISL of a set of sequences made up by different Rotations of a Legendre Sequence (RLS set) when N grows to infinity.

3. INTEGRATED SIDELOBE LEVEL OF SETS OF ROTATED LEGENDRE SEQUENCES

3.1 Legendre Sequences

The *Legendre Sequence* (LS) $\ell_0, \ldots, \ell_{N-1}$ exists for any prime N and is defined as

$$\ell_0 = 1$$
$$\ell_j = \begin{cases} 1 & \text{if } j \text{ is a square} \pmod N \\ -1 & \text{if } j \text{ is a nonsquare} \pmod N \end{cases}$$

A LS may be cyclically rotated t_a positions to the left to obtain a *Rotated Legendre Sequence* (RLS) a_j defined as

$$a_j = \ell_{j+t_a \pmod N} = \ell_{j+f_a N \pmod N}$$

with $f_a = t_a/N \in [0, 1]$.

The asymptotic value of $\mathbb{X}_{aa}$ for the family of RLS was calculated in [18] and [19] [1] noting that the asymptotic value of the modulus of the generating function of the LS ($|Q_\ell(\epsilon_j)|$) is independent of j, yielding

[1] The first contribution relies on a "Postulate of Mathematical Ergodicity" to arrive at a result which is formally proved by the second.

$$\frac{\mathbb{X}_{aa}}{N^2} \sim \frac{2}{3} - 4\left|f_a - \frac{1}{2}\right| + 8\left(f_a - \frac{1}{2}\right)^2 \tag{3.1}$$

We follow the same path as in [19] but for the calculation of the cross-correlations terms of the ISL $\mathbb{X}_{ab}$ [26].

To proceed, remember that the generating function of the LS is

$$Q_\ell(\epsilon_j) = \begin{cases} 1 + \ell_j\sqrt{N} & \text{if } j \neq 0 \text{ and } N = 1 \pmod 4 \\ 1 + \mathbf{i}\ell_j\sqrt{N} & \text{if } j \neq 0 \text{ and } N = 3 \pmod 3 \\ 1 & \text{if } j = 0 \end{cases} \tag{3.2}$$

Moreover, if we denote by $Q_a(\epsilon_j)$ the generating function of the RLS $a_j = \ell_{j+t_a \pmod N}$, then

$$Q_a(\epsilon_j) = \epsilon_j^{-t_a} Q_\ell(\epsilon_j)$$

Assume now that the two sequences a_j and b_j are obtained by rotating ℓ_j by, respectively, t_a and t_b positions to the left. We may compute S' as

$$S' = \sum_{j=0}^{N-1} \left|\epsilon_j^{-t_a} Q_\ell(\epsilon_j)\epsilon_j^{t_b} Q_\ell^*(\epsilon_j)\right|^2 = \sum_{j=0}^{N-1} \left|Q_\ell(\epsilon_j)\right|^4$$

from (3.2) we know immediately that $\left|Q_\ell(\epsilon_j)\right|^4 \sim N^2$, then $S' \sim N^3$. Let us now compute the asymptotic values of α, β, γ and δ in (2.7), (2.8), (2.9), (2.10) for any pair of RLS.

- For α in (2.7) we have

$$\alpha = \frac{1}{16}\left(\frac{1}{3}N^4 + \frac{2}{3}N^2\right) S' \sim \frac{1}{48}N^7$$

- For β in (2.8)

$$
\begin{aligned}
\beta = \frac{1}{8}N^2 \sum_{\substack{p,q=0 \\ p\neq q}}^{N-1} \Bigg\{ & \left(\frac{\epsilon_q + \epsilon_p}{\epsilon_p(\epsilon_q - \epsilon_p)^2} \right) \times \\
& \Big[\epsilon_p^2 \left| Q_\ell(\epsilon_p) \right|^2 \epsilon_{p-q}^{t_b} Q_\ell(\epsilon_p) Q_\ell^*(\epsilon_q) + \\
& \epsilon_p \epsilon_q \left| Q_\ell(\epsilon_p) \right|^2 \epsilon_{q-p}^{t_b} Q_\ell(\epsilon_q) Q_\ell^*(\epsilon_p) + \\
& \epsilon_p^2 \epsilon_{p-q}^{t_a} Q_\ell(\epsilon_p) Q_\ell^*(\epsilon_q) \left| Q_\ell(\epsilon_p) \right|^2 + \\
& \epsilon_q \epsilon_p \epsilon_{q-p}^{t_a} Q_\ell(\epsilon_q) Q_\ell^*(\epsilon_p) \left| Q_\ell(\epsilon_p) \right|^2 \Big] \Bigg\} \\
\sim \frac{1}{8}N^2 \sum_{\substack{p,q=0 \\ p\neq q}}^{N-1} \Bigg\{ & \left(\frac{\epsilon_q + \epsilon_p}{\epsilon_p(\epsilon_q - \epsilon_p)^2} \right) \times \\
& \Big(N^2 \ell_p \ell_q \epsilon_p^2 \epsilon_{p-q}^{t_b} + N^2 \ell_p \ell_q \epsilon_p \epsilon_q \epsilon_{q-p}^{t_b} + \\
& N^2 \ell_p \ell_q \epsilon_p^2 \epsilon_{p-q}^{t_a} + N^2 \ell_p \ell_q \epsilon_p \epsilon_q \epsilon_{q-p}^{t_a} \Big) \Bigg\} \\
= \frac{1}{8}N^4 \sum_{\substack{p,q=0 \\ p\neq q}}^{N-1} \Bigg\{ & \left(\frac{\ell_p \ell_q}{(1 - \epsilon_{p-q})^2} \right) \times \\
& \Big(\epsilon_{p-q}^{t_b+1} + \epsilon_{p-q}^{t_b+2} + \epsilon_{p-q}^{1-t_b} + \epsilon_{p-q}^{-t_b} + \\
& \epsilon_{p-q}^{t_a+2} + \epsilon_{p-q}^{t_a+1} + \epsilon_{p-q}^{1-t_a} + \epsilon_{p-q}^{-t_a} \Big) \Bigg\}
\end{aligned}
$$

$$= \frac{1}{8}N^4 \sum_{\substack{k=-N+1 \\ k\neq 0}}^{N-1} (X_{\ell\ell}(k) + X_{\ell\ell}(N-k)) \frac{\epsilon_k^{t_b+1} + \epsilon_k^{t_b+2} + \epsilon_k^{1-t_b} + \epsilon_k^{-t_b} + \epsilon_k^{t_a+2} + \epsilon_k^{t_a+1} + \epsilon_k^{1-t_a} + \epsilon_k^{-t_a}}{(1-\epsilon_k)^2}$$

Note that $X_{\ell\ell}(k) + X_{\ell\ell}(N-k)$ is the periodic correlation [24] of the LS. Then, from [18] and [27] we know that $|X_{\ell\ell}(k) + X_{\ell\ell}(N-k)| \leq 3$ for Legendre sequences. Then, using the fact that $\sum_{k=1}^{N-1} \frac{1}{|1-\epsilon_k|^2} = O(N^2)$ (see (3.3) and (3.6) below and set $t = 0$), and using the triangle inequality we get that $\beta = O(N^6)$.

- For the calculation of γ in (2.9), following the same steps we did for β we have

$$\gamma \sim \frac{1}{4}N^4 \sum_{\substack{p,q,r=0 \\ p\neq q\neq r}}^{N-1} \Bigg\{ - \frac{\ell_q \ell_r}{(1-\epsilon_{p-q})(1-\epsilon_{p-r})} \Big(\epsilon_{p-r}\epsilon_{q-r}^{t_b} + \epsilon_{p-q}\epsilon_{q-r}^{-t_b} + \epsilon_{p-r}^{t_b+1}\epsilon_{p-q}^{t_a+1} + \epsilon_{p-r}^{t_a+1}\epsilon_{p-q}^{t_b+1} + \epsilon_{p-r}^{-t_b}\epsilon_{p-q}^{t_a+1} + \epsilon_{p-r}^{t_a+1}\epsilon_{p-q}^{-t_b} + \epsilon_{p-q}^{-t_a}\epsilon_{p-r}^{-t_b} + \epsilon_{p-r}^{-t_a}\epsilon_{p-q}^{-t_b} + \epsilon_{p-r}\epsilon_{q-r}^{t_a} + \epsilon_{p-q}\epsilon_{q-r}^{-t_a} \Big) \Bigg\}$$

$$= \frac{1}{4}N^4 \sum_{\substack{u,v=-N+1 \\ u\neq v\neq 0}}^{N-1} \Bigg\{ - \frac{X_{\ell\ell}(v-u) + X_{\ell\ell}(N-(v-u))}{(1-\epsilon_v)(1-\epsilon_u)} \Big(\epsilon_u\epsilon_{u-v}^{t_b} + \epsilon_v\epsilon_{u-v}^{-t_b} + \epsilon_u^{t_b+1}\epsilon_v^{t_a+1} + \epsilon_u^{t_a+1}\epsilon_v^{t_b+1} + \epsilon_u^{-t_b}\epsilon_v^{t_a+1} + \epsilon_u^{t_a+1}\epsilon_v^{-t_b} + \epsilon_v^{-t_a}\epsilon_u^{-t_b} + \epsilon_u^{-t_a}\epsilon_v^{-t_b} + \epsilon_u\epsilon_{u-v}^{t_a} + \epsilon_v\epsilon_{u-v}^{-t_a} \Big) \Bigg\}$$

and again we have that $\gamma = O(N^6)$

- For δ in (2.10) we have

$$
\begin{aligned}
\delta =& \frac{1}{2}N^2 \sum_{\substack{p,q=0 \\ p\neq q}}^{N-1} \Bigg\{ \frac{-1}{(\epsilon_p - \epsilon_q)^2} \times \Big[\epsilon_p \epsilon_q \left| Q_\ell(\epsilon_p) Q_\ell(\epsilon_q) \right|^2 + \\
& \qquad \epsilon_p^2 \epsilon_{p-q}^{t_a} Q_\ell(\epsilon_p) Q_\ell^*(\epsilon_q) \epsilon_{p-q}^{t_b} Q_\ell(\epsilon_p) Q_\ell^*(\epsilon_q) + \\
& \qquad \epsilon_p \epsilon_q \epsilon_{p-q}^{t_a} Q_\ell(\epsilon_p) Q_\ell^*(\epsilon_q) \epsilon_{q-p}^{t_b} Q_\ell(\epsilon_q) Q_\ell^*(\epsilon_p) \Big] \Bigg\} \\
\sim & -\frac{1}{2}N^4 \sum_{\substack{p,q=0 \\ p\neq q}}^{N-1} \left\{ \frac{\epsilon_{q-p} + \epsilon_{q-p}^{-t_a-t_b} + \epsilon_{q-p}^{1-t_a+t_b}}{(1-\epsilon_{q-p})^2} \right\} \\
=& -\frac{1}{2}N^4 \sum_{\substack{k=-N+1 \\ k\neq 0}}^{N-1} \frac{\left(\epsilon_k + \epsilon_k^{-t_a-t_b} + \epsilon_k^{1-t_a+t_b} \right)}{(1-\epsilon_k)^2} (N - |k|) \\
=& -N^4 \sum_{k=1}^{N-1} \frac{\left(\epsilon_k + \epsilon_k^{-t_a-t_b} + \epsilon_k^{1-t_a+t_b} \right)}{(1-\epsilon_k)^2} (N - |k|)
\end{aligned}
$$

Larger values of the summand are those for k close to 1, which make the denominator close to zero and numerator $\sim cN$ for some constant c (for k close to $N-1$, the denominator becomes also close to zero but the numerator is $O(1)$).

Exploiting this and using the small angle approximation for the complex exponential, we may write

$$
\delta \sim -N^5 \sum_{k=1}^{N-1} \frac{\epsilon_k + \epsilon_k^{-t_a-t_b} + \epsilon_k^{1-t_a+t_b}}{-\frac{4\pi^2}{N^2}k^2} \tag{3.3}
$$

To continue, we recall the definition of the Dilogarithm function and its series expansion [28] valid for $|z| \leq 1$

$$\mathrm{Li}_2(z) = -\int_0^1 \frac{ln(1-zt)}{t}dt = \sum_{k=1}^{\infty}\frac{z^k}{k^2} \tag{3.4}$$

Taking the real part of (3.4) and evaluating on the unit circle gives [28, eq. (8.7)]

$$\mathrm{Re}\left\{\mathrm{Li}_2\left(e^{\mathbf{i}\theta}\right)\right\} = \mathrm{Re}\left\{\sum_{k=1}^{\infty}\frac{e^{\mathbf{i}k\theta}}{k^2}\right\} = \frac{1}{6}\pi^2 - \frac{1}{4}|\theta|(2\pi - |\theta|) \tag{3.5}$$

Exploiting (3.5) and concentrating on the first period $0 \leq \frac{t}{N} \leq 1$ we obtain

$$\mathrm{Re}\left\{\sum_{k=1}^{\infty}\frac{\epsilon_k^t}{k^2}\right\} = \pi^2\left[\frac{1}{6} - \left[\frac{t}{N}\right]_1\left(1 - \left[\frac{t}{N}\right]_1\right)\right] \tag{3.6}$$

where $[\cdot]_1 = \cdot \pmod 1$.

Hence, since we know that δ is real

$$\begin{aligned}
\delta \sim & \frac{1}{4}N^7\left\{\frac{1}{6} + \frac{1}{6} - \left[-\frac{t_a+t_b}{N}\right]_1\left(1 - \left[-\frac{t_a+t_b}{N}\right]_1\right) + \right. \\
& \left. \frac{1}{6} - \left[\frac{t_b-t_a}{N}\right]_1\left(1 - \left[\frac{t_b-t_a}{N}\right]_1\right)\right\} \\
= & \frac{1}{4}N^2\left\{\frac{1}{2} - [-f_a-f_b]_1(1-[-f_a-f_b]_1) - \right. \\
& \left. [f_b-f_a]_1(1-[f_b-f_a]_1)\right\} \\
= & \frac{1}{4}N^2\left\{\frac{1}{2} - [f_a+f_b]_1(1-[f_a+f_b]_1) - \right. \\
& \left. [f_a-f_b]_1(1-[f_a-f_b]_1)\right\}
\end{aligned}$$

where we have defined $f_a = \frac{t_a}{N}$ and $f_b = \frac{t_b}{N}$. Then, exploiting the symmetries of a

quadratic form of a modulus function we have for $0 \leq f_a, f_b \leq 1$

$$[f_a + f_b]_1 (1 - [f_a + f_b]_1) = \frac{1}{4} - \left(|f_a + f_b - 1| - \frac{1}{2} \right)^2$$
$$[f_a - f_b]_1 (1 - [f_a - f_b]_1) = \frac{1}{4} - \left(|f_a - f_b| - \frac{1}{2} \right)^2$$

so that

$$\delta \sim \frac{1}{4} N^7 \left[\left(|f_a + f_b - 1| - \frac{1}{2} \right)^2 + \left(|f_a - f_b| - \frac{1}{2} \right)^2 \right]$$

Based on the above we are now interested in computing the asymptotic value of

$$\begin{aligned} \frac{1}{N^2} \mathbb{X}_{ab} =& \frac{1}{2N^3} (S' + S'') \sim \frac{1}{2N^3} \left[N^3 + \frac{16}{N^4} (\alpha + \beta + \gamma + \delta) \right] \\ \sim& \frac{2}{3} + 2 \left(|f_a + f_b - 1| - \frac{1}{2} \right)^2 + 2 \left(|f_a - f_b| - \frac{1}{2} \right)^2 \end{aligned} \tag{3.7}$$

Going back to our original problem for calculation of the ISL value of a set of M sequences $x_j^{(p)}$ with $j = 0, \ldots, N-1$ and $p = 0, \ldots, M-1$, where each $x^{(p)}$ is made by a different rotation f_p of a LS (RLS set), replacing (3.1) and (3.7) into (2.4) we finally have that

$$\begin{aligned} \frac{\mathrm{ISL}}{N^2} \sim & \sum_{p=0}^{M-1} \frac{2}{3} - 4 \left| f_p - \frac{1}{2} \right| + 8 \left(f_p - \frac{1}{2} \right)^2 + \\ & \sum_{p=0}^{M-1} \sum_{\substack{q=0 \\ p \neq q}}^{M-1} \frac{2}{3} + 2 \left(|f_p + f_q - 1| - \frac{1}{2} \right)^2 + 2 \left(|f_p - f_q| - \frac{1}{2} \right)^2 \end{aligned} \tag{3.8}$$

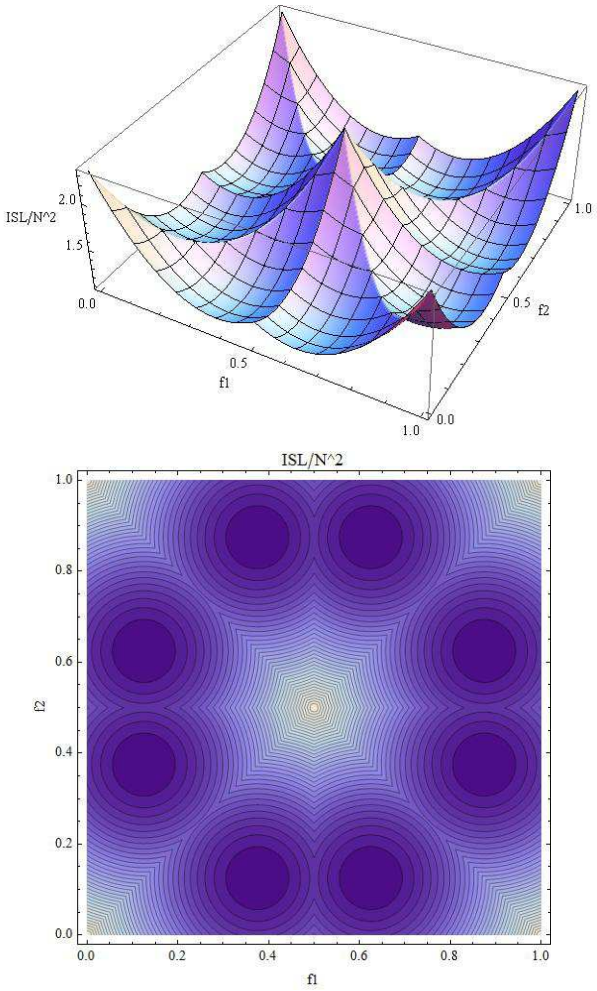

Fig. 3.1: Plots of ISL for $M = 2$ as a function of f_1 and f_2: top: 3D-view, bottom: iso-ISL lines

As an example, Figure 3.1 reports the 3D and contour plot of the right-hand side of (3.8) for $M = 2$. Direct visual inspection of that Figure confirms that minima exists and can be easily identified. In the next section we will exploit this result where an optimization procedure is developed to find the optimal rotations that minimize the ISL for any sequences length N.

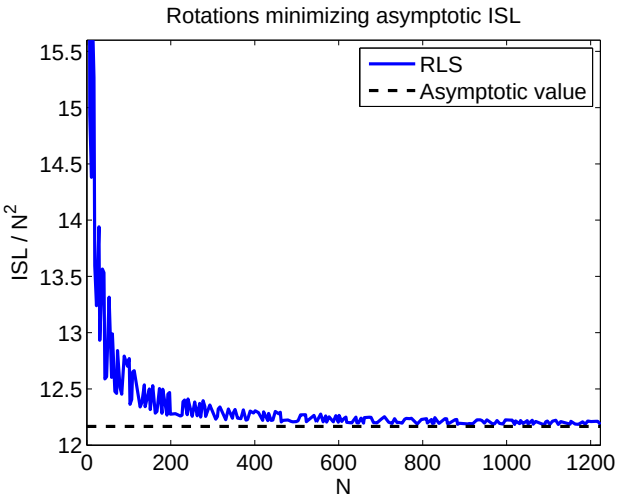

Fig. 3.2: Integrated Sidelobe Level as a function of sequence length. In blue, RLS with rotations minimizing asymptotic ISL; in black, asymptotic value.

