Compressive Sensing of Localized Signals: Application to Analog-to-Information Conversion

Juri Ranieri
ARCES
Università di Bologna
Bologna, Italy
email: juri.ranieri@gmail.com

Riccardo Rovatti
ARCES
Università di Bologna
Bologna, Italy
email: riccardo.rovatti@unibo.it

Gianluca Setti
ENDIF
Università di Ferrara
Ferrara, Italy
email: gsetti@unife.it

***Abstract*—Compressed sensing hinges on the sparsity of signals to allow their reconstruction starting from a limited number of measures. When reconstruction is possible, the SNR of the reconstructed signal depends on the energy collected in the acquisition. Hence, if the sparse signal to be acquired is known to concentrate its energy along a known subspace, an additional *"rakeness"* criterion arises for the design and optimization of the measurement basis. Formal and qualitative discussion of such a criterion is reported within the framework of a well-known Analog-to-Information conversion architecture and for signals localized in the frequency domain. Non-negligible inprovements are shown by simulation.**

I. Introduction

The convergence between Compressive Sensing (CS) techniques [1] [2] and Analog-to-Digital (AD) conversion architectures is a natural consequence of the gap between the performance of interface electronic circuits (expecially those working at radio frequencies) and that of processing cores.

In this framework, the ability of CS of reconstructing a sparse signal by feeding a limited number of measurements into a possibly computationally intensive procedure is an ideal candidate for future sensing systems.

The overall idea is that, though the bandwidth occupation of a signal is potentially large, its information content is not necessarily equally large, radar signals being one of the most striking examples [3]. In these cases, trying to match the measurement rate with the actual information rate of the signal may allow to reduce the sampling rate of the acquisition system.

This paradigm has been recently named Analog-to-Information (AI) conversion since digital samples play a secondary role in the process, that is designed and sized considering the mathematical structure of the signals instead of their mere bandwidth. The very same output of an AI converter is not necessarily the complete sensed waveform but, most often, the minimum number of data to encode the information contained therein, i.e., its compressed version.

In more formal terms, the mathematical assumption on which CS hinges is the sparsity of the signal, i.e., the possibility of expressing each interesting waveform as a linear combination of a very limited number of waveforms taken from a much more numerous waveform basis (which is called the *sparsity basis*).

We are here interested in signals for which an additional a-priori information is available, i.e., that the majority of their average energy is localized along a pre-defined waveform subspace.

Note that, though they both deal with how signal projections align with given subspaces, sparsity and localization are different properties. In fact, the sparsest possible signal - one that is always a scaled version of a single basis waveform - may be completely non-localized if its realizations excite each of the basis waveforms with a substantially non-vanishing probability.

Conversely, even if a signal localized along a low-dimensional subspace can be given a basis in which each realization has few dominant coefficients, this happens only "on average" to make most of the energy lie along that subspace and cannot guarantee that only those coefficients are nonvanishing.

In presence of sparsity only, one of the main criteria for the design of measures is *incoherence* [4], i.e., the fact that acquired projections do not align with any particular waveform in the sparsity basis. To make reconstruction robust to noise *restricted isometry* may be also required, i.e., the ability of the overall measuring operator to conserve the magnitude of sparse vectors.

When localization comes into play, additional criteria may arise that we will discuss in the development of a specific application.

In fact, in this paper, sparse and localized signals are acquired by means of a Random Modulation Pre-Integration (RMPI) converter [5]. Such an architecture exploits convolution with an antipodal chipping sequence (an hardware-friendly operation) and subsequent integration prior to sampling at sub-Nyquist rate. Incoherence between measurements and the sparsity basis is pursued imposing randomness of the chipping sequences that are usually generated to mimic a sequence of independent and identically distributed antipodal random variables.

Localization of signals is in the frequency domain, i.e., we assume that the power spectrum of the sparse signals has a strongly unimodal profile favoring frequencies in a predetermined interval.

