

Universal and efficient compressed sensing by spread spectrum and application to realistic Fourier imaging techniques

Gilles Puy^{1,2}, Pierre Vandergheynst¹, Rémi Gribonval³, Yves Wiaux^{1,3,4}

¹Institute of Electrical Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

²Institute of the Physics of Biological Systems, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

³Centre de Recherche INRIA Rennes-Bretagne Atlantique, F-35042 Rennes cedex, France.

⁴Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

⁵Department of Radiology and Medical Informatics, University of Geneva (UniGE), CH-1211 Geneva, Switzerland.

Email: Gilles Puy* - gilles.puy@epfl.ch; Pierre Vandergheynst - pierre.vandergheynst@epfl.ch; Rémi Gribonval - remi.gribonval@inria.fr; Yves Wiaux - yves.wiaux@epfl.ch;

*Corresponding author

Abstract

We advocate a compressed sensing strategy that consists of multiplying the signal of interest by a wide bandwidth modulation before projection onto randomly selected vectors of an orthonormal basis. Firstly, in a digital setting with random modulation, considering a whole class of sensing bases including the Fourier basis, we prove that the technique is *universal* in the sense that the required number of measurements for accurate recovery is optimal and independent of the sparsity basis. This universality stems from a drastic decrease of coherence between the sparsity and the sensing bases, which for a Fourier sensing basis relates to a spread of the original signal spectrum by the modulation (hence the name “spread spectrum”). The approach is also *efficient* as sensing matrices with fast matrix multiplication algorithms can be used, in particular in the case of Fourier measurements. Secondly, these results are confirmed by a numerical analysis of the phase transition of the ℓ_1 -minimization problem. Finally, we show that the spread spectrum technique remains effective in an analog setting with chirp modulation for application to realistic Fourier imaging. We illustrate these findings in the context of radio interferometry.

1 Introduction

In this section we concisely recall some basics of compressed sensing, emphasizing on the role of mutual coherence between the sparsity and sensing bases. We discuss the interest of improving the standard acquisition strategy in the context of Fourier imaging techniques such as radio interferometry and magnetic resonance imaging (MRI). Finally, we highlight the main contributions of our work, advocating a universal and efficient compressed sensing strategy coined spread spectrum, and describe the organization of this article.

1.1 Compressed sensing basics

Compressed sensing is a recent theory aiming at merging data acquisition and compression [1–7]. It predicts that sparse or compressible signals can be recovered from a small number of linear and non-adaptative measurements. In this context, Gaussian and Bernoulli random matrices, respectively with independent standard normal and ± 1 entries, have encountered a particular interest as they provide optimal conditions in terms of the number of measurements needed to recover sparse signals [4, 5]. However, the use of these matrices for real-world applications is limited for several reasons: no fast matrix multiplication algorithm is available, huge memory requirements for large scale problems, difficult implementation on hardware, etc.

Let us consider s -sparse digital signals $\mathbf{x} \in \mathbb{C}^N$ in an orthonormal basis $\Psi = (\psi_1, \dots, \psi_N) \in \mathbb{C}^{N \times N}$. The decomposition of \mathbf{x} in this basis is denoted $\boldsymbol{\alpha} = (\alpha_i)_{1 \leq i \leq N} \in \mathbb{C}^N$, $\boldsymbol{\alpha} = \Psi^* \mathbf{x}$ (\cdot^* denotes the conjugate transpose), and contains s non-zero entries. The original signal \mathbf{x} is then probed by projection onto m randomly selected vectors of another orthonormal basis $\Phi = (\phi_1, \dots, \phi_N) \in \mathbb{C}^{N \times N}$. The indices $\Omega = \{l_1, \dots, l_m\}$ of the selected vectors are chosen independently and uniformly at random from $\{1, \dots, N\}$. We denote Φ_Ω^* the $m \times N$ matrix made of the selected rows of Φ^* . The measurement vector $\mathbf{y} \in \mathbb{C}^m$ thus reads as

$$\mathbf{y} = A_\Omega \boldsymbol{\alpha}, \tag{1}$$

where

$$A_\Omega = \Phi_\Omega^* \Psi \in \mathbb{C}^{m \times N}. \tag{2}$$

We also denote $A = \Phi^* \Psi \in \mathbb{C}^{N \times N}$. Finally, we aim at recovering $\boldsymbol{\alpha}$ by solving the ℓ_1 -minimization problem

$$\boldsymbol{\alpha}^* = \arg \min_{\bar{\boldsymbol{\alpha}} \in \mathbb{C}^N} \|\bar{\boldsymbol{\alpha}}\|_1 \text{ subject to } \mathbf{y} = A_\Omega \bar{\boldsymbol{\alpha}}, \tag{3}$$

where $\|\bar{\alpha}\|_1 = \sum_{i=1}^N |\bar{\alpha}_i|$ ($|\cdot|$ denotes the complex magnitude), and the reconstructed signal \mathbf{x}^* satisfies $\mathbf{x}^* = \Psi \alpha^*$.