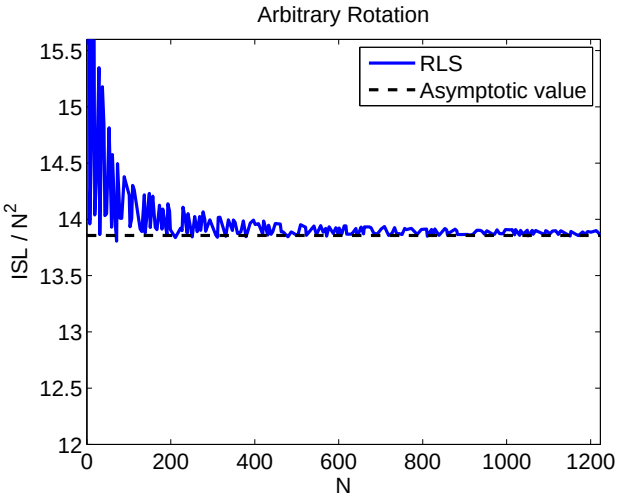

Fig. 3.3: Integrated Sidelobe Level as a function of sequence length. In blue, RLS with an arbitrary rotation; in black, asymptotic value.

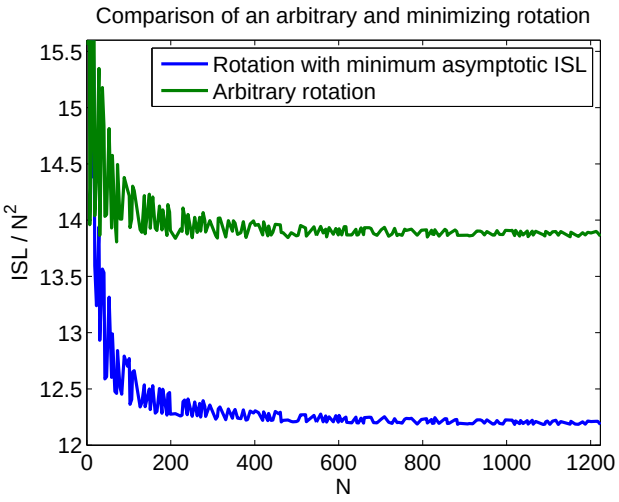

Fig. 3.4: Integrated Sidelobe Level as a function of sequence length. Comparison of an RLS with an arbitrary rotation and RLS with rotations minimizing asymptotic ISL.

As another example, in Figure 3.2, 3.3, 3.4 we plot the ISL for $M = 4$ as a function of the sequence length N.

In Figure 3.2 the values of rotation are those that minimize the asymptotic ISL, while in case Figure 3.3 we use an arbitrary rotation. In both cases we can see that the trend of the plots is in agreement with the asymptotic value calculated. In Figure 3.4, we plot together both curves to show that the one that achieves the minimum asymptotic value of ISL, also achieves the minimum ISL value for sequences length greater than approximately 20. For different choices of rotations and different number of sequences (M), the behavior is the same than presented.

3.2 *Sets of RLS minimizing* ISL

The key idea [29] is to set $x_j^{(p)} = \ell_j^{(f_p)}$ for properly chosen rotations f_p, $p = 0, \ldots, M-1$.

Since only N values for each f_p lead to distinct rotations, a complete scan requires "only" $\binom{N}{M}$ trials that, though far from the exponential explosion of a full scan (that would entail 2^{MN} trials) may soon become prohibitive.

To cope with larger values of N we may resort to the asymptotic analysis made in the previous section.

From equation (3.8) we see that asymptotic ISL is invariant if we change f_p into $1 - f_p$ for any p. Therefore, by assuming $f_0 \leq f_1 \leq \cdots \leq f_{M-1} \leq 1/2$ one may resolve all absolute values and easily compute the rotation values for which $\frac{\partial \mathrm{ISL}}{\partial f_p} = 0$. This yields

$$f_p = \frac{2p+1}{4M} \tag{3.9}$$

that result in a minimum attainable $\mathrm{ISL} = N^2\left[M(M-1) + \frac{1}{6}\right]$ and thus, in a performance figure

$$\epsilon^{\mathrm{RLS}} = \frac{6M(M-1)}{6M(M-1)+1} \tag{3.10}$$

indicating that, for large N, the performance of a set of RLS should be within 8% of the maximum possible, approaching it very rapidly as M increases.

Based on these asymptotic considerations it is easy to devise a much faster scan that drastically reduces the number of trials by considering for the j-th rotation only a narrow interval of possible values around $\frac{2p+1}{4M}$. Since the length of such an interval may be decreased as N increases, the resulting search burden goes from $\binom{N}{M}$ trials to $\theta(N)^M$ with $\theta(N)$ a function rapidly approaching a constant as N increases (experimentally we verified $\theta(N) \simeq 20$ for N larger than 200).

The results of such a scan yields the Optimum RLS set (ORLS) whose performance is compared with that of other known algorithms or sequence families in the following Section.

3.3 Numerical results

Beyond ORLSs that exist for every prime N, we consider

- Random sequences, that exist for any N and are generated by assigning $x_j^{(p)} = \pm 1$ with uniform probability and independently for each $p = 0, \ldots, M-1$ and $j = 0, \ldots, N-1$. For each N and M we generate 10^4 sets of sequences and record the best achieved performance.

- Gold sequences, that exist when $N = 2^q - 1$ for some integer q and are obtained from the well known maximum–length sequences to maintain low correlation and simultaneously be able to produce sets of sequences with relatively large cardinality. Though they are produced by linear-feedback shift registers, Gold sequences are designed to enjoy the same properties of random sequences. For each N we draw $10^3 \times NM$ M-tuples of Gold sequences at random from those available, and we record the least ISL.

- Q.CAN sequences, that exists for any N and are obtained by the CAN algorithm described in [3] when quantization is applied at the end of the iterative procedure. For the case of 1-bit quantization the option of leaving the algorithm operate with continuous phases and quantize only the final result, has been discarded , after experimentally verifying that it was leading to poorer performance. In this case, quantization was applied at every step of the iterative procedure.

- Optimally Rotated Best Sequences (ORBS) that leverage on the fact that for each N up to 304 one or more sequences are recognized as the state-of-the art solution to ISL minimization problem for $M = 1$ (some of them are known to be the true optimum solutions, some others are only the best *known* solutions). For each of those sequences, we build a set of M sequences by trying all the possible relative rotations and selecting the set of rotations yielding the minimum ISL.

In figure 3.5 we evaluate how state-of-the-art algorithm for the synthesis of ultra low-ISL sequences is affected when quantization is imposed. Note how the performance of Q.CAN is hardly impaired for low quantization depth, and reach the performance of ORLS when the number of quantization levels is greater than 13 for $M = 4$, 18 for $M = 4$ and 22 for $M = 12$ for the considered cases.

For different choices of M and N we have seen a similar trend, and that the performance of Q.CAN is only better than that of ORLS when the number of quantization bits of grater than 4, with an increasing trend as M increases. This shows a great advantage in the use of ORLS.

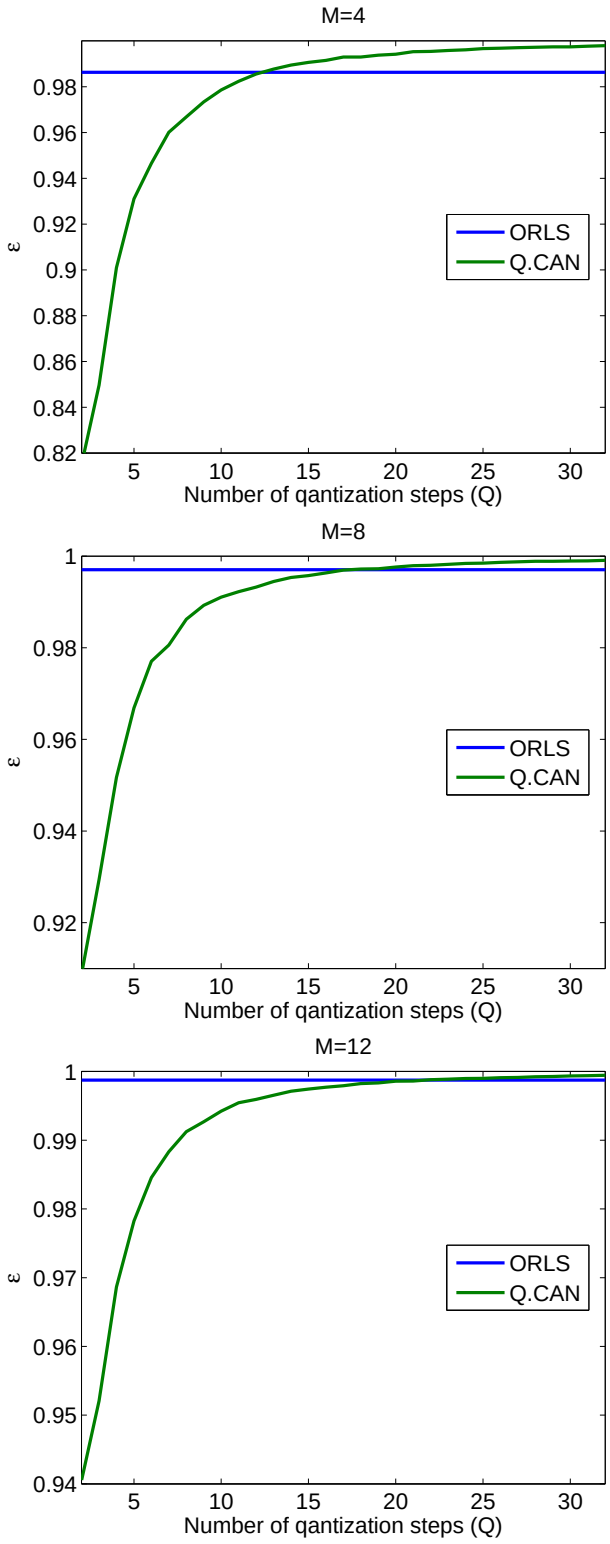

Fig. 3.5: Comparison between ORLS, with Q.CAN algorithm when quantization is imposed for (a) $M = 4$, (b) $M = 8$ and (c) $M = 12$, and fix value of $N = 1033$. The performance metric as a function of the number of quantization steps. The performance of Q.CAN exceeds the performance of the binary ORLS for Q grater than 13 for $M = 4$, 18 for $M = 4$ and 22 for $M = 12$.

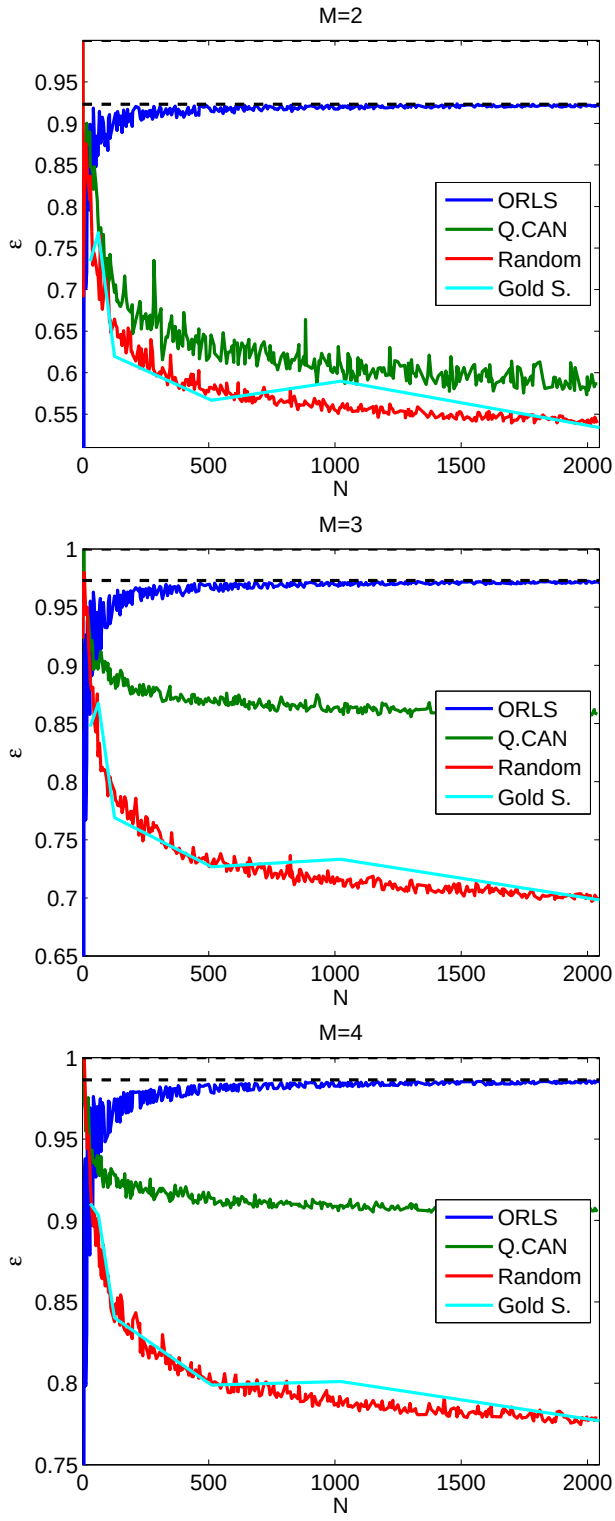

Fig. 3.6: Comparison between ORLS, random sequences, Gold sequences and Q.CAN sequences with binary quantization for (a) $M = 2$, (b) $M = 3$ and (c) $M = 4$. The solid horizontal line at 1 identifies the theoretical maximum performance while the dashed horizontal line marks the asymptotic performance achieved by RLS.

Figure 3.6 compares the performance of ORLS with that of random, Gold and Q.CAN sequences with binary quantization for $M = 2$, $M = 3$ and $M = 4$. Note how ORLS clearly outperform the other techniques for all reasonably large N (say for $N > 100$) also revealing a distinct improving trends approaching the theoretical limit as N increases.

On the contrary the performance of random, Gold, and Q.CAN sequences exhibits a clear decreasing trend. According to expectations, since Gold sequences are designed to mimic a random behavior, the corresponding performances follow an analogous trend.

Finally, though insufficient to reach ORLS, the optimization implicit in the construction of Q.CAN sequences make the corresponding performance clearly superior to that of random-like sequences.

In Figure 3.7 we compare each ORLS with the corresponding ORBS and with Q.CAN sequences with binary quantization for $M = 2$, $M = 3$ and $M = 4$. Again, ORLS perform uniformly better than ORBS for sufficiently large N; additionally ORBS do not exhibit a definite improvement with respect to Q.CAN at least for $M > 2$. This shows that the good performance of the proposed ORLS is only partially due to the exploitation of sequences that feature a good autocorrelation properties but also hinges on a structural property of Legendre Sequences.

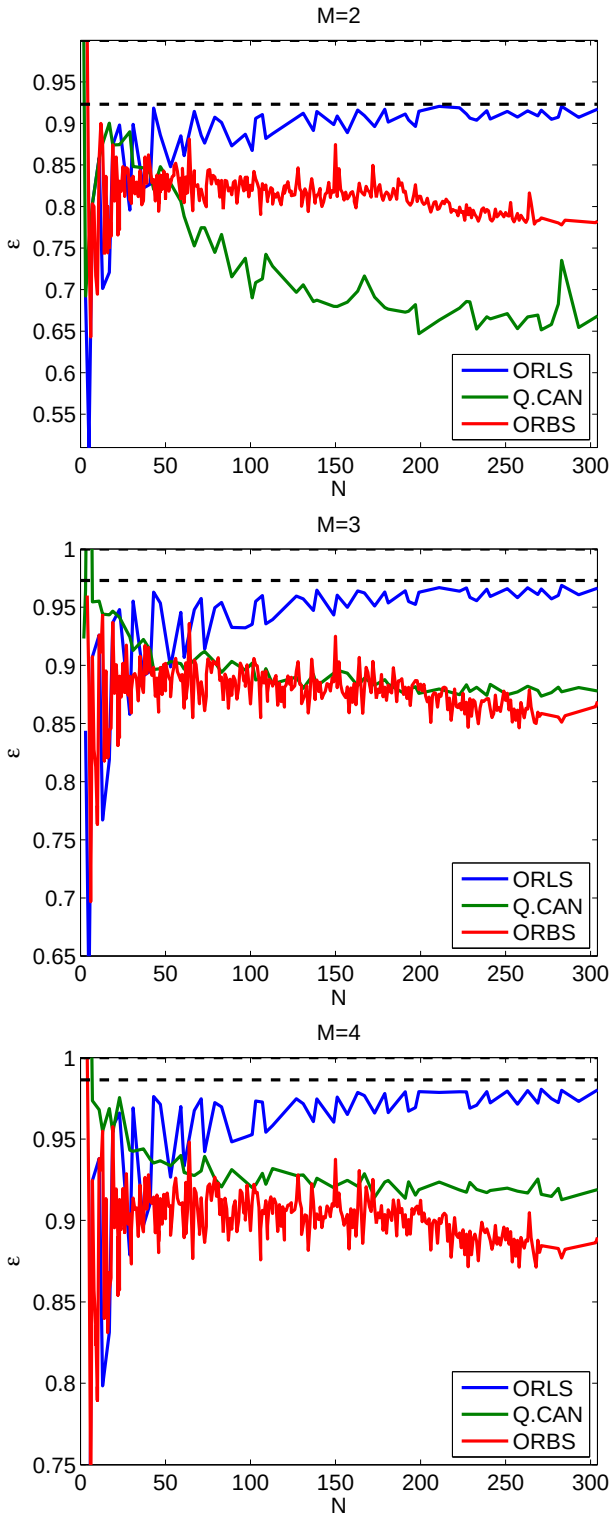

Fig. 3.7: Comparison between ORLS, ORBS and Q.CAN sequences for (a) $M = 2$, (b) $M = 3$ and (c) $M = 4$. The solid horizontal line at 1 identifies the theoretical maximum performance while the dashed horizontal line marks the asymptotic performance achieved by RLS.

3.4 Conclusion

We apply a method based on generating functions, which has already been proposed for the calculation of the ISL of a sequence, to the calculation of the cross-correlation components of the ISL of a set of sequences.

The apparent complexity of the resulting expressions can be tackled in the asymptotic conditions for sequences whose generating function has a relatively simple trend. Since this is the case of Legendre sequences, we are able to derive an analytical expression for the asymptotic ISL of sets of rotated Legendre sequences.

Based on the later result, we propose a simple procedure to construct sets of antipodal sequences with extremely low ISL. Each sequence in the set is a different rotation of the Legendre Sequence of the same length. Optimal rotations are found by an exhaustive scan whose complexity is greatly reduced by exploiting the asymptotic result yielding a general expression for the trend of the ISL of sets of infinitely long sequences.

The resulting performance largely exceeds that of classical methods for the direct generation of low-ISL sets of antipodal sequences. The method we propose also outperforms a well-known algorithm able to generate extremely-low ISL sets of unimodular continuous-phase sequences, which is nevertheless impaired by the strong quantization needed to satisfy antipodality constraint.

Part II

ANALOG TO INFORMATION CONVERSION

4. SIGNAL MODELS

4.1 Introduction

The classical acquisition approach based on the Nyquist-Shannon theorem states that for any analog band-limited signal, all its information content can be acquired by taking uniformed distributed samples at a rate that doubles the signal bandwidth.

While this is one of the fundamentals theorems of Signal Processing, by taking advantage on certain structures of the signal, a much clever acquisition strategy can be develop in order to reduce the number of measurements and still acquire its full information content.

In order to exploit the peculiarities of a given class of signal, we must be able to properly represent those signals of interest with accurate models. This models are useful to incorporate previous knowledge of a given class of signal, and to distinguish them from other classes of maybe no interest.

Many classes of signals, especially when representing physical signals, can be modeled to have a linear structure, i.e., if we sum two signals that belongs to that class, the new signal will also belong to the same class.

We will treat signals as real-valued functions having domains that are either continuous or discrete. In the case of a discrete signals, we can simply view them as vectors in N-dimensional Euclidean space $\mathbb{R}^N$.

For bandlimited analog signals with no frequency components above $N/2$, or Nyquist rate equal to N, we will also represent them as vectors with dimension

equal to its Nyquist rate. Both representation are equivalent in the sense that one can pass to another with standard techniques (sinc interpolation).

Note that the space dimension N of both kinds of signals described above defines the degrees of freedom they have. In particular, although analog signals can be more efficiently represented by other "representations", any analog bandlimited signal has at most N degrees of freedom, and we have choose this model in order to be able to directly compared with Nyquist-based acquisition.

Let Ψ denotes the $N \times N$ matrix with columns given by the set $\{\psi_i\}_{i=1}^N$. If the vectors in this set are linear independent, then they span a basis in $\mathbb{R}^N$, and any vector in this spaces has a unique representation as a linear combination of the elements of that basis.