We will leverage on some results on spectral shaping of antipodal sequences [6] [7] [8] to show that suitably designed chipping is able to improve the SNR of the signal that is reconstructed starting from the same number of measurement.

978-1-4244-5309-2/10/$26.00 ©2010 IEEE

II. CS AND RMPI ARCHITECTURE

To model sparse continuous-time signals we assume that each realization $x(t)$ of theirs can be expressed as a linear combination of a set of orthonormal waveforms $\psi_k(t)$. This is written as $x(t) = \Psi(t)^\top \alpha$ where $\cdot^\top$ stands for vector transposition, $\Psi(t)^\top = (\psi_0(t), \psi_1(t), \dots,)$ and $\alpha = (\alpha_0, \alpha_1, \dots)^\top$ are vectors of, respectively, functions of time and real numbers.

Since we work in a noisy environment, the signal fed into the converter is actually $x(t)+\nu(t)$ where $\nu(t)$ is a realization of a white Gaussian stochastic process. Measurements of $x(t) + \nu(t)$ are taken as its projections along M orthogonal waveforms $\phi_j(t)$, $j = 0, 1, \dots, M-1$ arranged in a vector of functions of time $\Phi = (\phi_0(t), \phi_1(t), \dots, \phi_{M-1}(t))^\top$ to produce an M-dimensional vector y with real entries

$$y = \left\langle \Phi(t), \Psi(t)^\top \alpha + \nu(t) \right\rangle = \Theta\alpha + n \tag{1}$$

where the scalar product is defined between real functions of time as $\langle f, g\rangle = \int_{-\infty}^{\infty} f(t)g(t)\mathrm{d}t$, $\Theta = \left\langle \Phi(t), \Psi(t)^\top \right\rangle$ is a matrix with M rows and a countable number of columns that contains the projections

$$\Theta_{j,k} = \left\langle \phi_j(t), \psi_k(t) \right\rangle, \tag{2}$$

and n is the vector of white Gaussian perturbations due to noise.

In the following we will consider the above model only in finite terms and assume that $N > M$ waveforms $\psi_k(t)$ are needed to express all the interesting realizations of $x(t)$ and that α is K-sparse, i.e. that only K of its N entries are non-vanishing.

With this, we fall into the most common setting for a CS problem, i.e., that of a linear, non-invertible operator linking the signal (through its defining coefficients) with the measurements.

CS theory puts constraint on the matrix Θ to guarantee that the K non-vanishing components of α can be retrieved.

A very common requirement stems from thinking that the ϕ_j are actually taken among N waveforms that form an orthonormal basis different from ψ_k and defining $\gamma = \max_{j,k} |\langle \phi_j, \psi_k \rangle|$ as the *coherence* between the two basis. It is then required that the two basis are *incoherent*, i.e., that γ is as close as possible to its lower bound that is $N^{-1/2}$.

Beyond formal justification, this is very sensible since a sparse signal has vanishing projections along many of the ψ_j and a measurement perfectly aligned with those directions has an high probability of being null and thus to convey no information.

Restricted isometry [9] requires that the ratio between the length of the mapped vector and the one of the original vector is bounded between $1 - \delta$ and $1 + \delta$ for a certain, hopefully small, isometry constant $\delta > 0$.

Within the model in (1), the RMPI architecture is nothing but an hardware-friendly way of building a matrix Θ that satisfies the CS requirements so that the information within the original analog signal (the vectors of coefficients α) can be computed starting from the measures, thus performing an AI conversion.

The sampling architecture is represented in Figure 1: a modulator multiplies the input analog signal by a rectangular PAM with random antipodal symbols (the chipping sequence)

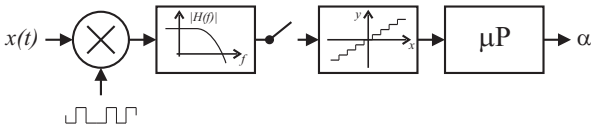

Fig. 1. Block diagram of an RMPI architecture.

at a frequency f_c not smaller than the Nyquist one, the modulated signal is filtered by an analog lowpass filter whose output is sampled by an AD converter at a sub-Nyquist rate.

Since this short contribution aims at giving a proof of concept, circuit blocks are idealized and the analog filter is replaced with the computation of the average of the chipped signal over subsequent and non-overlapping time windows of duration T. With this, the j-th converted sample is equivalent to the projection of the input signal over the slice of the chipping PAM contained in $[jT, (j+1)T]$.

The chipping PAM clipped to the j-th of those windows implicitly defines ϕ_j. Since the windows are non-overlapping, the ϕ_j are orthogonal. Once M subsequent projections are collected in the vector y, a proper reconstruction algorithm is run to yield the coefficient vector α.

Though no formal guarantee can be given that such a reconstruction is possible, incoherence and restricted isometry of the resulting Θ can be reasonably expected if we randomize the ϕ_j [10], i.e., if we randomize the ± 1 controlling their amplitude at each instant. In fact, both requirements are essentially prescribing that projections in Θ have a magnitude as uniform as possible (singularly as in incoherence, or along column subsets as in restricted isometry) and randomizing the measurement directions is surely an appealing way of avoiding predominant components.

III. ENERGY RAKING, SPECTRUM LOCALIZATION, AND CHIPPING SEQUENCE DESIGN

In the original architecture [5] the chipping sequence giving raise to the ϕ_j is generated by a *Linear Feedback Shift Register* (LFSR), so that its spectrum is a good approximation of a flat one.

Yet, when considering the noise injected in the acquisition flow, the use of chipping spectral profile as a further degree of freedom may lead to a significant optimization.

In fact, since $\nu(t)$ in (1) is white and the waveforms in $\Phi(t)$ are normalized, the power associated with the disturbance vector n does not depend on the choice of the measurement directions.

Yet, the power of the signal component in the j-th sample is

$$\mathbf{E}_{\phi_j,\alpha}\left[\frac{1}{T}\left|\langle \phi_j(t), x(t)\rangle\right|^2\right] \tag{3}$$

depends on the matching of the statistics of $\phi_j(t)$ with $x(t)$.

Note that, in general, for any two slices of independent stationary processes $a(t)$ and $b(t)$ as well as their power spectrum $\hat{a}(f)$ and $\hat{b}(f)$ we have

$$\begin{aligned}\rho(a,b) = \mathbf{E}_{a,b}\left[\frac{1}{T}\left|\langle a(t), b(t)\rangle\right|^2\right] = \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{a}(f')\hat{b}(f'')h_T(f'-f'')\mathrm{d}f'\mathrm{d}f''\end{aligned} \tag{4}$$

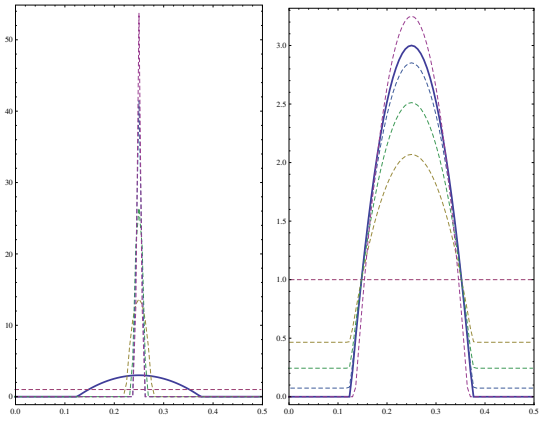

Fig. 2. Solutions of (5) when $\hat{x}(f)$ is the solid thick track and ϵ is very large (left) or very small (right).

with $h_T(f) = \frac{\sin^2(\pi f T)}{\pi^2 T f^2}$ that, when T is large, approximates $\delta(f)$ thus yielding $\rho(a,b) \overset{T \text{ large}}{\simeq} \int_{-\infty}^{\infty} \hat{a}(f)\hat{b}(f)\mathrm{d}f$.