For any $\boldsymbol{x} \in \mathbb{R}^N$ there exist $\boldsymbol{s} \in \mathbb{R}^N$ such that

$$\boldsymbol{x} = \Psi \boldsymbol{s} = \sum_{i=1}^{N} s_i \psi_i$$

For analog signals, note that this representation is equivalent to:

$$x(t) = \sum_{i=1}^{N} s_i \psi(t)_i$$

, where the set of continuous time waveforms $\{\psi(t)_i\}_{i=1}^N$ are the sinc-interpolated signals obtains from the vectors in $\{\psi_i\}_{i=1}^N$ (or equivalently, the vector in $\{\psi_i\}_{i=1}^N$ are form by taking samples of the waveform in $\{\psi(t)_i\}_{i=1}^N$ at a rate N).

4.2 Sparse Signals

With the models given above, we are able to represent any linear signal (discrete or analog) of dimensionality equal to N, and with N degrees of freedom. However, many natural signal that are found in real situations have a smaller number of

degrees of freedom with respect to its dimensionality. In other words, not all possible vectors in $\mathbb{R}^N$ represents valid signals for a given class.

Many natural signals can be expressed as a linear combination of only just a few vectors from a given basis. This class of signals are called to be sparse signals, since only a small amount of its coefficients, when represented on that basis, are different from zero. The information content of this class of signals is concentrated only on the values of the non-zero components and on the position of those components.

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

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

, its sparsity $\operatorname{spar}(\boldsymbol{a})$ (sometimes indicated as L_0 norm) as the cardinality of $\operatorname{supp}(\boldsymbol{a})$, and its usual p-norm as

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

We will assume that a suitable basis exists whose vectors are the columns of the $N \times N$ matrix Ψ, and that the signal of interest is K-sparse, which means that for any instance of $\boldsymbol{x}$ there is an N-dimensional vector $\boldsymbol{s}$ such that $\boldsymbol{x} = \Psi \boldsymbol{s}$ and $\operatorname{spar}(\boldsymbol{s}) \leq K$.

Although the sparse model given above is of broad interest, it is difficult to find real life signals to be truly sparse. However, many natural signals can be very well approximated by sparse models. This classes of signals are called to be compressible signals, and can be approximated by setting the smallest components to zero and keeping the biggest K.

In the following, we will treat compressible signals and sparse signals as to have a simple sparse representation. The error produced by this approximation

will be considered as if the sparse signal would have an intrinsic noise independently of the source where it is generated.

5. COMPRESSIVE SENSING

5.1 Introduction

The newly introduced paradigm of Compressive Sensing (CS) [7, 8, 9] exploits special signal features to extract its information content with a smaller amount of samples (or measurements in the general case) with respect to acquisition based on the Nyquist-Shannon sampling theorem.

According to the sampling theorem, we can perfectly reconstruct any bandlimited signal by its samples provided that the sampling rate exceeds twice the maximum frequency in the bandlimited signal. However, as we have seen before, the information content of some classes of signals is concentrated in only few coefficients for a given representation.

Taking advantage on the knowledge of the structure of the signal, more sophisticated sampling methods can be developed in order to reduce the number of samples necessary to reconstruct the signal. Compressive sensing theory exploits the "sparsity" representation in order to reduce well below the number of measurements stated by the Nyquist-Shannon theorem, and still be able to perfectly reconstruct the original signal.

Reducing the number of measurements has noteworthy advantages. It can reduce the hardware complexity, storage capacity, power consumption, channel bandwidth, etc.

In the compressive sensing framework, few nonadaptive linear measurements of the signal are taken, i.e. projections of the signal over vectors of a given basis.

Based on these projections, by means of a non-linear algorithm, it is possible to recover the signal.

To make the discussion more concrete, consider the general case where the signal $\boldsymbol{x} \in \mathbb{R}^N$ is measured through M inner products of the form:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{e}$$

where $\boldsymbol{y} \in \mathbb{R}^M$ is the measured vector, Φ is an $M \times N$ measurement matrix, and $\boldsymbol{e} \in \mathbb{R}^M$ is a vector representing measurement noise.

In general, given $M < N$, the matrix Φ represents a dimensionality reduction, i.e., it maps a vector in $\mathbb{R}^N$ into a vector in $\mathbb{R}^M$. Under this condition, there are infinite different signals $\boldsymbol{x}$ that satisfy the above equation given the measurements $\boldsymbol{y}$.

At this point there are two main questions to be done: $a)$ Under what conditions the application of the matrix Φ preserves the information of the signal $\boldsymbol{x}$? How it is recovered the original signal $\boldsymbol{x}$ from the reduced set of measurements $\boldsymbol{y}$?

We will try to answer these question in the following sections.

5.2 The Restricted Isometry Property

To partially answer the first question, lets first write the vector $\boldsymbol{x}$ as

$$\boldsymbol{x} = \Psi \boldsymbol{s}$$

, and

$$\boldsymbol{y} = \Phi \Psi \boldsymbol{s} + \boldsymbol{e} = \Theta \boldsymbol{s} + \boldsymbol{e}$$

Relaying on the a-priori knowledge that $\mathrm{spar}\,(\boldsymbol{s}) \leq K$, it is possible to define

a subset of $\mathbb{R}^N$ containing all the interesting instances of $\boldsymbol{x}$. Then, the acquisition mechanism should map this subset into the measurement space $\mathbb{R}^M$ "quasi-bijectively" in a sense that will be made more precise in the following.

One of the most striking, and useful, facts that appear at this point is that, when sparsity is one of the priors, if Θ 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 Θ provides, with high probability, the needed "quasi-bijection".

More formally, we say that a matrix Θ is a *restricted isometry* [30] when there is a constant $0 \leq \delta_K < 1$ such that

$$(1 - \delta_K) \|\boldsymbol{s}\|_2^2 \leq \|\Theta \boldsymbol{s}\|_2^2 \leq (1 + \delta_k) \|\boldsymbol{s}\|_2^2$$

whenever $\mathrm{spar}\,(\boldsymbol{s}) \leq K$. 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 Θ is made of independent random entries characterized by a sub-Gaussian distribution then, with an overwhelming probability, the matrix Θ is a restricted isometry with a constant δ provided that [31, 32, 33]

$$M \geq CK \log(N/K) \tag{5.1}$$

where C is some constant depending on each instance.

If Θ is a restricted isometry, once that $\mathrm{supp}\,(\boldsymbol{s})$ is known, we may restrict Θ to that domain and obtain an injective mapping. If the measurements in $\boldsymbol{y}$ 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 $\boldsymbol{s}$.

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 Θ and design a system in which Φ 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 Φ depending on the specific Ψ but relies on randomness to implicitly "scan" all possible sparsity bases.

5.3 *CS Reconstruction Algorithms*

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

Though reconstruction mechanisms may 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 [34].

We will here concentrate on the most frequently adopted methods, and note that those techniques fall in one of two categories: optimization-based reconstruction [30, 35, 36, 37, 38, 39] and iterative support-guessing reconstruction [40, 41, 42, 43, 44, 45].

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

5.3.1 *Optimization Based Reconstruction Algorithms*

The key fact behind optimization-based methods is that, among all the possible counterimages $\boldsymbol{s}$ of the vector $\boldsymbol{y} = \Theta\boldsymbol{s}$ the one that we are looking for is the "most

sparse", i.e., the one for which $\mathrm{spar}\,(\boldsymbol{s})$ is minimum.

Since we usually have $\mathrm{spar}\,(\boldsymbol{s}) \leq K \ll N$ this assumption is sensible. Moreover, it leads to some beautiful results on the possibility of recovering $\boldsymbol{s}$ by means of simple optimization problems [30].

More formally, it can be shown that, if Θ is a restricted isometry with constant $\delta \leq \sqrt{2} - 1$ then the $\hat{\boldsymbol{s}}$ solution of the optimization problem

$$\begin{aligned} \min \quad & \|\hat{\boldsymbol{s}}\|_1 \\ \text{s.t.} \quad & \|\Theta\hat{\boldsymbol{s}} - y\|_2 \leq \epsilon \end{aligned} \tag{5.2}$$

is such that

$$\|\hat{\boldsymbol{s}} - \boldsymbol{s}\|_2 \leq C\epsilon$$

for some constant $C > 0$.

Hence, if we use ϵ 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) we 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 ϵ 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 $\|\Theta\hat{\boldsymbol{s}} - \boldsymbol{y}\|_2 \leq \epsilon$ with $\Theta\hat{\boldsymbol{s}} = \boldsymbol{y}$ 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 $\boldsymbol{u} = (1, \ldots, 1)^\top$ and by introducing the auxiliary unknown

vector $\boldsymbol{w} = (w_0, \ldots, w_{n-1})^\top$ to write

$$\begin{aligned} \min \quad & \boldsymbol{u}^\top \boldsymbol{w} \\ & \Theta \hat{\boldsymbol{s}} = \boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{w} \geq 0 \\ & -\boldsymbol{w} \leq \hat{\boldsymbol{s}} \leq \boldsymbol{w} \end{aligned} \tag{5.3}$$

where vector inequalities are meant to hold component-wise.

The equality constraints in (5.3) 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, we know that the true value of the j-th measurement 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 Δy_j the corresponding quantization step.

Hence, in presence of a coarse quantization, it is sensible to substitute the equality constraints $\Theta \hat{\boldsymbol{s}} = \boldsymbol{y}$ in (5.3) with $\boldsymbol{y} - \Delta \boldsymbol{y}/2 \leq \Theta \hat{\boldsymbol{s}} \leq \boldsymbol{y} + \Delta \boldsymbol{y}/2$, where $\Delta \boldsymbol{y} = (\Delta y_0, \ldots, \Delta y_{m-1})^\top$. 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 Δy_j to be substantially different one from the other.

It is interesting to note that optimization-based reconstruction algorithms work 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.

5.3.2 Support-Guessing Reconstruction Algorithms

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

Iterative support-guessing methods are, in general, procedures that alternate a rough, non-necessarily sparse, solution of $\boldsymbol{y} = \Theta\boldsymbol{s}$ from which an estimate of $\mathrm{supp}(\boldsymbol{s})$ 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` [40] that has some definite advantages. First, it works for matrices Θ that are restricted isometries and, if K is known and the isometry constant δ_{2K} for vectors with $2K$ non-zero components can be bounded by $\delta_{2K} \leq 0.025$, then, given a tolerance $\epsilon > 0$, the reconstructed vector $\hat{\alpha}$ satisfies

$$\|\hat{\boldsymbol{s}} - \boldsymbol{s}\|_2 \leq C \max\left\{\epsilon, \frac{\|\boldsymbol{s}'\|_2}{\sqrt{K}} + \|\Phi\nu\|_2\right\}$$

where $\boldsymbol{s}'$ is the vector that can be obtained by $\boldsymbol{s}$ 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 $\boldsymbol{s}$ is known and that the isometry constant δ_{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 $\boldsymbol{s}$ 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.

5.4 Analog-to-Information Converters

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

- **encoder**: a hardware system 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 Φ 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 Φ, reconstructs the original signal.

In this section we will discuss various strategies for designing systems for acquiring compressive measurements of real-world signals.

Note that, in practical implementation, we do not want to communicate Φ to the decoder and thus most often exploit 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 the two stages happens along a digital channel thus implying an ADC block with a finite range (we will assume $[-V^{\max}, V^{\max}]$ for a certain $V^{\max}$) and a finite number of levels.

In the following we 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 is actually a "bit budget" since it may be partitioned into digital words of

different depths corresponding to different measurements. Additionally, in many applications the total number of bits is constrained, which suggests a tradeoff between the number of measurements and the number of bits per measurement.

5.4.1 Random-Modulation-Pre-Integration – RMPI

This is probably the most straightforward implementation of compressive sensing concepts [46].

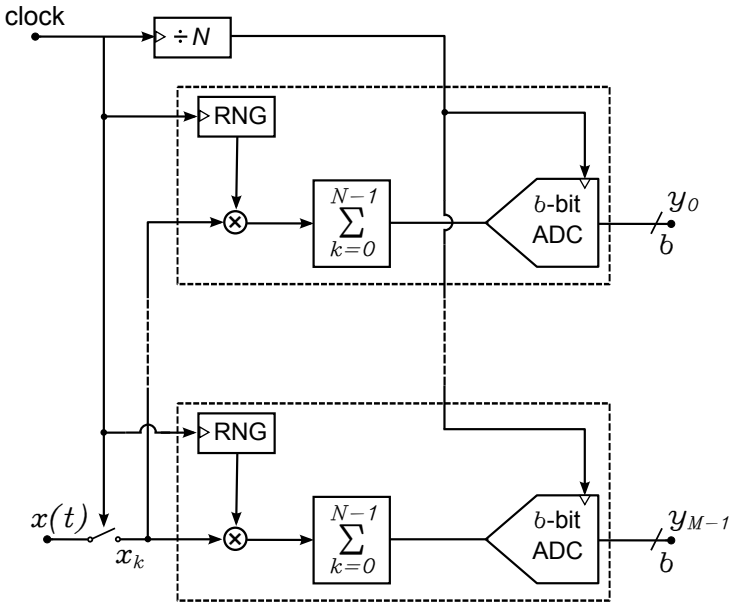

Fig. 5.1: Block scheme of an RMPI encoder. The samples of the input signal are multiplied by M different random sequences and accumulated up to time N. The accumulated values are then quantized by a b bit AD converter.

With reference to Figure 5.1 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 allegedly $y_j = \sum_{k=0}^{N-1} \Phi_{j,k} x_k$, and relying on the following assumptions:

- $\boldsymbol{x}$ and Φ are independent stochastic processes;
- the $\Phi_{j,k}$ are *independent* and *identically distributed* (either Gaussian or binary antipodal) random variables, with zero mean and unity variance;
- the energy of x in the accumulation time window is constant;

the random variable y_j will converge to a normal random variable independently of the input signal x.

Given the above observation, the measurements y obtained with an RMPI architecture will have a range that is potentially $\sqrt{N}$-times larger than that of x (e.g., $\pm 3\sigma$ around the signal average). When comparing an RMPI solution with a direct application of a Nyquist based AD converter, and considering a uniform quantizer in both cases, in order to maintain the same amount of quantization error the number of bits needs to be increased for an RMPI implementation. Moreover, 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.

On one hand, RMPI architecture allows to reduce the number of measurements for the acquisition of a given class of signals with respect to classical Nyquist based sampling. On the other hand, in order to obtain a given performance in terms of reconstruction error, the number of bits needed to encode each of the measurement would be bigger than for Nyquist based acquisition. This suggests a tradeoff between the number of measurements M and the number of bits per measurement b.

RMPI architecture presents a direct implementation of the compressive sensing concepts developed in this section. However, as it has been shown, some design consideration are needed to be taken. More precisely, the choice of a proper AD converter is of crucial importance in order to obtain a given performance. Moreover, RMPI architecture requires the used of a huge amount of circuitry (continuous-time or discrete-time analog multiply-and-accumulate blocks, multibit AD converters) leading to an expensive system implementation in terms of cost, power consumption, and design effort.

5.4.2 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 rate). Compressive sensing architectures 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.

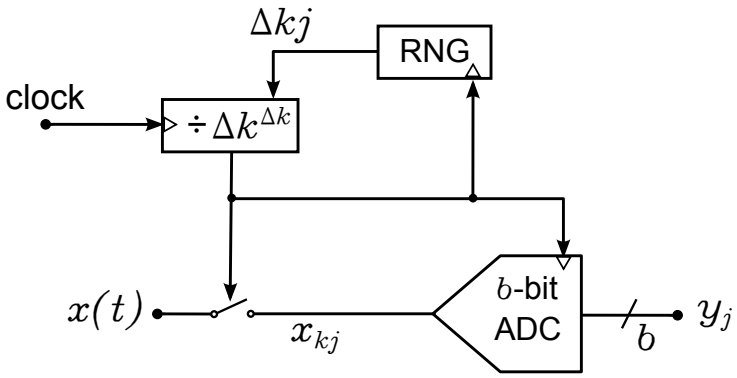

Fig. 5.2: Block scheme of an random sampling encoder. The samples are taken at random positions in time, over a predefined grid.

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 5.2 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 $x' = (x'_0, \ldots, x'_{\upsilon N-1})^\top$ that oversamples a bandlimited $x(t)$ by a factor υ with respect to $x = (x_0 \ldots, x_{N-1})^\top$ containing the Nyquist samples. The two vectors x' and x are linked by $x' = \boldsymbol{A}x$, being $\boldsymbol{A}$ an upsampling matrix.

With this, the $M \times N$ matrix Φ is nothing but the product $\Phi = \boldsymbol{PA}$, where $\boldsymbol{P}$ 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

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

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 $\Delta k_{\min}$ we 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) we 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.

RSAM is only subject 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.

5.4.3 1-bit Compressive Sensing - 1bRMPI

Given a total bit budget B, the trade-off between the number of measurements M and the number of bits $b = B/M$ spent to encode each of them is a classical theme in signal acquisition and coding and applies also to CS architectures.

Among other issues, it may help coping with the unavoidable saturation of the ADC since the extreme solution $b = 1$ identifies the ADC with a pure saturation centered in 0, thus completely eliminating the problem.

In particular, RMPI systems may be optimized in each particular setting to see how much information in our original signal can be inserted into B bits [47] and is possible to think that each measurement is represented by a single bit encoding its sign [48, 49].

There are several benefits to the 1-bit CS technique. Given that the quantizer can be implemented as a simple comparator that merely tests if a measurement is

above or below zero, an efficient hardware quantizer can be built to operate at high speeds. Furthermore, 1-bit quantizers do not suffer from dynamic range issues nor linearity problems inherent of the implementation of a multibit AD converter.

Since signs give no hint on the magnitude of the involved signals, the problem in (5.2), with $\epsilon = 0$, with $y = \mathrm{sign}(\Theta\hat{\alpha})$ and where the $\mathrm{sign}(\cdot)$ operator applied component-wise, is recast into [48]

$$\begin{aligned} \min \quad & \|\hat{\alpha}\|_1 \\ \text{s.t.} \quad & y \circ \Theta\hat{\alpha} \geq 0 \\ & \|\Theta\hat{\alpha}\|_2 = 1 \end{aligned} \tag{5.4}$$

where $\circ$ stands for component-wise product[1] and the second, unit-energy constraint is introduced as a scale-fixing prior. This approach is referred in [48] as 1-bit CS.

The above optimization problem is a non-convex problem and must be addressed by specialized algorithms. Two state-of-the-art algorithms were presented in the lasts years to address this problem. The *Restricted Step Shrinkage* [50] that will be indicated here as `RSS`, and the *Binary Iterative Hard Thresholding* [51], indicated here as `BIHT`. These algorithm are proved to achieve a higher average recovery SNR, and are an order of magnitude faster than other previous proposed algorithms in [48] and [49].

Regrettably, even with `BIHT`, typical performance of an 1bRMPI architecture are largely inferior with respect to multibit RMPI or RSAM solutions for comparable bit budgets.

[1] so that this result in a set of M component-wise inequalities

6. $RADS$ CONVERTER

6.1 Introduction and Motivation

As it has been shown in the previous sections, there is a trade-off between the number of measurements needed to uniquely identify a given class of signals, and the number of bits that is necessary to represent them in order to obtain a given precision [47].

On one hand, under certain assumptions on the signal structure, compressive sensing theory allows to reduce the number of measurement by increasing the hardware architecture complexity. On the other hand, Delta-Sigma converters allows to reduced the number of bits per measurement, even to the extreme case of only 1-bit, by increasing the number of measurements and mixing time encoding information.

The fundamental question is: is it possible to combine the advantages of both theories in one single device that allows to reduce the total number of bits in a measurement, and simplify the hardware system implementation?

The answer to this question is YES, and it is what we have called *The $RADS$ Converter* [52, 53, 54].

In this section we will introduce the $RADS$ Converter which constitutes the main contribution of this thesis. We will start by describing its hardware architecture, and modeling the operations performed by the architecture in the frequency domain. This model will lead to an intuitive understanding of its working principle, and will give some insight on how the decoding stage can be efficiently

implemented.

We will next introduce a time domain analysis that, starting by the analysis of a Δ/Σ modulator, and followed by the analysis of the whole $RADS$ Converter architecture, will lead to a deeper understanding of capabilities of the system.

We will close this section by presenting a set of measurements performed on an "off-the-shelf" implementation of the $RADS$ Converter that constitutes a proof of concept of the proposed architecture, and we discuss how it can be efficiently implemented on a single silicon device.