We may think of maximizing the power raked by our measurements saying that all the projection waveforms share the same power spectrum $\hat{\phi}$ and solving the optimization problem

$$\begin{array}{ll} \max_{\hat{\phi}} & \rho(\phi, x) \\ & \int_{-\infty}^{\infty} \hat{\phi}(f)\mathrm{d}f = 1 \\ \text{s.t.} & \hat{\phi}(f) \geq 0 \\ & \hat{\phi}(-f) = \hat{\phi}(f) \end{array}$$

For large T this yields the optimal projection spectrum $\hat{\phi}(f) \simeq {}^1\!/\!_2\delta(f - f_0) + {}^1\!/\!_2\delta(f + f_0)$ where f_0 is any of the frequencies at which $\hat{x}$ achieves its maximum.

Yet, this makes the $\phi_j(t)$ stretches of a *statistically predictable* process thus going against the use of randomness to ensure incoherence and thus solvability of the inverse problem by $\mathbb{L}_1$-based techniques (this is especially true in RMPI architecture where the set of ϕ_j is restricted to antipodal PAM).

A randomness constraint must then be introduced in the optimization of the projection waveforms. One of the possible choices is to require that the average correlation between $\phi_{j'}$ and $\phi_{j''}$ cannot exceed a prescribed threshold ϵ. With this, the best $\hat{\phi}$ should solve

$$\begin{array}{ll} \max_{\hat{\phi}} & \rho(\phi, x) \\ & \rho(\phi, \phi) \leq \epsilon \\ & \int_{-\infty}^{\infty} \hat{\phi}(f)\mathrm{d}f = 1 \\ \text{s.t.} & \hat{\phi}(f) \geq 0 \\ & \hat{\phi}(-f) = \hat{\phi}(f) \end{array} \tag{5}$$

Such a problem entails a linear obective function with a set of linear and quadratic constraints, and its solutions can be approximated by careful discretization. In Figure 2 we report such solutions (plotted only against the positive semiaxis of normalized frequencies) for ϵ in different ranges when $\hat{x}(f)$ consists of a symmetric parabolic profile taking up half of the total available bandwidth (the thick and solid track).

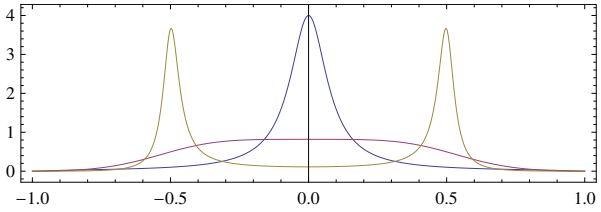

Fig. 3. Spectra of antipodal PAM signals modulated by a two-state symmetric Markov chain transition probability $p_0 = 0.55$ (flat), $p_0 = 0.2$ (lowpass), and $p_0 = 0.8$ (highpass). Frequencies are normalized to the inverse of T_c.

Note how, when ϵ is large (left of Figure 2), the randomness constraint is inactive and $\hat{\phi} \simeq {}^1\!/\!_2\delta(f - f_0) + {}^1\!/\!_2\delta(f + f_0)$ (that discretization approximates with a very steep profile centered at f_0). Yet, as ϵ is very small (right of Figure 2), the peaks in the solution of (5) are widened to follow the profile of the higher part of $\hat{x}(f)$. In fact, even if the randomness constraint forbids single-frequency spectra, frequencies at which $\hat{x}(f)$ is relatively large are always favored in maximizing $\rho(\phi, x)$. In the limit, as ϵ decreases to enforce the randomness constraint there is no room for optimization and the procedure yields a flat spectrum.

Note that the presence of the randomness constraint (and the underlying need for incoherence as a prerequisite for truly compressive sensing) distinguishes this path from the classical derivation of a matched-filter receiver, that is concerned with the direct and linear reconstrucion of the signal independently of its sparse structure.

In a sense, our considerations lead us to a less prescriptive and more flexible framework in which, though not necessarily optimal in the sense of (5), smooth spectral profiles that rise in correspondence to frequencies at which $\hat{x}(f)$ is large can be reasonably expected to be good solutions to the trade-off between the need to rake the energy of a frequency localized signals and the need to provide a random measuring base that is incoherent with the sparsity base of the same signal.

This is actually good news since, in RMPI architectures, the ϕ_j are antipodal PAM signals whose spectrum may be difficult to shape in an arbitrary way.

Few steps in this direction are actually taken in [6], [7] and most recently in [8]. In the case at hand we resort to the simplest generator of antipodal sequences with an adjustable spectrum, i.e., a two-states Markov chain in which a single probability p is assigned to both the transition from $+1$ to -1 and to the opposite one.

The power spectrum of a PAM $\mu(t)$ with pulse duration $T_\mathrm{c} = 1/f_\mathrm{c}$ and driven by the output of such a Markov chain

$$\hat{\mu}(f) = \frac{\sin^2(\pi T_\mathrm{c} f)}{(\pi f)^2} \frac{2(1-p)p}{1 - 2(1-p)p - (1-2p)\cos(2\pi f T_\mathrm{c})}$$

is roughly controlled by the parameter p in that it is substantially flat for $p = p_0 \simeq 0.55$, is increasingly lowpass for $p \leq p_0$, and increasingly highpass for $p \geq p_0$ (see Figure 3).

IV. Simulations

To substantiate with numerical evidence the suggestion that, to maintain incoherence while increasing energy rakeness, the projection waveforms should mimic the spectral behavior of

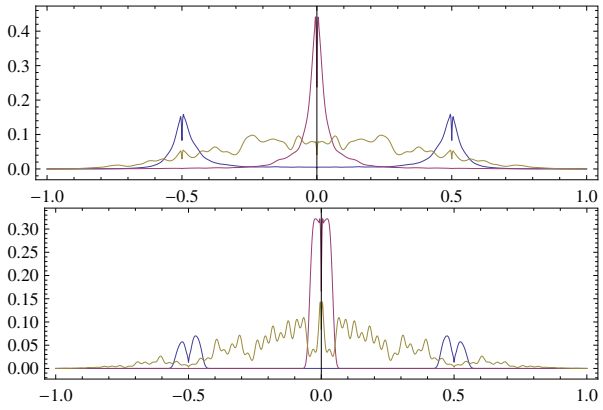

Fig. 4. Spectra of low-pass, flat, and high-pass input signals in the PAM (upper) and sinusoidal (lower) cases.

	SER (dB)					
	mode 1			mode 2		
chipping	LP	FL	HP	LP	FL	HP
LP	31.40	26.88	23.51	29.73	18.11	16.74
FL	27.72	27.64	27.85	25.01	25.09	24.31
HP	23.14	27.42	31.79	17.76	18.79	28.58

TABLE I
QUALITY OF RECONSTRUCTION ACHIEVED BY LOWPASS (LP), FLAT (FL), AND HIGHPASS (HP) CHIPPING WHEN DIFFERENT SPARSE AND LOCALIZED SIGNALS IN DIFFERENT MODES ARE ACQUIRED BY AN RMPI ARCHITECTURE.

the input signals, we simulated an RMPI architecture acquiring sparse signals with spectral profiles similar to those in Figure 3.

Such signals where produced following two modes of operations. In both modes the signals were a linear combination of $K = 5$ out of $N = 32$ basis waveform. In mode 1 the sparsity basis was made of independent stretches of antipodal PAM signals with pulse time equal to the chipping time T_c and modulating symbols with a low-pass, high-pass, or flat Markov behavior.

In mode 2 the sparsity basis was made of sinusoids with frequencies equally spaced either in a range of frequencies close to DC (low pass), or in a range far from DC (high-pass), or in a range encompassing both low and high frequencies.

In both modes, noise was injected so that the input waveform featured an SNR of 30dB.