In order to evaluate the performance of the converter, we have extensively appealed to numerical simulations. Performance is evaluated by matching the reconstructed vector $\hat{\boldsymbol{s}}$ with the original vector $\boldsymbol{s}$ and using two merit figures: the Probability of Support Reconstruction (PSR) and the Reconstruction Signal-to-Noise Ratio (RSNR), i.e.,

$$\mathrm{PSR} \;=\; \Pr\left\{\mathrm{supp}\left(\boldsymbol{s}\right) \subseteq \mathrm{supp}_{\min\{\boldsymbol{s}\}/5}\left(\hat{\boldsymbol{s}}\right)\right\}$$

$$\mathrm{ARSNR(dB)}=\mathbf{E}\left[\mathrm{dB}\left(\frac{\|\boldsymbol{s}\|_2^2}{\|\boldsymbol{s}-\hat{\boldsymbol{s}}\|_2^2}\right)\right]=\mathbf{E}\left[\mathrm{dB}\left(\frac{\|x\|_2^2}{\|x-\hat{x}\|_2^2}\right)\right]$$

where the thresholded support is conventionally defined as

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

Probabilities and expectations were estimated using Monte Carlo simulations for which statistics was gather after 5000 trials.

6.1.1 Preliminaries: Delta-Sigma Modulation

In this subsection we will make a short review of the main concepts that applies to Delta-Sigma (Δ/Σ) modulators.

Let model a basic 1^{st}-order Δ/Σ modulator structure as in Figure 6.1 where the block [Q] represents a general quantizer and the block [D] represents a one time-step delay.

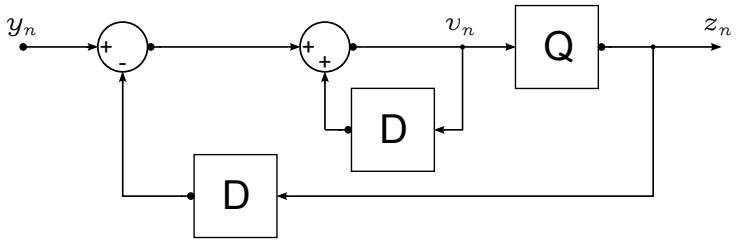

Fig. 6.1: A time-domain block diagram of a first order Δ/Σ modulator.

The input sequence y feeds the Δ/Σ modulator that produces a lower resolution output sequence z at every time step n.

The quantization stage of the modulator is usually implemented with a very low resolution quantizer. Single bit quantizers are the most common option for the implementation of this kind of converters, since it is particularly appealing for hardware implementations. The quantizer takes the form of a comparator to zero, an extremely inexpensive and fast hardware device. Furthermore, 1-bit quantizers do not suffer from dynamic range issues (the sign of the measurement remains valid even if the quantizer saturates).

Though beneficial, 1-bit quantization is a very non-linear operation that makes difficult to obtain simple models for the operation of the converter. In order to provide an insight into the operation of the modulator, the analysis is usually tackled in the z-domain [55, 56, 57, 58], for which the quantizer has been replaced by it linear model as shown in Figure 6.2.

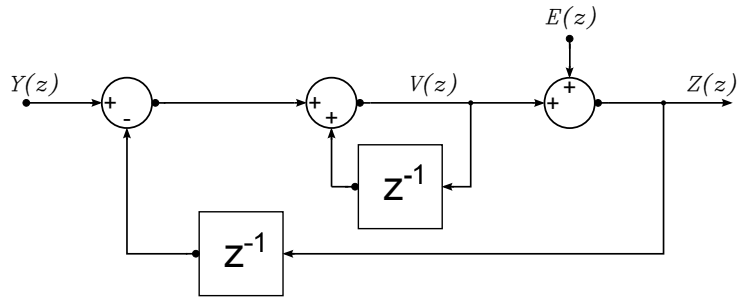

Fig. 6.2: A z-domain linear model of a first order Δ/Σ modulator.

From the diagram we can write

$$V(z) = z^{-1}V(z) + X(z) - z^{-1}Z(z)$$

Thus

$$Z(z) = V(z) + E(z) = z^{-1}V(z) + X(z) - z^{-1}Z(z) + E(z)$$

and rearranging we get

$$Z(z) = X(z) + (1 - z^{-1})E(z) \tag{6.1}$$

Equation (6.1) can be written in the general form

$$Z(z) = STF(z)X(z) + NTF(z)E(z) \tag{6.2}$$

where the STF refers to the *Signal Transfer Function*, that in this case is unity, and the NTF refers to the *Noise Transfer Function* and is equal to

$$NTF = 1 - z^{-1} \tag{6.3}$$

Equation (6.2) is the basic equation Δ/Σ modulator, and shows how the output can be expressed as a sum of a term accounting for the signal, and a term accounting for the quantization noise.

For the case presented above, the NTF has clearly a high pass response, which suppresses the quantization noise near dc, and amplifies it out of the signal band. This is the so called noise shaping capabilities of the Δ/Σ modulators.

By replacing z in equation (6.3) by $e^{\mathrm{i}2\pi f/M}$, where M is the sampling frequency, the *power spectral density* (PSD) of the output noise is found to be

$$S_q(f) = 2(\sin(\pi f/M))^2 S_e(f)$$

, where $S_e(f)$ is the PSD of the quantization noise of the internal quantizer of the converter.

Consider a signal bandwidth of B Hertz, and approximate $S_e(f) = 2e_{rms}^2/M$. By integrating $S_e(f)$ in the signal band, we get that the in-band noise, i.e., the quantization noise present in the signal band, can be approximated as

$$q_{rms} = e_{rms}\frac{\pi}{\sqrt{3}}\left(\frac{M}{B}\right)^{-3/2} \tag{6.4}$$

As it can be seen from equation (6.4), the in band noise decreases with increasing the oversampling ratio, i.e., the ratio between the sampling frequency and the signal bandwidth.

In oder to increase resolution, by replacing the quantizer stage in the block diagram of Figure 6.1 by a new copy 1^{st}-order Δ/Σ modulator, we will get a second order Δ/Σ modulator. This procedure can be continued to obtain an L^{th}-order Δ/Σ modulator.

By extending the analysis we have made for the 1^{st}-order, we can get a basic expression for the NTF of an L^{th}-order Δ/Σ modulator as

$$NTF = (1 - z^{-1})^L \tag{6.5}$$

By integrating the above equation in the signal band, we get that the power of quantization noise of an L^{th}-order Δ/Σ modulator is

$$q_{rms} = e_{rms} \frac{\pi^L}{\sqrt{2L+1}} \left(\frac{M}{B}\right)^{-(L+\frac{1}{2})} \tag{6.6}$$

The equation given above is an approximation for the calculation of the in-band quantization noise of an Δ/Σ modulator. This approximation does not take into consideration quantizer overload thus increasing the total power of quantization noise. Moreover, for higher order modulators, it is possible to change the shape of the NTF to produce different behavior. However, for the sake of concreteness, it is enough the analysis made so far.

6.2 RADS Converter architecture

The RAndom Delta-Sigma ($RADS$) Converter illustrated in Figure 6.3 is nothing but a conventional Δ/Σ converter whose input signal is pre-multiplied by a random sequence of symbols. $RADS$ Converter exploits the noise shaping capabilities of Delta Sigma (Δ/Σ) structures and produce a number of measurements ($M \geq N$) each coarsely quantized (actually with only 1 bit). The use of $RADS$ Converter with a proper exploitation of sparsity gives as a result a substantial compression in the number of acquired bits with respect to classical acquisition or to simply Δ/Σ modulation. The simplicity of the architecture also allows to operate at very high frequencies, making possible, for example, to acquire frequency sparse signals that are spread over a large bandwidth with a very high resolution.

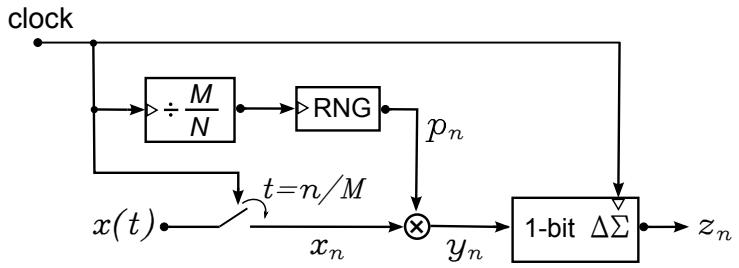

Fig. 6.3: Block scheme of a $RADS$ Converter. The input signal is multiplied by a random sequence and fed into a Δ/Σ converter made of a 1-bit ADC and a loop filter in charge of noise shaping.

The loop filter and the nonlinear dynamics of the $\Delta\Sigma$ produce a progressive encoding of widening windows from the original signals so that there is no one-to-one relation between single bits and projections.

On one hand, such a technique has the desired effect to allow squeezing amplitude information into a sequence of sign informations. On the other hand, its nonlinearity avoids the writing of a simple linear model linking the signal samples x_n with the bits produced by the encoder z_n. Although, this is formally true, it

will be necessary to waive the detailed modeling of the $\Delta\Sigma$ operations to concentrate on its high-level functionality of oversampling converter with noise-shaping abilities. In doing so, we will obtain a model that can be effectively plugged into reconstruction algorithms.

Without any lost of generality, we may focus on a normalized acquisition time of one second, and model the signal $x(t)$ to be sampled at the Nyquist rate N by collecting $x_k = x\left(\frac{k}{N}\right)$ for $k = 0, \ldots, N-1$. Clearly $\boldsymbol{x} \in \mathbb{R}^N$ and $\boldsymbol{x}$ is sparse if there is an $N \times N$ matrix Ψ such that $\boldsymbol{x} = \Psi \boldsymbol{s}$ for some vector $\boldsymbol{s} \in \mathbb{R}^N$ in which at most $K << N$ components are non-zero.

Given that the analog waveform $x(t)$ corresponding to the samples in $\boldsymbol{x}$ is sampled at frequency M, that in general is larger than N, defining the oversampled signal $x'_n = x\left(\frac{n}{M}\right)$ for $n = 0, \ldots, M-1$ we can link the two vectors $\boldsymbol{x}'$ and $\boldsymbol{x}$ by a linear operation $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$, being $\boldsymbol{A}$ an upsampling matrix that considers the components of $\boldsymbol{x}$ as the Nyquist samples of a bandlimited signal. Hence, Δ/Σ operations do not apply to the original components of the vector but to a vector oversampled by a factor M/N.

The sinc-interpolation matrix $\boldsymbol{A} \in \mathbb{R}^{M \times N}$ is defined as:

$$\boldsymbol{A}_{j,k} = \operatorname{sinc}\left[\frac{N-1}{M-1}(j-1) - (k-1)\right] \qquad \begin{matrix} j = 1, \ldots, M \\ k = 1, \ldots, N \end{matrix}$$

and for the case of $N = M$ we have that $A = I$.

Note that since we are dealing with 1-bit measurements we have $M = B$, so oversampling does not imply an increment in the total number of bits.

With reference to Figure 6.3, the samples in $\boldsymbol{x}'$ are multiplied by a Nyquist-rate random sequence $p_1, p_2, \ldots, p_N$. Applying a further linear operator indicated with the symbol $\boldsymbol{P}$ which is defined by

$$\boldsymbol{P}_{j,k} = \begin{cases} p_{\lceil j\frac{N}{M} \rceil} & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \tag{6.7}$$

, therefore, the input of the $\Delta\Sigma$ is the vector $\boldsymbol{Px}' = \boldsymbol{PAx}$.

The binary output of the Δ/Σ at time n can be expressed as the sum of the corresponding input sample and a term accounting for the quantization noise which spectral profile is dictated by the Noise Transfer Function (NTF) of the converter loop [55]. Hence

$$\boldsymbol{z} = \boldsymbol{PAx} + \zeta$$

where ζ accounts for the quantization noise introduced by the Δ/Σ converter.

Conventional Δ/Σ approaches have $\boldsymbol{P}$ equal to the identity and exploit this construction by noting that low-pass filtering $\boldsymbol{z}$ is equivalent to low-pass filtering $\boldsymbol{PAx} + \zeta = \boldsymbol{Ax} + \zeta$ and thus invert upsampling to recover $\boldsymbol{x}$ with an error equal to the low-pass filtering of ζ, a term that can be made very small by playing with the NTF, i.e., making it as high-pass as possible given other implementation constraints.

In our case, the matrix $\boldsymbol{P}$ is designed to introduce spreading in order to allow that higher frequency components of the upsampled signal enter the baseband range in which the bits in $\boldsymbol{z}$ are processed. This alias normally prevents signal reconstruction. Yet, sparsity can be exploited to counter alias and allow the acquisition of signal components that would otherwise fall out of the reach of the $\Delta\Sigma$ range (or, conversely, allow smaller oversampling to acquire the same signal).

Figure 6.4 shows the spectrum at different points of the system. For simplicity, assume that the spectrum of p_n is "flat" in the interval $(-N, N)$ and negligible outside that interval. Note that y is now band limited to $\frac{3}{2}N$ and that, depending on the value of the sampling frequency M, the replicas of the spectrum will alias on the discrete time signal y_n.

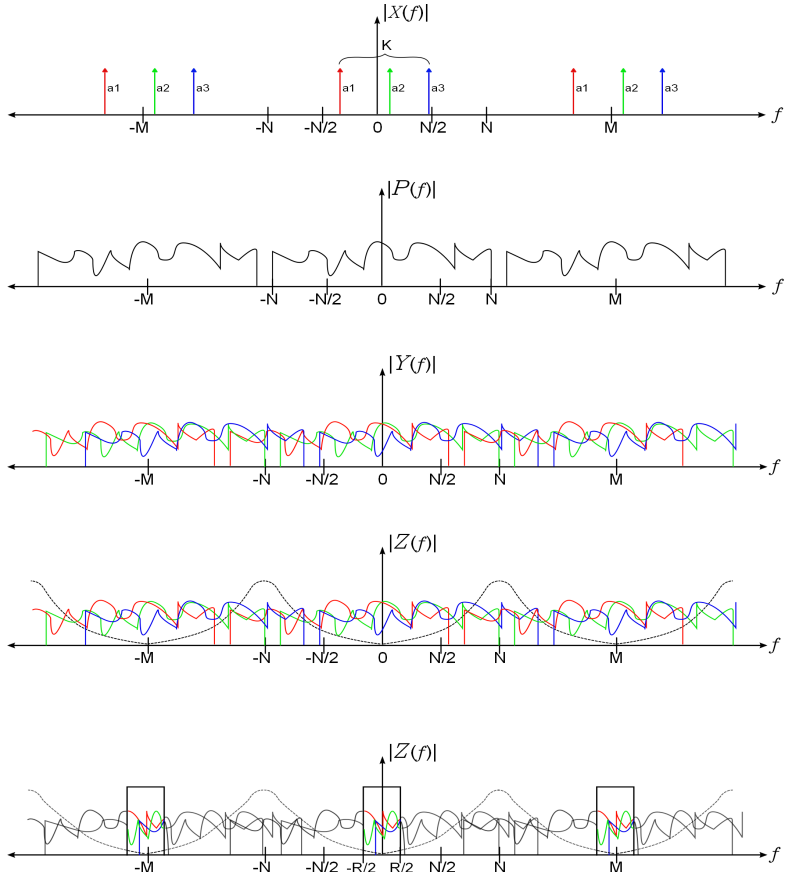

Fig. 6.4: Frequency occupancy at the different point of the system. From top to bottom: spectrum of the input sparse signal in the Fourier domain x; spectrum of the modulating signal p; spectrum of the modulated signal y as a sum of different shifts of the modulating signal; spectrum of the output signal z with the addition of the quantization noise shaped by the NTF of the Δ/Σ modulator; remaining spectrum after low-pass filtering.

Given that we are multiplying the input signal by a pseudorandom sequence, with a very high probability and independent of the sparsity basis, the resulting signal after the modulation will be spread over a large bandwidth. Furthermore, since the rate of change of the pseudorandom sequence is equal to the Nyquist rate of the input signal, there will be always a contribution of every component of

the original signal into the low part portion of the bandwidth.

6.3 Decoding and reconstruction

In order to reconstruct the original signal $x(t)$ form the 1-bit samples z_n, sparsity is only one of the two priors we have, the other is the high-pass nature of the disturbance introduced by the Δ/Σ modulator. This further piece of information allows us to remove the biggest amount of energy of the quantization noise, while leaving enough information to reconstruct the original $x(t)$ in the low-pass portion of the spectrum. Note that while signal energy decreases linearly as the band shrinks around DC, disturbance energy decreases polynomially thanks to the NTF of the modulator [55].

The block diagram shown in Figure 6.5 is used to recover the original signal $x(t)$ from the 1-bit samples z_n. The left-hand part of the block diagram is a low-pass filter that removes the biggest contribution of the quantization noise, and it is followed by a decimation operation that removes redundant samples.

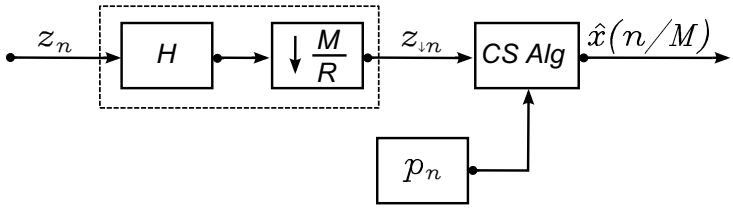

Fig. 6.5: Decoding and reconstruction scheme for $RADS$ Converter. The 1-bit input signal is first filtered, decimated, and then processed by a compressive sensing reconstruction algorithm.

Consider a low-pass filter with a cutoff frequency $R/2$. Depending on the ratio R/N, and considering a perfect filter with a decimation operation that leaves only R significant samples, the remaining samples form a system of linear equations, that for $R/N < N$ (which is the most common case) is undetermined. This system of equations can efficiently be solved by the right-hand part of the diagram that represents any of the CS reconstruction algorithm we have seen in section 5.3 in the previous chapter.

Note that the band between $-R/2$ and $R/2$ contains the contribution of all possible shifts of the spectrum of p_n determined by the frequency values present in $x(t)$, as it is shown at the bottom of Figure 6.4. If spreading were not applied before Δ/Σ modulation, only a portion of the signal would enter in such a band.

To determine the value of the cutoff frequency of the filter $R/2$, it is desirable to take R as small as possible which contributes to remove the quantization noise produced by the Σ/Δ converter.

On the other hand, the signal obtained after filtering should contain enough information for the recovery of the original sparse signal. In other words, the number of significant components of the filtered signal must be large enough to guarantee that the CS reconstruction algorithm can recover the original signal with a high probability of success while removing as much noise as possible. The correct choice of the bandwidth R will determine the system performance.

To model the filtering process we use an l-order FIR filter ($l \leq M$) and arrange its coefficients $h_1, h_2, \ldots, h_l$ as the rows of a matrix $\boldsymbol{H}$ of $M \times M$ elements.

$$\boldsymbol{H}_{j,k} = \begin{cases} h_i & \text{if } j = k + i - 1 \\ 0 & \text{if } j \neq k + i - 1 \end{cases} \qquad \begin{matrix} j, k = 1, \ldots, M \\ i = 1, \ldots, l. \end{matrix}$$

As an example with $M = 8$ and $\boldsymbol{h} = [h1, h2, h3]$

$$\boldsymbol{H} = \begin{bmatrix} h_1 & & & & & & & \\ h_2 & h_1 & & & & & & \\ h_3 & h_2 & h_1 & & & & & \\ & h_3 & h_2 & h_1 & & & & \\ & & h_3 & h_2 & h_1 & & & \\ & & & h_3 & h_2 & h_1 & & \\ & & & & h_3 & h_2 & h_1 & \\ & & & & & h_3 & h_2 & h_1 \end{bmatrix}.$$

Encoding the downsampling operator in the matrix

$$\boldsymbol{D}_{j,k} = \begin{cases} 1 & \text{if } j = \frac{R}{M}k \\ 0 & \text{if } j \neq \frac{R}{M}k \end{cases} \qquad \begin{matrix} j = 1, \ldots, R \\ k = 1, \ldots, M \end{matrix}$$

we may finally link the sparse vector $\boldsymbol{s}$ with the filtered and downsampled measurement $\boldsymbol{z}_{\downarrow}$ as

$$\boldsymbol{z}_{\downarrow} = \boldsymbol{DHPA}\Psi\boldsymbol{s} + \boldsymbol{DH}\zeta.$$

Define the matrix $\Theta = \boldsymbol{DHPA}\Psi$ and the noise vector $\boldsymbol{e} = \boldsymbol{DH}\zeta$ to have $\boldsymbol{z}_{\downarrow} = \Theta\boldsymbol{s} + \boldsymbol{e}$, where $\Theta \in \mathbb{R}^{NxR}$ with $R < M$. This recasts the classical Compressive Sensing problem presented in section 5.1 of the previous chapter, and can be efficiently solved with a greedy algorithm or an *L1-norm* minimization to find the sparse vector $\boldsymbol{s}$.

6.3.1 Reconstruction Signal to Noise Ratio Estimation

In this subsection, we estimate the performance in terms of $RSNR$ achieved by the $RADS$ Converter. Since no other errors are modeled in the previous analysis, quantization noise limits the performance of the reconstruction algorithm and of the whole architecture.

It is possible calculate the total power of quantization noise from equation (6.6) considering the remaining bandwidth determined by the filter cutoff frequency. On top of that, we have payed a price when we decided to have a contribution of every possible component of the signal in the low pass portion of the band, i.e. we have spread the energy of every single component over the whole bandwidth of the original signal. Given that signal energy is conserved as it pass trough the random multiplication (we have multiplied by a ± 1 sequence), the

magnitude of the signal that remains at the end of the reconstruction algorithm will be inversely proportional to the original signal bandwidth.

Finally, we can estimate the $RSNR$ as

$$RSNR = 20log_{10}\left(\frac{\|\boldsymbol{x}\|_2}{Ne_{rms}\frac{\pi^L}{\sqrt{2L+1}}\left(\frac{M}{R}\right)^{-(L+\frac{1}{2})}}\right) \tag{6.8}$$

Note that it is possible to achieve a significant improvement with respect to classical Δ/Σ conversion by making R as small as possible, given that since the oversampling ratio in the above expression is calculated with respect to the filter bandwidth instead the signal bandwidth.

6.3.2 Numerical Experiments

In this subsection, we present the results of a set of numerical experiments designed to verify and validate the $RADS$ Converter architecture.

All the simulated points showed in the plots are the mean value over more than 5000 simulations where a new signal was generated with a random support in every trial.

The 1-bit encoding was made using third order Δ/Σ modulator designed with *delsig* [59] toolbox for Matlab [60].

For all the simulations of this section we have fixed a set of parameters that illustrates the most significant cases. The number of samples was always fixed to $M = 2048$ independently of the time scale used. We have considered an input signal that is K-sparse in the Fourier domain , i.e. it is constructed of up to K different periodic tones, $\boldsymbol{x} = \boldsymbol{Fs}$ where

$$\boldsymbol{F}_{j,k} = \text{real}\left\{e^{-\frac{2\pi}{N}(j-1)(k-1)}\right\} \qquad\qquad j,k = 1,\ldots,N$$

and the value of N is varied across the simulation.

In order to show deferents behaviors of the system, we have simulated a set of different sparsity values of K ($K = 4$, 12, 20 and 28). The power of the input signal was kept constant along all values of K, which implies a decrease in the value of every single component as the sparsity value increases.

The Δ/Σ modulator was chosen to be a third order modulator, which is a typical configuration for this kind of converters and due to the fact that in higher order modulators instabilities are more frequent to happen in the loop filter [55].

The CS reconstruction algorithm at the end of the chain is `CoSaMP`, and the number of iteration is fixed to 200.

In the first set of simulation we have estimated the $RSNR$ and the PSR as a function of the cutoff frequency of the reconstruction filter. The results are plot in Figure 6.6.

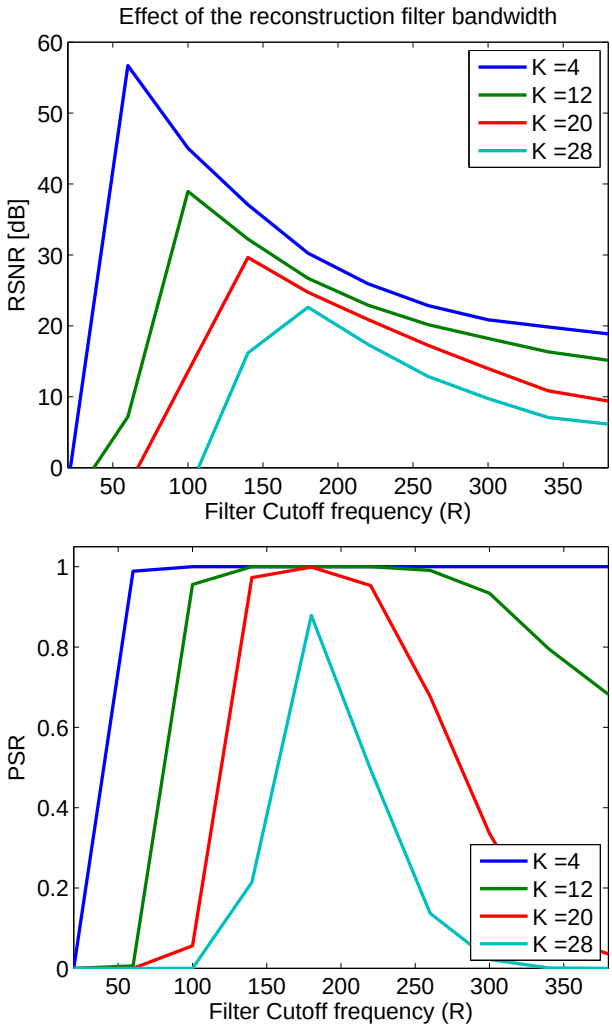

Fig. 6.6: Effect of the reconstruction filter bandwidth for a $RADS$ Converter with a third order Δ/Σ modulator for different sparsity levels. On top: $RSNR$ as a function of filter bandwidth R; bottom: PSR as a function of R. For every combination of K and N there is an optimal value for R.

As it is shown in the plots, for small values of R, it is not possible to reconstruct the original signal. This is due to the fact that only a small amount of information is left after filtering, and it is not possible to distinguish which are the

components present in the signal. In different words, there is a large probability that more than one signal could be a good candidate solution with this reduced set of measurements.

As we increase the value of R, the probability of reconstruction jumps from almost 0 to almost 1, for small values of K. This behavior is consistent for different values of R at every K we have simulated. At this point, there is enough information to distinguish which are the components present in the original signal from the low pass portion of the mixed signal. The reconstruction error is limited by the quantization noise that is left in this portion.

As we continue increasing the filter bandwidth, there is a decreased in the performance in terms of $RSNR$, as well as PSR. The deterioration in the $RSNR$ can be easily explained due to the fact that a larger bandwidth produces an increment in the power of the quantization noise (see eq. (6.8)) , i.e., the residual noise energy is large compared to the signal energy.

On the other hand, for large values of R, as we increase K there is a decrease in performance in terms of PSR. As we have the same amount of quantization noise power for a fixed R, increasing the value of K reduces the power of every single component present in the input signal, making it harder for the reconstruction algorithm to identify those components in a noisy environment (the magnitude of the noise is comparable with the magnitude of the signal). This fact illustrates the existing trade-off in the selection of the R parameter. As we increase the value of R in order to obtain a better performance in terms of PSR, there is a detriment in terms of $RSNR$. The optimal value of R will be the smallest value that produces a PSR near to one, and this value is a function of K as well as of N.

We have extensively studied the optimal value of R trough an empirical approach using numerical simulation. We have found that to obtain a $PSR \geq 0.99$ then

$$R \geq 1.4Klog\left(\frac{N}{K} + 6\right) + 30 \tag{6.9}$$

. This equation is in accordance with equation (5.1) presented in section 5.2 of the previous chapter. However, in the following section we will see how a further exploitation of the $RADS$ architecture will lead to an improvement in both figures of merits.

In Figure 6.7, we show the simulation results for the optimal value of R given by equation (6.9) as we vary the value of N. We have also added the curves given by equation (6.8) in order to compare the simulation result with the predicted theoretical $RSNR$.

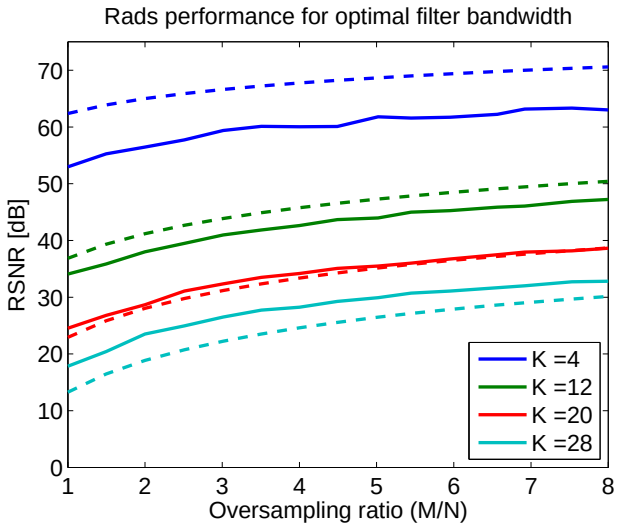

Fig. 6.7: Simulation of the performance of $RADS$ Converter using the optimal reconstruction filter bandwidth for different sparsity levels. $RSNR$ as a function of the oversampling ratio M/N. The support was correctly recovered 100% of the time. The solid line represents the simulation result, while the dashed line the theoretical result from eq. (6.8).

The estimated $RSNR$ follows the behavior of the simulated system, in terms of variation of parameters K and N (the same occurs with M, not shown). The differences are due to the linear model used in the approximation of the in-band noise of the Δ/Σ converter (which a very non-linear system), and the non-ideal

behavior of the filters used for the reconstruction. Clearly, the expression in equation (6.8) can be used as a design guideline.

Note also that by taking just 1-bit measurements at Nyquist rate we can get resolutions of up to $52dB$, obtaining a compression rate of about 8 times with respect to Nyquist sampling for the same resolution.

Finally, in Figure 6.8 we have simulated the same setting as before. In this case, we have added some intrinsic noise to the original signal (*i.i.d.* additive Gaussian noise) to get an input SNR of 30 dB.

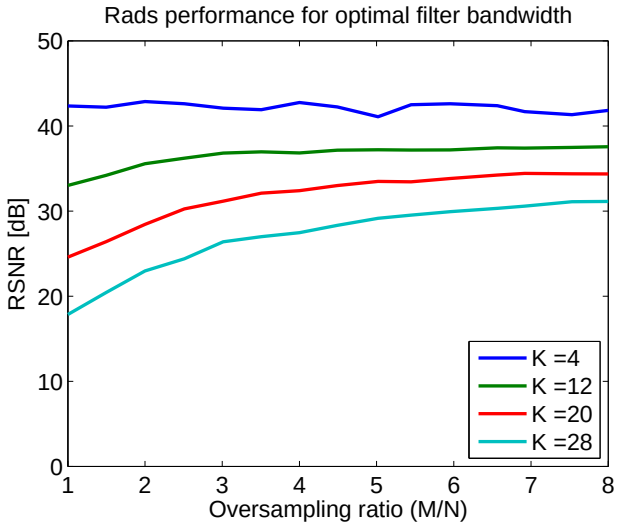

Fig. 6.8: Simulation of the performance of the $RADS$ Converter using the optimal reconstruction filter bandwidth for different sparsity levels. The input signal has an intrinsic SNR of 30 dB. $RSNR$ as a function of the oversampling ratio M/N. The support was correctly recovered 100% of the time.

As we can see in the plots, the performance of the converter is limited by the intrinsic noise present in the signal, even if there is a sort of denoising for the smallest values of K.

In this section we have shown how the $RADS$ Converter can be employed

to acquire analog-sparse signals with a total number of bits much smaller than Nyquist multi-resolution analog-to-digital converters. Simulation results have shown that the proposed architecture collects the necessary information to successfully reconstruct sparse signals.

In the next section we will see how we can exploit the peculiarities of the acquisition strategy to produce an improved estimate of the signal in terms of accuracy and probability of successful reconstruction over different sparsity conditions. This further exploitation will derive in an algorithm that we have called `FCoSaMP` which is free of the parameters that compromise both figures of merit.

6.4 FCOSAMP

In general, signal reconstruction for a CS acquisition scheme can be split into two parts: support recovery, i.e. the identification of the location of the nonzero components, and amplitude estimation over that support. Consider first the situation in which the support is already known. If the columns of the measurement matrix Θ indexed by the location of the nonzero components form a full-rank matrix, the natural approach is to reconstruct the signal by least squares, and the approximation error will be only limited by the power magnitude of the noise introduced by the measurement process.

On the other hand, in the general case where the support is not known, most algorithms can be ensured to work based on the RIP of the measurement matrix as stated in the previous section. For some matrix construction with entries that are Gaussian or sub-Gaussian, the RIP is satisfied with overwhelming probability if the number of measurements is bigger than a multiple of the signal sparsity $M \geq CK \log(N/K)$. If the number of measurements falls below a certain minimum number, the probability of successful reconstruction change from a very high to a very poor one (see e.g. [61]). This phenomena, in terms of compressive sensing, is the so called $phase\ transition$ effect.

As we have seen before, thanks to spreading, every non-zero entry in $\boldsymbol{s}$ implies a waveform whose energy can be detected at practically any frequency including those where the quantization error is reduced by the Δ/Σ. Hence, to remove the quantization noise it is desirable to take only a small bandwidth around zero where the high-pass nature of the disturbance has only a small contribution. On the other hand, the considered signal should contain enough information for the recovery of the original sparse signal.

This trade-off fits particularly well into algorithms iterating an elementary step that estimates $\mathrm{supp}\,(\boldsymbol{s})$ and then calculates the corresponding non-null entries.

In these algorithms, it is possible to low-pass filter (and decimate to remove redundant samples) the input vector at each iteration. Doing so, as the recon-

struction proceeds, its refinement happens with values that are progressively less affected by disturbances since, while signal energy decreases linearly as the band shrinks around DC, disturbance energy decreases polynomially thanks to the NTF.

The nature of the architecture allows to develop an iterative algorithm that recovers the support with a very high probability, since we start the estimation process with a big number of measurements, and reduces the quantization noise to the minimum possible depending on the sparsity level.

The algorithm we propose to exploit this intuition is reported in Algorithm 1 and will be referred as `FCoSaMP` in the following.

We can model the filter process as the application of an l-order FIR filter ($l \leq m$) and arrange its impulse response coefficients $h_1, h_2, \ldots, h_l$ as the rows of a matrix $\boldsymbol{H}^{(m)}$ of $m \times m$ elements, where m is the length of the sequence to be filtered.

$$\boldsymbol{H}^{(m)}_{j,k} = \begin{cases} h_i & \text{if } j = k + i - 1 \\ 0 & \text{if } j \neq k + i - 1 \end{cases} \qquad \begin{matrix} j, k = 1, \ldots, m \\ i = 1, \ldots, l \end{matrix}$$

Depending on the filter cutoff frequency, the filtered sequence can be decimated by a factor d. We can model this operator in the matrix $\boldsymbol{D}^{(d,m)}$ of $\left\lfloor \frac{m}{d} \right\rfloor \times m$ elements

$$\boldsymbol{D}^{(d,m)}_{j,k} = \begin{cases} 1 & \text{if } j = \left\lfloor \frac{k}{d} \right\rfloor \\ 0 & \text{if } j \neq \left\lfloor \frac{k}{d} \right\rfloor \end{cases} \qquad \begin{matrix} j = 1, \ldots, \left\lfloor \frac{m}{d} \right\rfloor \\ k = 1, \ldots, m \end{matrix}$$

By writing the number of measurements as $M = 2Kd_0d_1 \ldots d_{J-1}$ with d_j being a small downsampling factor (typically 2 or 3) and J being the total number of downsampling steps of the algorithm, at the j-th iteration the outer loop filters the signal and downsample it by a factor d_j to reduce quantization noise. In our

implementation, low-pass filtering was obtained by sinc frequency profiles with lobes matched with the subsampling ratio.

Downsampling continues until the number of available samples is $2K$ since this is the minimum information needed to discriminate between two different K-sparse vectors.

The inner loop is performed a fixed number of times and is based on `CoSaMP` to iteratively produce an improved estimation of $\boldsymbol{s}$ by least squares over a reduced support made of the support of the previous iteration plus the support of the largest components of the residuals of the previous iteration.

Algorithm 1 Reconstruct $\boldsymbol{x}$ from 1-bit vector $\boldsymbol{z}$

Θ^*	Complex conjugate transpose of Θ.
$\Theta^\dagger$	Pseudoinverse of Θ. $\Theta^\dagger = (\Theta^*\Theta)^{-1}\Theta^*$.
$\boldsymbol{w}_{\|K}$	Set to zero all $\boldsymbol{w}$ but the K biggest component.
$\text{supp}(\boldsymbol{w})$	Indexes of the nonzero components of $\boldsymbol{w}$.
d_j	Downsampling ratios such that $M = 2Kd_0d_1 \dots d_{J-1}$

Require: Sampling matrix Θ, 1-bit vector $\boldsymbol{z}$, sparsity level K.
  $m \leftarrow M$
  $\boldsymbol{s} \Leftarrow (0, \dots, 0)^T$
  for $j = 1, \dots, J-1$ **do**
    $\boldsymbol{z} \Leftarrow \boldsymbol{D}^{(d_j,m)}\boldsymbol{H}^{(m)}\boldsymbol{z}$
    $\Theta \Leftarrow \boldsymbol{D}^{(d_j,m)}\boldsymbol{H}^{(m)}\Theta$
    $\boldsymbol{v} \Leftarrow \boldsymbol{z} - \Theta\boldsymbol{s}$
    $m \leftarrow \lfloor m/d_j \rfloor$
    for $i = 1, \dots, I$ **do**
      $\boldsymbol{w} \Leftarrow \Theta^*\boldsymbol{v}$
      $T \Leftarrow \left\{\text{supp}(\boldsymbol{w}_{|K}) \cup \text{supp}(\boldsymbol{s}_{|K})\right\}$
      $\boldsymbol{b}(T) \Leftarrow \Theta(\cdot, T)^\dagger\boldsymbol{s}$
      $\boldsymbol{b}(\{1, \dots, N\}\backslash T) \Leftarrow (0, \dots, 0)^T$
      $\boldsymbol{s} \Leftarrow \boldsymbol{b}_{|K}$
      $\boldsymbol{v} \Leftarrow \boldsymbol{z} - \Theta\boldsymbol{s}$
    end for
    $T \Leftarrow \text{supp}(\boldsymbol{s}_{|K})$
    $\boldsymbol{b}(T) \Leftarrow \Theta(\cdot, T)^\dagger\boldsymbol{s}$
    $\boldsymbol{b}(\{1, \dots, N\}\backslash T) \Leftarrow (0, \dots, 0)^T$
    $\boldsymbol{s} \Leftarrow \boldsymbol{b}_{|K}$
  end for
  $\hat{\boldsymbol{x}} \Leftarrow \Psi\boldsymbol{s}$

Intuitively, the high probability of correct support recovery comes from the fact that we estimate it under large noise condition, but large number of measurements. Once the support is identified (at every iteration the support estimation is improved), the bandwidth is decreased in order to reduce the quantization noise. The key fact is to note that the signal energy decreases linearly when the frequency decreases, while noise energy decreases polynomially thanks to the Δ/Σ

noise shaping properties [55]. This combination of filtering and estimation, has the benefit of recovering the signal with a very high probability of success, while reducing the quantization noise to the minimum.

6.4.1 Numerical Experiments

The simulations run in this section share the same set of parameters and configuration as the simulations draw in Section 6.3.2.

In the first experiment (Figure 6.9), we have run the same encoding as that in Figure 6.7, and we have made the decoding with `FCoSaMP`. Note the increment in terms of $RSNR$ of at least 30 dB in all the cases. It is also important to note the large $RSNR$ that is achieved especially for very sparse signals. Note that for M/N = 1 we are just taking 1-bit measurement at Nyquist rate and obtaining a $RSNR$ of up to $90dB$, which translates into a compression factor of about 15 times with respect to Nyquist sampling.