Figure 4 shows the spectra of the two sets of signals estimated a-posteriori on the simulation data to clarify that, though the two modes entail completely different signal structure, they both yield the same qualitative spectral behavior.

Acquisition was performed by projection along 20 independent stretches of antipodal PAM controlled by a two-state Markov chain with and adjustable p_0 probability.

The projections were fed into a standard *l1_ls* algorithm [11] to estimate the coefficient vector α. Each acquisition and reconstruction was repeated 1000 times to yield a reliable estimate of the quality of the result.

Table I report the Signal-to-Error Ratio (SER) of the estimated coefficient vector. As expected, when the spectral profile of the chipping sequences matches that of the input performance is maximized.

This is less evident for flat spectra (the second and fifth column of Table I) in which energy is almost uniformly distributed and thus can be raked with the same efficiency by projection in any direction.

For strongly uneven spectral profiles, the effect of matching is more evident and allow to reconstruct the coefficients α with approximately the quality of the raw analog incoming signals (30dB).

V. CONCLUSIONS

We have introduced the concept of signal localization and paired it with signal sparsity to suggest that, within a CS framework, the criteria to design the measurement basis may take advantage of additional this additional a-priori knowledge on the signals to acquire.

Though stemming from general considerations, the idea was here exemplified with reference to an idealized RMPI architecture for AI conversion when sparse signals are also localized in the frequency domain. A mathematical optimization problem was laid down whose solutions provided qualitative hints for the design of effective chipping sequences. Sequences implementing those hints can be generated by leveraging on some recent results on spectrum shaping for antipodal processes.

We have simulated a simplified version of the overall acquisition and reconstruction process to show that the quality of the conversion is indeed improved by a sequence design following the newly proposed guidelines.

REFERENCES

[1] D.L. Donoho, "Compressed Sensing," IEEE TRANSACTIONS ON INFORMATION THEORY, vol.52, pp. 1289–1306, 2006
[2] E.J. Candès, M.B. Wakin, "An introduction to compressive sampling," IEEE SIGNAL PROCESSING MAGAZINE, vol. 25, pp. 21–30, 2008
[3] M. Vetterli, P. Marziliano, T. Blu, "Sampling signals with finite rate of innovation," IEEE TRANSACTIONS ON SIGNAL PROCESSING, vol. 50, pp. 1417–1428, 2002.
[4] E. Candès, J. Romberg, "Sparsity and incoherence in compressive sampling," Inverse Problems, vol. 23, pp. 969–985, 2007
[5] J.N. Laska, S. Kirolos, M.F. Duarte, T.S. Ragheb, R.G. Baraniuk, Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," *IEEE Intenational Symposium on Circuits and Systems*, pp. 1959–1962, 2007
[6] G. Setti, G. Mazzini, R. Rovatti, S. Callegari, "Statistical Modeling of Discrete-Time Chaotic Processes - Basic Finite-Dimensional Tools and Applications,"PROCEEDINGS OF THE IEEE, vol. 90, pp. 662-690, 2002
[7] R. Rovatti, G. Mazzini, G. Setti, A. Giovanardi, "Statistical Modeling of Discrete-Time Chaotic Processes - Advanced Finite-Dimensional Tools and Applications,", PROCEEDING OF THE IEEE, vol. 90, pp. 820-841, 2002
[8] R. Rovatti, G. Mazzini, G. Setti, "Memory-m Antipodal Processes: Spectral Analysis and Synthesis," IEEE TRANSACTION ON CIRCUITS AND SYSTEMS - I, Vol. 56, pp. 156-167, 2009
[9] E. Candès, T. Tao, "Decoding by linear programming," IEEE TRANSACTIONS ON INFORMATION THEORY, vol. 51, pp. 4203–4215, 2005
[10] R.G. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," Constructive Approximation, vol. 28, pp. 253–263, 2008
[11] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, "An Interior-Point Method for Large-Scale l1-Regularized Least Squares," IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING, vol 1, pp. 606–617, 2007