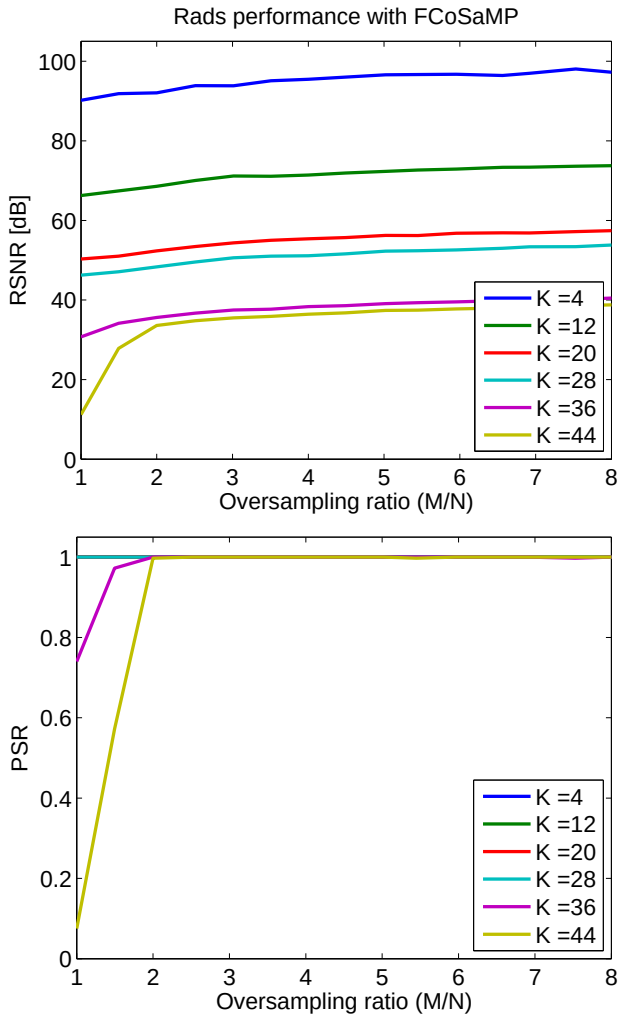

Fig. 6.9: Simulation of the performance of $RADS$ Converter using the `FCoSaMP` algorithm in the reconstruction for different sparsity levels. On top: $RSNR$ as a function of the oversampling ratio M/N; bottom: PSR as a function of M/N. The large $RSNR$ that is achieved translates into a compression factor of up to 15 times with respect to Nyquist based acquisition

Another interesting fact is that the support recovery was always correct for $K < 36$, while it was substantially less than 100% only for $K \geq 36$ when low oversampling ratios are considered. This is mainly due to the large bandwidth that

remains at the end of the algorithm, i.e., to the residual noise energy that is large with respect to the signal energy therefore large noise energy compared with the signal energy.

In the second experiment, we have added intrinsic noise to the signal by adding *i.i.d.* Gaussian noise to get an input SNR of 30 dB. The results are shown in Figure 6.10.

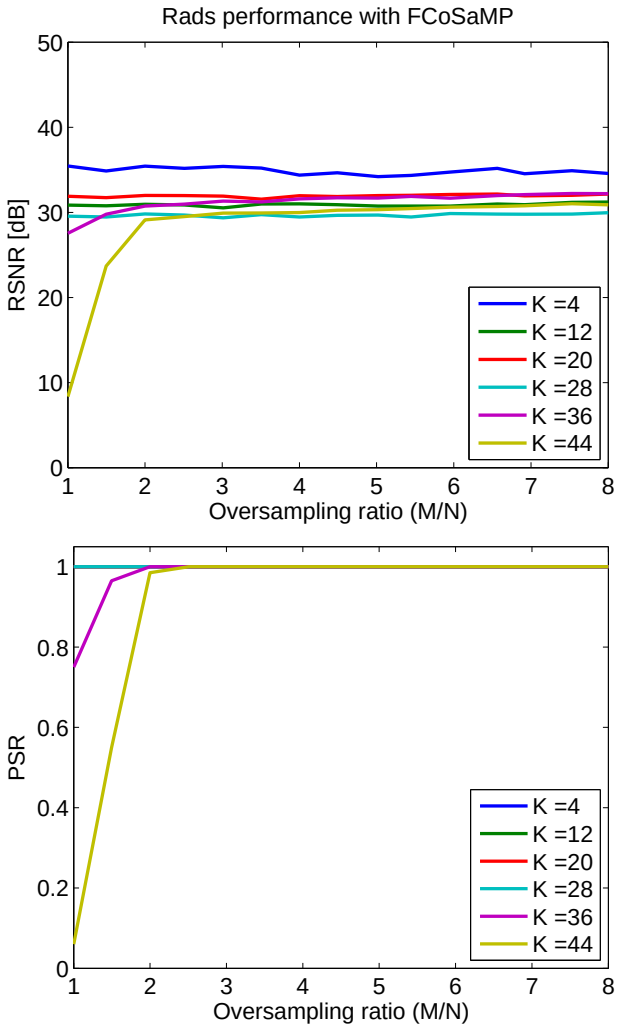

Fig. 6.10: Simulation of the performance of $RADS$ Converter using the `FCoSaMP` algorithm for reconstruction, for different sparsity levels. The input signal has an intrinsic SNR of 30 dB. On top: $RSNR$ as a function of the oversampling ratio M/N; bottom: PSR as a function of M/N. The encoding process and the reconstruction algorithm show to be robust against strong noise condition.

As in the experiment presented in previous section, the performance of the converter is limited by the intrinsic noise of 30 dB. However the denoising effect is less evident in this simulation. This is due to the fact that as we filter and

decimate iteratively, given the not ideal behavior of the filters, part of the noise is aliased into the low part of the band and added to the total noise energy at the end of the algorithm.

In spite of this, it is shown that the encoding process of the $RADS$ Converter, as well as the behavior of reconstruction algorithm are robust against strong noise condition showing a behavior in terms of PSR similar to that of the previous simulation.

To avoid possible biases due to the choice of a particular sparsity basis, in the third experiment we have changed the sparsity basis and we have simulated the acquisition of a signal that is sparse along a random basis obtained by orthonormalizing a matrix with Gaussian independent entries with zero average. The results are shown in Figure 6.11. Comparing this results with those in Figure 6.9, note that there is a slight difference in terms of $RSNR$. This is due to the fact that the former may contain some components that when looked in the time domain concentrate most of its energy in small time intervals. In other words, the signal energy is not uniformly distributed along the time axis, making many of the samples taken by the $RADS$ Converter useless or without information.

In the extreme case, when all the energy is concentrated in a small period of time compared with the time-window used for the processing, $RADS$ Converter will fail to decode this kind of signals.

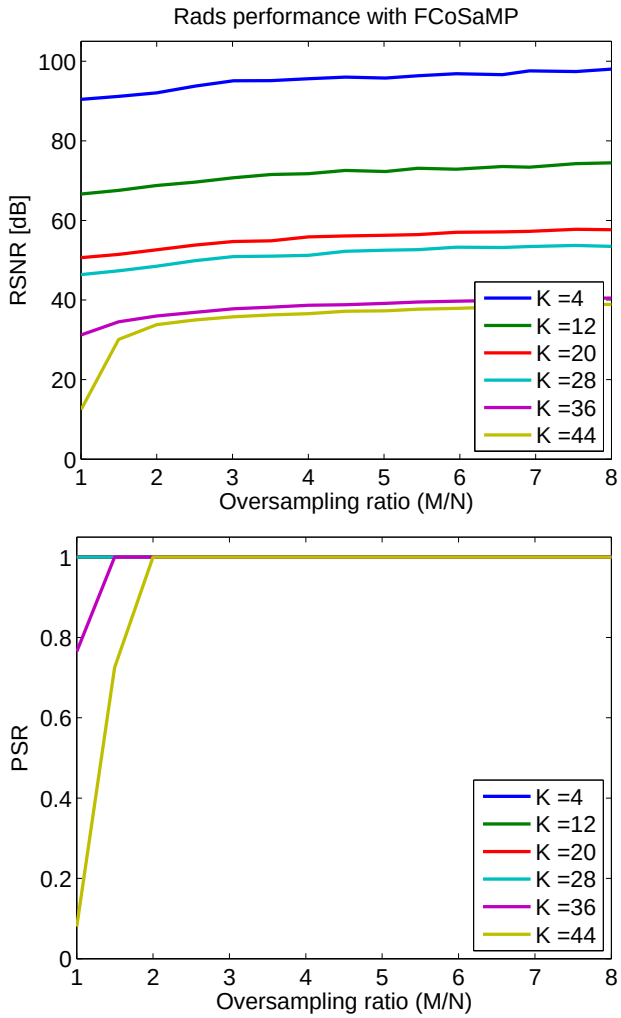

Fig. 6.11: Simulation of the performance of $RADS$ Converter by using the `FCoSaMP` algorithm for reconstruction, for different sparsity levels. The input signal is sparse in a random basis. On top: $RSNR$ as a function of the oversampling ratio M/N; bottom: PSR as a function of M/N. The proposed architecture shows to work independently of the sparsity basis provided it is spread on the time axis.

Finally, in the last experiment we have compared the performance achieved by our system with two state of the art 1-bit compressive sensing algorithms, i.e.,

the *Restricted Step Shrinkage* (RSS) [62] and *Binary Iterative Hard Thresholding* (BIHT) [51], that are generic schemes working on measurement matrices with nice theoretical properties (Figure 6.12).

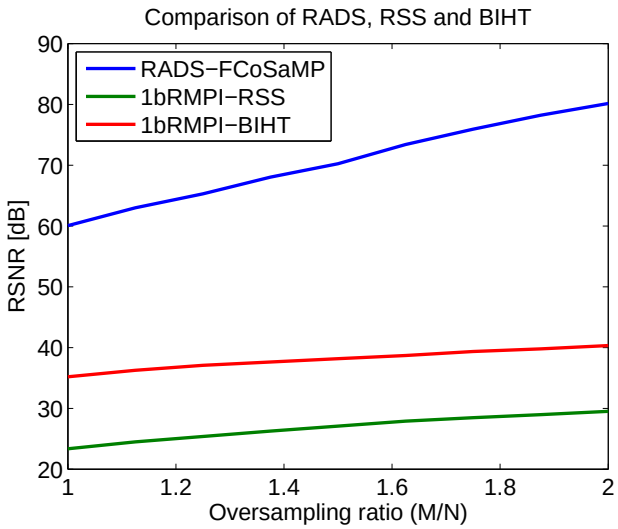

Fig. 6.12: Comparison of the performance of $RADS$ Converter decoded with FCoSaMP with $1bRMPI$ decoded with the RSS and BIHT algorithms. $RSNR$ as a function of the oversampling ratio M/N for fix signal length $N = 1024$ and sparsity level of $K = 10$. The same amount of 1-bit samples are considered for every case.

In every trial we have simulated the acquisition of a signal that is 10-sparse along a random basis. Independently of the architecture, the same amount of 1-bit samples z are considered as input for the reconstruction algorithm.

In all cases, the $RADS$ scheme was able to perfectly reconstruct the support of the original signal and, as shown in Figure 6.12, it achieves an $RSNR$ largely superior to that of the references.

6.5 Time Domain Analysis

Up to now, we have made a high level analysis based on frequency domain assumptions of the functionality of the $RADS$ converter, and we have evaluated two different reconstruction algorithms that produce different results in terms of the selected performance metrics.

It seems that in the studied cases, the achieved performance is limited by the selected reconstruction algorithm, and not by the acquisition architecture itself.

The main question we want to answer in this section is: what is the maximum achievable performance of the $RADS$ converter (independently of the reconstruction algorithm)?

To answer this question we cannot perform only the high level analysis we have made so far, but we need a deeper understanding of the encoding process. For this purpose, we will make a time-domain analysis of the converter, starting by a time-domain analysis of single 1^{st}-order Δ/Σ modulator, then generalizing it to a L^{st}-order Δ/Σ modulator, and finally analyzing the whole $RADS$ Converter architecture. In addition, we will also show how to exploit the time-domain analysis made for the $RADS$ converter in order to reconstruct the original signal from the one bit measurements.

6.5.1 Δ/Σ modulator time-domain analysis

1^{st}-order Δ/Σ modulator time-domain analysis

Consider first a discrete time 1^{st}-order Δ/Σ modulator as in figure 6.13 with zero initial conditions, where the discrete sequence y feeds the modulator, that outputs the discrete sequence z, and where the internal state is defined by the state variable v at any time n. The block called [Q] represents a general quantizer and the block called [D] represents a one time-step delay.

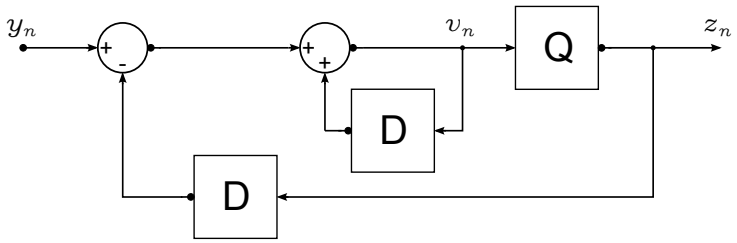

Fig. 6.13: First order Δ/Σ modulator schematic diagram.

Following the signal path we can write the following equation at any time n:

$$v_n = y_n - z_{n-1} + v_{n-1} \tag{6.10}$$

and of course

$$v_{n-1} = y_{n-1} - z_{n-2} + v_{n-2} \tag{6.11}$$

Replacing (6.11) into (6.10) we get

$$v_n = y_n - z_{n-1} + y_{n-1} - z_{n-2} + v_{n-2}$$

Extending the same reasoning up to $n = 1$ and for $v_1 = 0$ (zero initial conditions) we can write

$$v_n = y_n - z_{n-1} + y_{n-1} - z_{n-2} + y_{n-2} - z_{n-3} + v_{n-3}$$

$$v_n = y_n + \cdots + y_1 - z_{n-1} - \cdots - z_1 = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n-1} z_i \tag{6.12}$$

This equation relates the current state variable v at any time n with the whole history of input y and output z up to time $n - 1$.

On the other hand, since we are using a 1-bit quantizer [Q], without loss of generality by encoding the value $'1'$ for all positive inputs of the quantizer (including zero), and the value $'-1'$ for all negative values of the input, we have that

$$v_n z_n \geq 0 \qquad \forall n \tag{6.13}$$

Defining the vectors

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_n \end{bmatrix}; \boldsymbol{v} = \begin{bmatrix} v_1 \\ . \\ . \\ . \\ v_n \end{bmatrix}; \boldsymbol{z} = \begin{bmatrix} z_1 \\ . \\ . \\ . \\ z_n \end{bmatrix}$$

and the matrices

$$\boldsymbol{Z} = \begin{bmatrix} z_1 & 0 & 0 & . & . & . & 0 \\ 0 & z_2 & 0 & . & . & . & 0 \\ 0 & 0 & z_3 & . & . & . & 0 \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ 0 & 0 & 0 & . & . & . & z_n \end{bmatrix};$$

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 1 & 1 & 0 & . & . & . & 0 \\ 1 & 1 & 1 & . & . & . & 0 \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ 1 & 1 & 1 & . & . & . & 1 \end{bmatrix}; \Delta = \begin{bmatrix} 0 & 0 & 0 & . & . & . & 0 \\ 1 & 0 & 0 & . & . & . & 0 \\ 1 & 1 & 0 & . & . & . & 0 \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ 1 & 1 & 1 & . & . & . & 0 \end{bmatrix}$$

we can write

$$\boldsymbol{v} = \Sigma \boldsymbol{y} - \boldsymbol{\Delta z} \tag{6.14}$$

and

$$\boldsymbol{Z v} \geq 0 \tag{6.15}$$

where the last inequality is component-wise.

In this way, combining equation (6.14) with equation (6.15), and given the measurements $\boldsymbol{z}$ we can define a solution space for any input $\boldsymbol{y}$:

$$\boldsymbol{Z\Sigma y} \geq \boldsymbol{Z\Delta z}$$

This space contains all possible instances of the input $\boldsymbol{y}$ that are solutions of the Δ/Σ modulation process.

L^{th}-order Δ/Σ modulator time-domain analysis

Consider now a discrete time L^{th}-order Δ/Σ modulator as in figure 6.14 with zero initial conditions.

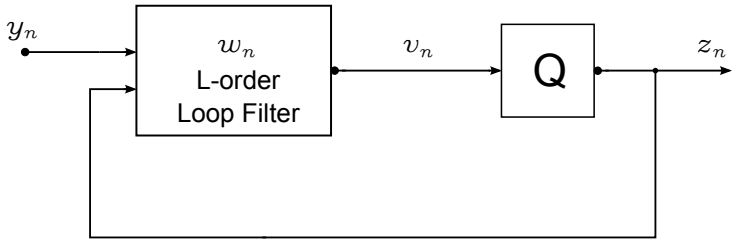

Fig. 6.14: L^{th}-order Δ/Σ modulator schematic diagram.

In this case, the vector $\boldsymbol{w} \in \mathbb{R}^L$ defines the state vector of the loop filter, while

the state variable v is equal to the last component of the state vector ($w_{:,L}$).

The current state at time n can be computed based on the previous states as

$$\boldsymbol{w}_n = \boldsymbol{B}y_{n-1} + \boldsymbol{C}z_{n-1} + \boldsymbol{A}\boldsymbol{w}_{n-1}$$

where the matrix $\boldsymbol{A} \in \mathbb{R}^{LxL}$ and the vectors $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{L}$ contain the coefficients that determine the transfer function of the loop filter.

The output of the loop filter is simply $v_n = w_{n,L}$ and the output of the modulator is calculated as $z_n = sign(v_n)$, where $w_{:,L}$ is the L^{th} component of the state vector $\boldsymbol{w}$.

Analogously as proceeded with the 1^{st}-order modulator, we can write

$$\boldsymbol{w}_n = \boldsymbol{B}y_{n-1} + \boldsymbol{C}z_{n-1} + \boldsymbol{A}\boldsymbol{B}y_{n-2} + \boldsymbol{A}\boldsymbol{C}z_{n-2} + \boldsymbol{A}^{(2)}\boldsymbol{w}_{n-2}$$

$$\boldsymbol{w}_n = \boldsymbol{A}^{(0)}\boldsymbol{B}y_{n-1} + \cdots + \boldsymbol{A}^{(n-2)}\boldsymbol{B}y_1 + \boldsymbol{A}^{(0)}\boldsymbol{C}z_{n-2} + \cdots + \boldsymbol{A}^{(n-2)}\boldsymbol{C}z_1$$

$$\boldsymbol{w}_n = \boldsymbol{B}\sum_{i=1}^{n-1}\boldsymbol{A}^{(i-1)}y_i + \boldsymbol{C}\sum_{i=1}^{n-1}\boldsymbol{A}^{(i-1)}z_i \tag{6.16}$$

since $\boldsymbol{w}_1 = \boldsymbol{0}$ (zero initial conditions).

We also have that,

$$v_n z_n \geq 0 \qquad \forall n$$

Defining the vectors

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_{n-1} \end{bmatrix}; \boldsymbol{v}' = \begin{bmatrix} w_{2,1} \\ w_{2,2} \\ . \\ . \\ w_{2,L} \\ \boldsymbol{w}_3 \\ . \\ . \\ . \\ \boldsymbol{w}_{n-1} \end{bmatrix}; \boldsymbol{z} = \begin{bmatrix} z_1 \\ . \\ . \\ . \\ z_{n-1} \end{bmatrix}$$

and the matrices

$$\boldsymbol{Z} = \begin{bmatrix} z_1 & 0 & 0 & . & . & . & 0 \\ 0 & z_2 & 0 & . & . & . & 0 \\ 0 & 0 & z_3 & . & . & . & 0 \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ 0 & 0 & 0 & . & . & . & z_{n-1} \end{bmatrix};$$

Define

$$\boldsymbol{\sigma}_i = \boldsymbol{A}^{(i)}\boldsymbol{B}$$

$$\boldsymbol{\delta}_i = \boldsymbol{A}^{(i)}\boldsymbol{C}$$

then,

$$\boldsymbol{\Sigma}' = \begin{bmatrix} \boldsymbol{\sigma}_0 & 0 & 0 & . & . & . & 0 \\ \boldsymbol{\sigma}_1 & \boldsymbol{\sigma}_0 & 0 & . & . & . & 0 \\ \boldsymbol{\sigma}_2 & \boldsymbol{\sigma}_1 & \boldsymbol{\sigma}_0 & . & . & . & 0 \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ \boldsymbol{\sigma}_{n-2} & \boldsymbol{\sigma}_{n-3} & \boldsymbol{\sigma}_{n-4} & . & . & . & \boldsymbol{\sigma}_0 \end{bmatrix};$$

$$\boldsymbol{\Delta}' = - \begin{bmatrix} \boldsymbol{\delta}_0 & 0 & 0 & . & . & . & 0 \\ \boldsymbol{\delta}_1 & \boldsymbol{\delta}_0 & 0 & . & . & . & 0 \\ \boldsymbol{\delta}_2 & \boldsymbol{\delta}_1 & \boldsymbol{\delta}_0 & . & . & . & 0 \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ \boldsymbol{\delta}_{n-2} & \boldsymbol{\delta}_{n-3} & \boldsymbol{\delta}_{n-4} & . & . & . & \boldsymbol{\delta}_0 \end{bmatrix}$$

We can now write equation (6.16) in matrix form as

$$\boldsymbol{v}' = \Sigma' \boldsymbol{y} - \boldsymbol{\Delta}' \boldsymbol{z} \tag{6.17}$$

In order to keep only the last component of the state vector, define the matrix $\boldsymbol{K} \in \mathbb{R}^{L(N-1)xN-1}$ as

$$\boldsymbol{K} = \begin{bmatrix} 0 \ldots 0, 1 & \boldsymbol{0} & \boldsymbol{0} & . & . & . & \boldsymbol{0} \\ \boldsymbol{0} & 0 \ldots 0, 1 & 0 & . & . & . & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & 0 \ldots 0, 1 & . & . & . & \boldsymbol{0} \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & . & . & . & 0 \ldots 0, 1 \end{bmatrix};$$

where $\boldsymbol{0}$ represents a row vector of size L with all zeros in its inputs.

Defining $\boldsymbol{v}$ as

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ . \\ . \\ . \\ v_{n-1} \end{bmatrix}$$

to get

$$\boldsymbol{v} = \boldsymbol{K\Sigma' y} - \boldsymbol{K\Delta' z}$$

, and

$$\boldsymbol{v} = \boldsymbol{\Sigma y} - \boldsymbol{\Delta z} \tag{6.18}$$

where $\boldsymbol{\Sigma} = \boldsymbol{K\Sigma'}$ and $\boldsymbol{\Delta} = \boldsymbol{K\Delta'}$.

As in the 1^{st}-order case we have,

$$\boldsymbol{Zv} \geq 0 \tag{6.19}$$

and combining equation (6.18) with (6.19) to have

$$\boldsymbol{Z\Sigma y} \geq \boldsymbol{Z\Delta z} \tag{6.20}$$

The space defined by equation (6.20) contains all possible instances of the input $\boldsymbol{y}$ that are solution of the Δ/Σ modulation process, given the measurements $\boldsymbol{z}$.

Now, we can highlight some observations for the above equation:

- The input signal defines a point in the multi-dimensional space that is contained in the solution space defined by the set of equations in (6.20).
- Fixing the dimension of the input signal, as we add new measurements, every measurement will split the space into two sub-spaces. Only one of those sub-spaces will contain possible solutions.
- The minimum number of measurements needed to define a closed region is equal to the dimension of the input signal plus one.
- Adding a new measurement, does not imply a reduction in the solution space.
- A smaller solution space implies an estimation of the input signal with a bigger accuracy. In other words, the smaller the solution space, the bigger the SNR of the estimated signal.

6.5.2 $RADS$ Converter time-domain analysis

We now have all we need to analyze the whole $RADS$ converter architecture. The set of inequalities $\boldsymbol{Z}\Sigma\boldsymbol{y} \geq \boldsymbol{Z}\Delta\boldsymbol{z}$ define the solution space given by the Δ/Σ modulator. In the same way as above, to completely model the $RADS$ Converter and the input signal itself, we can easily write

$$\boldsymbol{y} = \boldsymbol{DAx} = \boldsymbol{DA}\Psi\boldsymbol{s}$$

to have

$$\boldsymbol{Z}\Sigma\boldsymbol{DA}\Psi\boldsymbol{s} \geq \boldsymbol{Z}\Delta\boldsymbol{z} \tag{6.21}$$

where the matrix $\boldsymbol{A}$ is the upsampling operator and the matrix $\boldsymbol{D}$ represents the pre-modulation process, as we have defined before.

We have now a complete description of the modulation process of the $RADS$ Converter architecture in time domain. The main differences between the $RADS$ Converter and a Δ/Σ alone are two: first, the pre-modulation increases the probability that every new measurement modify the solution space, increasing in this way the accuracy in the estimation. Secondly, under the assumption that $\boldsymbol{s}$ is sparse in a given domain, it is possible reduce the solution space to only those candidates that satisfy this condition, reducing even more the solution space. Since sparsity is not a dimensionality reduction, it is not possible to know a priori which are the directions to look at, but as we proceed with the measurements, there will be many candidates to discard since they are not sparse enough to be a possible solution.

6.5.3 Space Dimension Analysis

As we have seen before, it is possible to define a monotonic dependence between the size of the solution space and the SNR of the estimation of the input signal.

As a measure of size, and considering that any point in the solution space is a candidate with the same probability, it is possible to consider the hyper-volume of that solution space as a measure of precision of the estimation of the input signal.

Regrettably, an analytical expression for the calculation of the hyper-volume in high-dimensional spaces is a difficult task, and we need to resort to numerical integration. For that purpose, we will use Monte-Carlo integration [63, 64] in order to have an estimation of the hyper-volume of the solution space as a function of the number of measurements.

Monte Carlo integration is a technique for numerical integration that uses random numbers, and is particularly useful for higher dimensional integrals. Informally, to estimate the volume of a given domain D, we have first, to pick a simple domain E whose volume is easily calculated and which D is contained. Then, we generate a sequence of random points that fall within E, some of which will also fall within D. Finally, we calculate the area of D as the area of E by the fraction of points that fall between E.

In this case, we have set the container volume as an hyper-cube of $1 \times 1 \times 1 \times \cdots \times 1$ and we have generated $500,000 \times M$ random points with uniform distribution within this range for every point in the plot.

We have plot in Figure 6.15 the hyper-volume of the solution space as a function of the number of measurements for two cases: using a single Δ/Σ converter (equation (6.20)); and using the $RADS$ Converter architecture (equation (6.21)). As a reference, we have also plot a line with slope $-1/2^M$. This line will occur only when every cut of the space produced by a new measurement divides the solution space exactly into two equal parts.

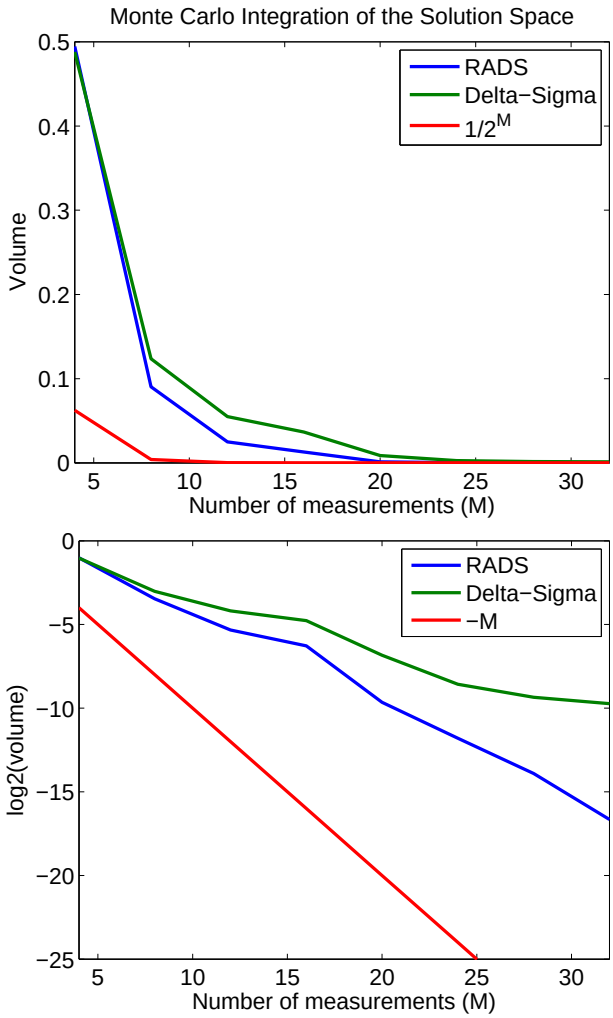

Fig. 6.15: Monte Carlo integration of the hyper-volume of the solution space for $RADS$ Converter and for Δ/Σ converter. On top: volume in linear scale as a function of the number of measurements M; bottom: volume in logarithmic scale as a function of M. The encoding performed by $RADS$ is more effective in reducing the size of the solution space.

As can be observed in Figure 6.15 the difference in size of the solution space obtained using the $RADS$ Converter approach is orders of magnitude more convenient than using classical Δ/Σ modulation. This difference is much more evident

as the number of measurements is increased. However, as the plot shows it is still possible to obtain an important improvement, since the line indicating the optimal cuts is far away from the one described by the $RADS$ Converter.

6.5.4 *L1-norm Minimization*

It was demonstrated above that the set of inequalities given by equation (6.21) define the solution space of the $RADS$ modulation process. This space still contains many possible input vectors $\boldsymbol{s}$, but we are particularly interested in the sparsest vector that exist in this space. In order to find such a vector we can recast to a *L1-norm* minimization, since from the observed in the previous chapter, it enforces sparsity across all possible solutions.

We can write the following minimization problem

$$\hat{s} = argmin \sum_{i=1}^{N} |s_i| \quad \text{s.t.} \quad \boldsymbol{Z\Sigma DA\Psi s} \geq \boldsymbol{Z\Delta z} \tag{6.22}$$

which from now on we will call $L1min$.

6.5.5 *Numerical Experiments*

In this section we will show the results from a series of simulation we have run in order to evaluate the minimization problem presented in equation (6.22).

We have setup the same conditions for the simulation in section 6.3.2, except that in this case we have reduced the number of measurements from 2048 to 1024. It was necessary to reduce this number for the simulation to be computational feasible, since every measurement produce a new constrain to be pass to the solver. The minimization problem was solved by the software `cplex` [65].

Figure 6.16 shows the performance achieved by `FCoSaMP` compared with that obtained using the minimization problem of equation (6.22), by fixing the sparsity number to $K = 8$.

As it is observed in the plot, $L1min$ outperforms `FCoSaMP` by around 10 dB in the whole range. This behavior can be verified using different experimental setups not shown.

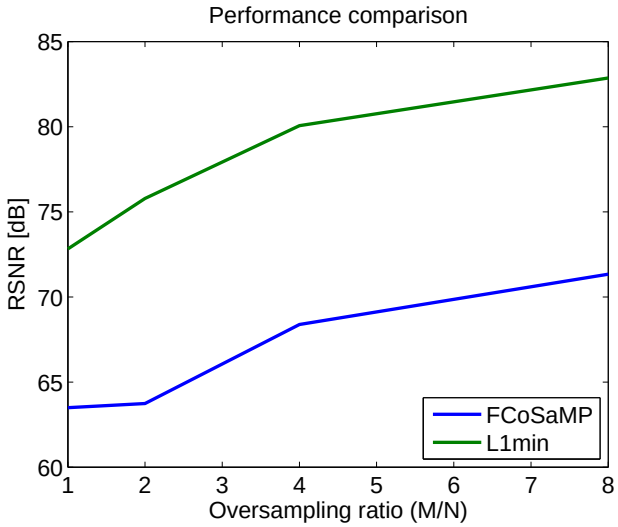

Fig. 6.16: Performance comparison of `FCoSaMP` and $L1min$. $RSNR$ as a function of oversampling ratio M/N for an 8-sparse signal encoded with $RADS$ converter.

The main drawback of this reconstruction algorithm is the running time needed to solve the minimization problem. In Figure 6.17 we have plotted the relationship between the average simulation time taken by $L1min$ over the time taken by `FCoSaMP`.

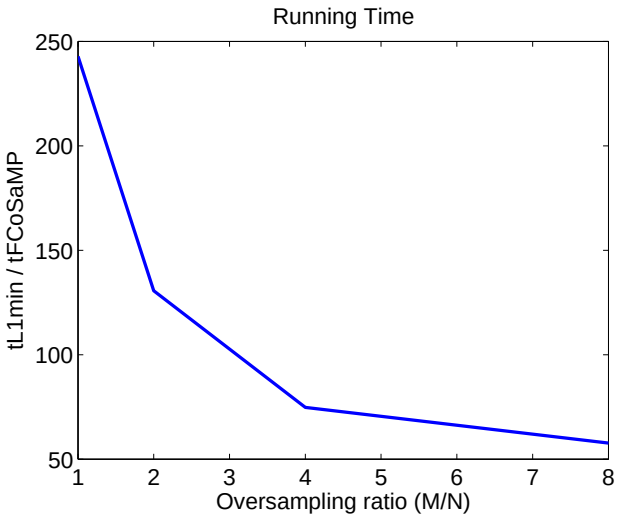

Fig. 6.17: Reconstruction algorithm execution time for `FCoSaMP` and $L1min$. Relationship between the time taken by $L1min$ and `FCoSaMP` as a function of the oversampling ratio M/N.

As it is shown in the plot, the time needed for $L1min$ is between 50 and 250 times longer than that of `FCoSaMP`. As we increase the oversampling ratio more equations enter into play, which reduce the search space of the minimization algorithm. However, this long reconstruction time make this algorithm only feasible in particular cases.

6.6 Hardware Implementation

In this section we propose a hardware implementation of the $RADS$ Converter in order to validate the ideas presented above by a real application.

The implementation was made in a reduced size PCB with off-the-shelf components. Some constrains were imposed in the design of the board, since the output of the Δ/Σ converter must be a 1-bit output, but it is rather difficult to find a commercial Δ/Σ converter with this characteristic in the market (most converters include the decimation filter as well).

Figure 6.18 shows a simplified schematic diagram of the implemented architecture, and the aspects of the implemented board can be observed in Figure 6.19.

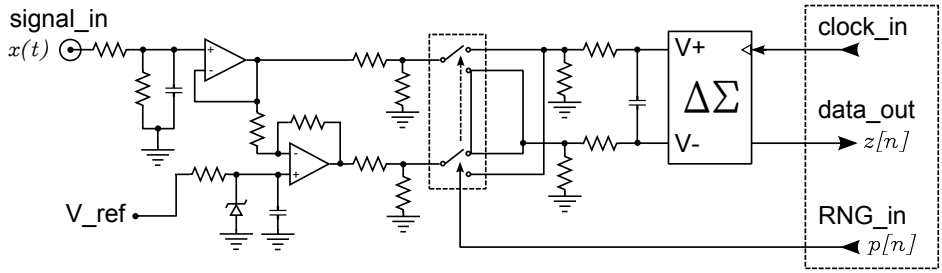

Fig. 6.18: Simplified schematic diagram of the hardware implementation of the $RADS$ Converter.

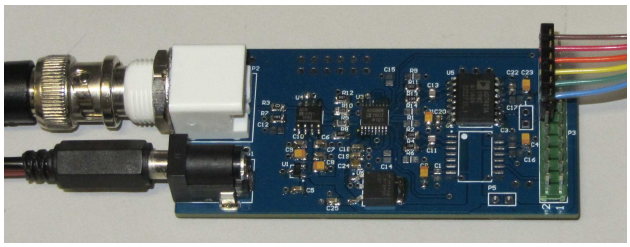

Fig. 6.19: Picture of the hardware implementation of the $RADS$ Converter.

From the user point of view the converter presents two inputs: a clock and a

random sequence, and one output: the 1-bit $RADS$ Converter output. The clock input and input for the random sequence must be synchronized, and the frequency relationship must be the oversampling ratio determined by the application. The 1-bit output is synchronized with the input clock and must be read before the next rising edge of the incoming clock.

The first stage of the converter is nothing but an amplifier, which functions is to convert the signal from single-ended to differential, and to adapt the signal input level to the Δ/Σ converter level.

After this stage, the signal is passed trough a combination of switches, that change the polarity of the signal as it is commanded by the RNG input. This processing is equivalent to the multiplication stage showed in Figure 6.3.

The last block is a conventional Δ/Σ converter, which produces a 1-bit output stream. The chosen converter was an AD7401A, from Analog Devices, which is a 2nd order discrete time modulator with a maximum sampling rate of 20MSPS.

Note that in an integrated implementation the whole architecture can be directly implemented with a slight modification of the first stage of a discrete time Δ/Σ converter.

6.6.1 Measurement Setup

Figure 6.20 and Figure 6.21 show the measurement setup. The complete setup is composed by the $RADS$ converter board, a Spartan 6 Development Board responsible for to generating the pseudorandom sequence and to interface it to a PC trough a USB port, a signal generator with GPIB interface, a power supply, and a laptop for the control and the acquisition of the measurements.

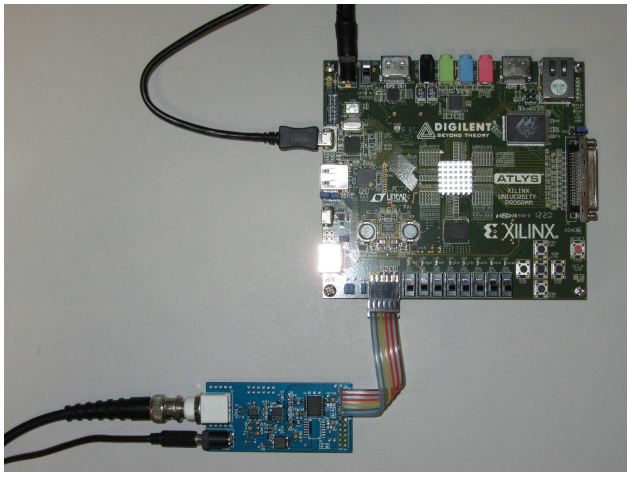

Fig. 6.20: Picture of the $RADS$ Converter connected to a Spartan 6 FPGA development kit.

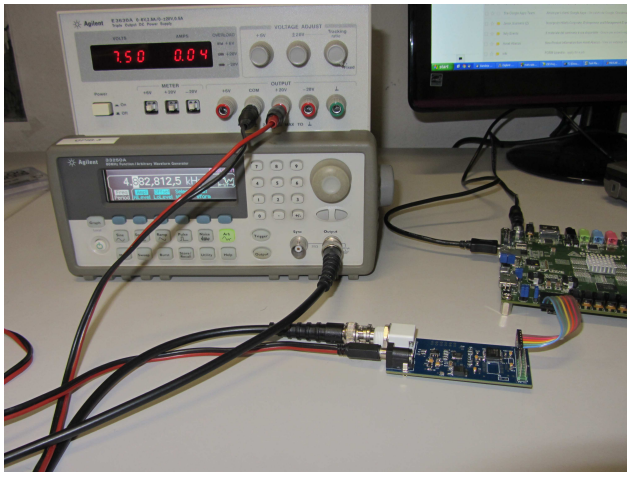

Fig. 6.21: Measurement setup for the evaluation of the hardware implementation of the $RADS$ Converter.

The measurement procedure is described below:

- Set the number of measurements (M), the Nyquist rate of the input signal

to be generated (N), the sparsity level (K), and choose a basis of sparsity for the signal.

- Generate the samples of the signal to be acquired in the PC with Matlab, and send them through the GPIB interface to the signal generator.
- Start the acquisition with the $RADS$ board, save the measurements temporally in the Spartan 6 development board, and transfer them to the PC trough USB.
- Process the acquired samples in the PC with `FCoSaMP` and compare the reconstructed signal with the synthetically generated signal.

The proposed measurement setup is very flexible and allows to exploit the whole space of parameters of the acquisition process.

6.6.2 Measurements and Validation

We have made a series of measurements in order to validate the functioning of $RADS$ Converter. We have fixed the sampling frequency to $10MHz$ and we have vary the time window in order to change the number of acquired measurements.

As an example, Figure 6.6.2 shows a plot of an 8-sparse (in a random basis with a Nyquist rate of $5MHz$) synthetic signal generated in Matlab, and superimposed to it, it is the signal acquired with $RADS$ Converter and reconstructed with `FCoSaMP`. The obtained $RSNR$ was of $32dB$.

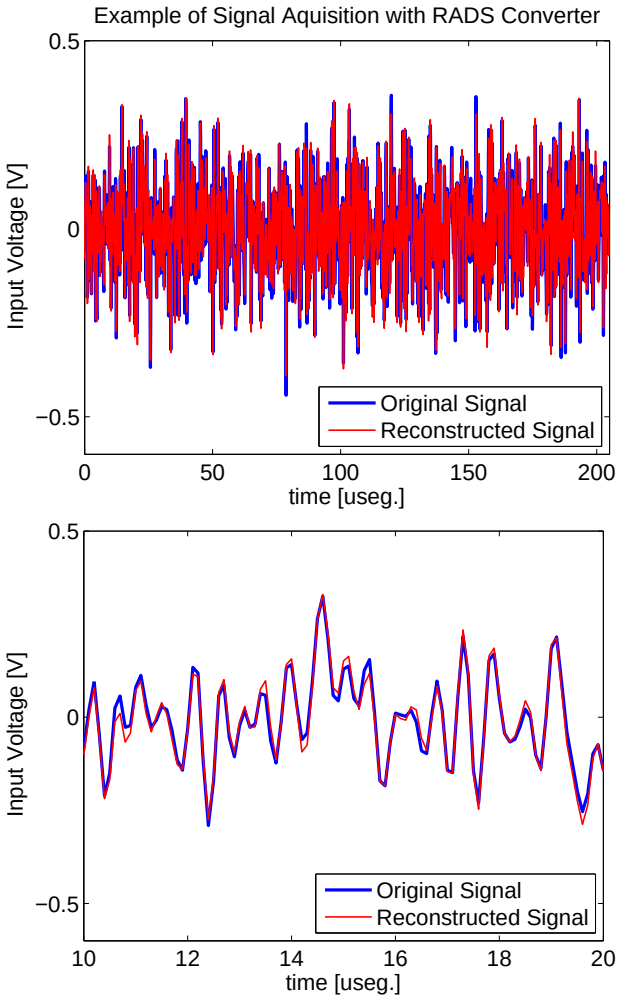

Fig. 6.22: Acquisition of an analog signal with $RADS$ Converter. The input signal is 8-sparse in a random basis and with a Nyquist rate half the sampling frequency. On top: the synthetic signal superimposed to the reconstructed signal for the whole acquisition window; on bottom: a zoom-in of the same acquired signal.

Figure 6.6.2 shows the signal input spectrum. As it is shown, the spectrum occupancy is of $2.5MHz$, which implies a Nyquist rate of $5MHz$. With the bit budget utilized by $RADS$ Converter in the acquisition of this signal, it would be obtained a maximum SNR of $12dB$ by the used of a classical Nyquist converter.

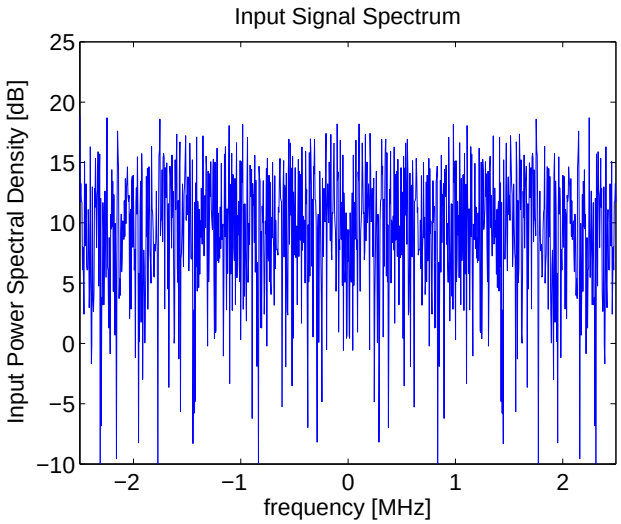

Fig. 6.23: Spectrum of the input signal acquired by $RADS$ Converter. The spectrum has a full occupancy for frequencies up to $2.5MHz$

Another example using a sparse signal in the Fourier domain is presented in the 6.24. The plot shows the mean value over 10 measurements obtained by the $RADS$ converter board.

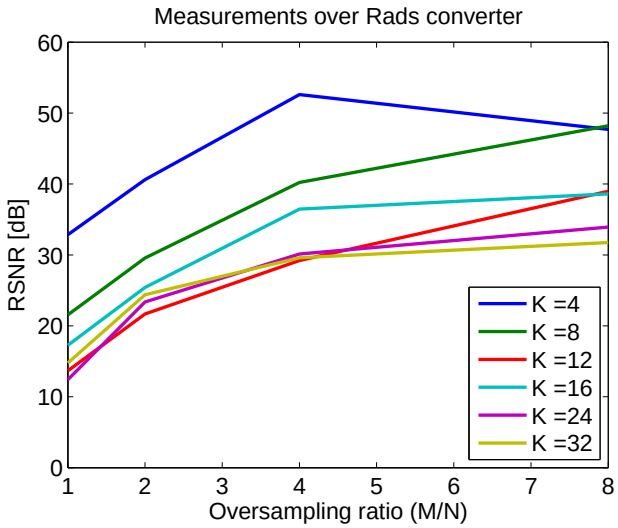

Fig. 6.24: Performance of the hardware implementation of the $RADS$ Converter by using the `FCoSaMP` algorithm for reconstruction, for different sparsity levels. $RSNR$ as a function of the oversampling ratio M/N. The support recovery was always correct for the 10 measurements.

As can be observed, the trend is the same to that obtained in the simulation, but the performance archived in terms of $RSNR$ is inferior to the expected by the simulation. These differences can be due to, imperfections in the utilized switches (on/off resistance, switching time, frequency response), different sources of noise (power supply noise, noise introduced by the amplifier, thermal noise, quantization noise in the signal generator) and most important, the bandwidth of the input stage of the utilized Δ/Σ converter (the modulated signal that enter into the Δ/Σ converter exceeds greatly the converter specification).

In spite of this, the implementation of the converter has shown that this architecture is promising as an analog-to-information converter, significantly reducing the total number of bits with respect to Nyquist based sampling for specific classes of signals.

6.7 Conclusion

In this chapter we have introduced the $RADS$ Converter, we have evaluated its performance through theoretical results, numerical simulations and a hardware implementation of the acquisition architecture. We have proposed a number of reconstruction algorithms among those we highlight the `FCoSaMP` and the *L1-norm* minimization.

The proposed architecture allows a "simple" hardware implementation for the acquisition of large bandwidth signals that are sparse over a variety of supports, obtaining a very high resolution after reconstruction. This contrast with classical sampling methods, where the resolution drastically decreases with the sampling frequency.

We have also evaluated numerically the quality of the algorithm to retrieve a correct support under different input signal condition, obtaining a very high probability over a wide range of sparsity levels.

Finally we have proposed a different approach for the study of the $RADS$ Converter and for Δ/Σ modulators in general. Contrary to what is found in the literature for this kind of converters, usually evaluated in the frequency domain [55, 56, 57, 58], the proposed approach is based on a time-domain analysis.

7. CONCLUSIONS

This thesis builds on the field of signal processing, and illustrates with two different applications how, by increasing the efforts in the digital domain, it is possible to reduce the requirements for the implementation of analog hardware.

Specifically, we have focused on the analysis of the use of very coarse quantization, more precisely 1-bit quantization, with the aim of obtaining a simplification in the implementation of both, analog to digital converters, and digital to analog converters. We have shown that a proper exploitation of binary quantization can lead to performances that are similar, and sometimes even better, than those obtained using multibit approaches.

In the first part we have proposed the use of Legendre sequences (binary sequences) for the utilization in MIMO active sensing systems. We have proposed the construction of set of sequences, where each of the sequences in the set is made from a different rotation of the same Legendre sequence. We have found that optimal rotations exist, and that the set formed by this binary sequences has a performance in terms of ISL beyond the one obtained by other sets of binary sequences. We have also found that the performance obtained by our sequences is comparable to state-of-the-art algorithms that produce real value sequences, when quantization is imposed to them up to a certain level of quantization depth.

In order to obtain the optimal rotations, we have presented an analytical expression for the calculation of the cross-correlation components of the ISL of a set of sequences. This expression, put together with a previously obtained expression for the calculation of the ISL of a single sequence, allowed the creation

of a complete expression for the ISL of a set of sequences. Under asymptotic conditions, this expression can be used to calculate the ISL of sequences whose generating function has a relatively simple trend. Since this is the case of Legendre sequences, we were able to derive an analytical expression for the asymptotic ISL of sets of rotated Legendre sequences. Such an expression was exploited to drive the optimization procedure needed to construct small-ISL sets of antipodal sequences of any sequence length with potential applications to communication and active sensing systems.

We have started the second part of this thesis by introducing the models necessary to represent the classes of signals of interest, i.e. sparse signals. We have shown how many high-dimensional signals actually have a limited number of degrees of freedom compared to its dimensionality. These classes of signals are known as sparse signals, which are one of the main ingredients for the development of the compressive sensing theory.

In this part of the thesis we have dealt particularly with the design and development of a hardware architecture for the implementation of a compressive sensing system. Based on the motivation of this thesis work, one of the requisite we have impose for the implementation of such a system, was that it must lead into a simple hardware/system implementation.

In this way , we have introduced a new architecture for an Analog to Information converter that was called the $RADS$ Converter. The proposed architecture is based on a well-known Δ/Σ converter that produces 1-bit measurements of the incoming signal. Starting from a Δ/Σ converter, a straightforward modification of the input stage topology lead to the implementation of the $RADS$ Converter architecture.

The reconstruction performance obtained using the proposed converter was found to depend on the signal information content, instead of depending on the signal bandwidth, as it is in the case for a classical Δ/Σ converter. This results in the possibility of acquisition of large bandwidth signals that are sparse over a vari-

ety of supports, with an extremely high accuracy after being processed. Based on compressive sensing concepts, $RADS$ Converter is able exploit the sparse signal structure to capture all its information content by taking single bit measurements.

An important finding of this work, was that by exploiting the peculiarities of the acquisition strategy we were able to develop a new reconstruction algorithm that produces an improved estimate (with respect to general algorithms) of the signal in terms of accuracy and probability of successful reconstruction. This suggest that, while most of the reconstruction algorithms for compressive sensing are based on guaranties on the structure of the measurement matrix (RIP based algorithms), it is possible to get a profit by generating more clever algorithms that match with the acquisition architecture itself.

The modeling of the $RADS$ Converter in the frequency domain has led to an intuitive understanding of the encoding process, and has given light on how proceed to efficiently reconstruct the input signal from the measurements.

However, in order to get a deeper insight into the functioning of the proposed converter, we were able to develop a time-domain model of the operations performed to the signal in the encoding process. With this aim we have raised an algebraic analysis of the space determined by the measurements, and its reduction as new measurements come into consideration. The study of the size of that space, evidences the difference between the $RADS$ encoding and the Δ/Σ encoding, and allows the calculation/ estimation of the theoretical maximum limit that can be expected by taking 1-bit measurements of any form.

The different perspective given by the time domain modeling of the encoding process, has led to the proposal of a new reconstruction algorithm for the $RADS$ Converter architecture. This algorithm is based on classical compressive sensing concepts that promotes sparsity through the minimization of the $L1$-norm. It has been demonstrated that the use of this algorithm can produce a better estimate of the signal than its frequency-based counterpart. However, the complex task of minimizing the $L1$-norm over the huge amount of constraints, makes this im-

provement be achieved at the expense of an increase in execution time, making this application only feasible for certain applications.

Besides the extensively numerical simulations performed during the development of this thesis to validate the results, we have implemented the $RADS$ Converter architecture in a reduced size PCB with off-the-shelf components.

The implementation of the converter has demonstrated that this architecture is promising as an analog to information converter, significantly reducing the total number of bits with respect to Nyquist based sampling, for specific classes of signals.

Although the performance attained by the hardware implementation differs from the one achieved in simulations, we believe that a proper implementation of the $RADS$ Converter in a specifically designed integrated device can lead to an increase in the performance close to the obtained in the simulation.

BIBLIOGRAPHY

[1] J. Oppermann and B. Vucetic, "Complex spreading sequences with a wide range of correlation properties," *IEEE Transactions on Communications*, vol. 45, no. 3, pp. 365 –375, mar. 1997.

[2] H. Deng, "Polyphase code design for orthogonal netted radar systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 11, pp. 3126 – 3135, nov. 2004.

[3] J. L. H. He, P. Stoica, "Designing unimodular sequence sets with good correlations - including an application to mimo radar," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4391–4405, Nov. 2009.

[4] H. He, P. Stoica, and J. Li, "Designing unimodular sequence sets with good correlations-including an application to mimo radar," *Signal Processing, IEEE Transactions on*, vol. 57, no. 11, pp. 4391 –4405, nov. 2009.

[5] H. Khan, Y. Zhang, C. Ji, C. Stevens, D. Edwards, and D. O'Brien, "Optimizing polyphase sequences for orthogonal netted radar," *IEEE Signal Processing Letters*, vol. 13, no. 10, pp. 589 –592, oct. 2006.

[6] C. Chen and P. Vaidyanathan, "Mimo radar ambiguity properties and optimization using frequency-hopping waveforms," *Signal Processing, IEEE Transactions on*, vol. 56, no. 12, pp. 5926–5936, 2008.

[7] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, 2006.

[8] 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.

[9] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289 –1306, april 2006.

[10] J. Li, P. Stoica, L. Xu, and W. Roberts, "On parameter identifiability of mimo radar," *Signal Processing Letters, IEEE*, vol. 14, no. 12, pp. 968–971, 2007.

[11] E. Fishler, A. Haimovich, R. Blum, L. Cimini Jr, D. Chizhik, and R. Valenzuela, "Spatial diversity in radars-models and detection performance," *Signal Processing, IEEE Transactions on*, vol. 54, no. 3, pp. 823–838, 2006.

[12] D. Bliss and K. Forsythe, "Multiple-input multiple-output (mimo) radar and imaging: degrees of freedom and resolution," in *Signals, Systems and Computers, 2003. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1. IEEE, 2003, pp. 54–59.

[13] M. Skolnik, "Radar handbook, 2nd ed." 1990.

[14] J. Li, P. Stoica, and X. Zheng, "Signal synthesis and receiver design for mimo radar imaging," *Signal Processing, IEEE Transactions on*, vol. 56, no. 8, pp. 3959–3968, 2008.

[15] J. Li, L. Xu, P. Stoica, K. Forsythe, and D. Bliss, "Range compression and waveform optimization for mimo radar: a cramer–rao bound based study," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 218–232, 2008.

[16] J. L. P. Stoica, H. He, "New algorithms for designing unimodular sequences with good correlation properties," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1415–1425, Apr. 2009.

[17] H. He, P. Stoica, and J. Li, "On aperiodic-correlation bounds," *Signal Processing Letters, IEEE*, vol. 17, no. 3, pp. 253–256, 2010.

[18] M. Golay, "The merit factor of legendre sequences," *IEEE Transactions on Information Theory*, vol. 29, no. 6, pp. 934–936, Oct. 1983.

[19] H. J. T. Hø holdt, "Determination of the merit factor of legendre sequences," *IEEE Transactions on Information Theory*, vol. 34, no. 1, pp. 161–164, Jan. 1988.

[20] L. B. M. Antweiler, "Merit factor of chu and frank sequences," *IEE Electronics Letters*, vol. 26, no. 25, pp. 2068–2070, Dec. 1990.

[21] R. K. P.B. Rapajic, "Merit factor based comparison of new polyphase sequences," *IEEE Communications Letters*, vol. 2, no. 10, pp. 269–270, Oct. 1998.

[22] J. Jedwab, "A survey of the merit factor problem for binary sequences," *T. Helleseth et al, eds., Lecture Notes in Computer Science, Sequences and their Applications - Proceedings of SETA 2004, Springer-Verlag*, vol. 3486, pp. 30–55, jan. 2005.

[23] Best known sequences. [Online]. Available: http://www2.chemistry.msu.edu/faculty/linedantus/merit_factor_records.html

[24] T. Høholdt, "The merit factor problem for binary sequences," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2006, vol. 3857, pp. 51–59.

[25] G. S. d. G. Polya, *Aufgeben und Lehrs¨atze aus der Analyse II.* Berlin: Springer, 1925.

[26] S. J. Haboba, R. Rovatti, and G. Setti, "Determination of the integrated sidelobe level of sets of rotated legendre sequences," *arXiv preprint arXiv:1012.3638*, 2010.

[27] G. G. S. W. Golomb, *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar*. Cambridge University Press, 2005.

[28] L. Maximon, "The dilogarithm function for complex argument," *Proceedings of the Royal Society, part A: Mathematical, Physical & Engineering Sciences*, vol. 459, pp. 2807–2819, 2003.

[29] J. Haboba, R. Rovatti, and G. Setti, "Integrated sidelobe level of sets of rotated legendre sequences," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2632–2635.

[30] E. Candes and T. Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, 2005.

[31] 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.

[32] S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann, "Uniform uncertainty principle for bernoulli and subgaussian ensembles," *Constructive Approximation*, vol. 28, no. 3, pp. 277–289, 2008.

[33] E. J. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on pure and applied mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[34] 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.

[35] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM journal on scientific computing*, vol. 20, no. 1, pp. 33–61, 1998.

[36] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *Information Theory, IEEE Transactions on*, vol. 52, no. 1, pp. 6–18, 2006.

[37] J. A. Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *Information Theory, IEEE Transactions on*, vol. 52, no. 3, pp. 1030–1051, 2006.

[38] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *The Annals of statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[39] D. Donoho and Y. Tsaig, "Fast solution of l1-norm minimization problems when the solution may be sparse," *Information Theory, IEEE Transactions on*, vol. 54, no. 11, pp. 4789 –4812, nov. 2008.

[40] 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.

[41] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 310 –316, april 2010.

[42] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655 –4666, dec. 2007.

[43] M. Fornasier and H. Rauhut, "Iterative thresholding algorithms," *Applied and Computational Harmonic Analysis*, vol. 25, no. 2, pp. 187–208, 2008.

[44] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 629–654, 2008.

[45] K. Herrity, A. Gilbert, and J. Tropp, "Sparse approximation via iterative thresholding," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 3, may 2006, p. III.

[46] 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 *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*. IEEE, 2007, pp. 1959–1962.

[47] J. Laska and R. Baraniuk, "Regime change: Bit-depth versus measurement-rate in compressive sensing," *Signal Processing, IEEE Transactions on*, vol. 60, no. 7, pp. 3496–3505, 2012.

[48] P. Boufounos and R. Baraniuk, "1-bit compressive sensing," in *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*. IEEE, 2008, pp. 16–21.

[49] P. Boufounos, "Greedy sparse signal reconstruction from sign measurements," in *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*. IEEE, 2009, pp. 1305–1309.

[50] J. Laska, Z. Wen, W. Yin, and R. Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *Signal Processing, IEEE Transactions on*, vol. 59, no. 11, pp. 5289–5301, 2011.

[51] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *arXiv preprint arXiv:1104.3160*, 2011.

[52] J. Haboba, M. Mangia, R. Rovatti, and G. Setti, "An architecture for 1-bit localized compressive sensing with applications to eeg," in *Biomedical Circuits and Systems Conference (BioCAS), 2011 IEEE*. IEEE, 2011, pp. 137–140.

[53] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, vol. 2, no. 3, pp. 443–459, 2012.

[54] J. Haboba, R. Rovatti, and G. Setti, "Rads converter: An approach for analog to information conversion," in *Electronic Circuits and Systems (ICECS), 2012 IEEE International Conference on*. IEEE, 2012, pp. 49–52.

[55] R. Schreier and G. C. Temes, *Understanding delta-sigma data converters*. IEEE press Piscataway, NJ, USA, 2005, vol. 74.

[56] S. R. Norsworthy, R. Schreier, G. C. Temes *et al.*, *Delta-sigma data converters: theory, design, and simulation*. IEEE press New York, 1997, vol. 97.

[57] J. C. Candy and G. C. Temes, *Oversampling delta-sigma data converters: theory, design, and simulation*. IEEE press New York, 1992.

[58] G. I. Bourdopoulos, A. Pnevmatikakis, V. Anastassopoulos, and T. L. Deliyannis, *Delta-sigma modulators: modeling, design and applications*. World Scientific Publishing Company, 2003.

[59] R. Schreier. (2009) Delta sigma toolbox. [Online]. Available: www.mathworks.com/matlabcentral/fileexchange/19-delta-sigma-toolbox

[60] The mathworks, inc. [Online]. Available: http://www.mathworks.it/products/matlab/

[61] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *Information Theory, IEEE Transactions on*, vol. 56, no. 1, pp. 520 –544, jan. 2010.

[62] J. Laska, Z. Wen, W. Yin, and R. Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *Signal Processing, IEEE Transactions on*, 2011, to appear.

[63] M. E. Newman and G. T. Barkema, "Monte carlo methods in statistical physics," *Monte Carlo methods in statistical physics/MEJ Newman and GT Barkema. Oxford: Clarendon Press, 1999.*, vol. 1, 1999.

[64] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes with Source Code CD-ROM 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

[65] Ibm ilog cplex optimizer. [Online]. Available: http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/