The theory of compressed sensing already demonstrates that a small number $m \ll N$ of random measurements are sufficient for an accurate and stable reconstruction of \mathbf{x} [6, 7]. However, the recovery conditions depend on the mutual coherence μ between Φ and Ψ . This value is a similarity measure between the sensing and sparsity bases. It is defined as $\mu = \max_{1 \leq i, j \leq N} |\langle \phi_i, \psi_j \rangle|$ and satisfies $N^{-1/2} \leq \mu \leq 1$. Performances are optimal when the bases are perfectly incoherent, i.e. $\mu = N^{-1/2}$, and unavoidably decrease when μ increases.

1.2 Fourier imaging applications and mutual coherence

The dependence of performance as a function of the mutual coherence μ is a key concept of compressed sensing of significant interest for applications such as Fourier imaging, in particular radio interferometry or magnetic resonance imaging (MRI), where signals are probed in the orthonormal Fourier basis. In radio interferometry, one of the main challenge is to reconstruct accurately the original signal from a limited number of accessible measurements [8–10]. In MRI, accelerating the acquisition process by reducing the number of measurements is of huge interest for applications ranging from structural imaging [11–13] to dynamic imaging [14]. The theory of compressed sensing shows that Fourier acquisition would represent the best sampling strategy if the signals probed were sparse in the Dirac basis. The sensing system would indeed be optimally incoherent. Unfortunately, natural signals are usually rather sparse in multi-scale bases, e.g. wavelet bases, which are coherent with the Fourier basis. Many measurements are thus needed to reconstruct accurately the original signal. In the perspective of accessing better performance, sampling strategies with improved incoherence of the system should be considered.

1.3 Main contributions and organization

In the present work, we advocate a compressed sensing strategy coined spread spectrum that consists of a wide bandwidth pre-modulation of the signal \mathbf{x} of interest before projection onto randomly selected vectors of an orthonormal basis. In the particular case of Fourier measurements, the pre-modulation amounts to a convolution in the Fourier domain which spreads the power spectrum of the original signal \mathbf{x} (hence the name “spread spectrum”), while preserving its norm. Equivalently, this spread spectrum phenomenon acts on each sparsity basis vector describing \mathbf{x} so that information of each of them is accessible whatever the Fourier coefficient selected. This effect implies a decrease of coherence between the sparsity and sensing

bases and enables an enhancement of the reconstruction quality.

In Section 2, we study the spread spectrum technique in a digital setting for arbitrary pairs of sensing and sparsity bases (Φ, Ψ) . We consider a digital pre-modulation $\mathbf{c} = (c_l)_{1 \leq l \leq N} \in \mathbb{C}^N$, with $|c_l| = 1$ and random phases identifying a random Rademacher or Steinhaus sequence. We show that the recovery conditions do not depend anymore on the coherence of the system but on a new parameter $\beta(\Phi, \Psi)$ called *modulus-coherence* and defined as

$$\beta(\Phi, \Psi) = \max_{1 \leq i, j \leq N} \sqrt{\sum_{k=1}^N |\phi_{ki}^* \psi_{kj}|^2}, \quad (4)$$

where ϕ_{ki} and ψ_{kj} are respectively the k^{th} entries of the vectors ϕ_i and ψ_j . We then show that this parameter reaches its optimal value $\beta(\Phi, \Psi) = N^{-1/2}$ whatever the sparsity basis Ψ , for particular sensing matrices Φ including the Fourier matrix, thus providing *universal* recovery performances. It is also *efficient* as sensing matrices with fast matrix multiplication algorithms can be used, thus reducing the need in memory requirement and computational power. In Section 3, these theoretical results are confirmed numerically through an analysis of the phase transition of the ℓ_1 -minimization problem for different pairs of sensing and sparsity bases. In Section 4, we show that the spread spectrum technique remains effective in an analog setting with chirp modulation for application to realistic Fourier imaging, and illustrate these findings in the context of radio interferometry. Finally, we briefly conclude in Section 5.

Note that the spread spectrum technique was already briefly introduced by the authors for compressive sampling of pulse trains in [15], and applied to radio interferometry [16–18] and MRI [12, 13]. This paper provides theoretical foundations for the technique, both in the digital and analog settings.

2 Compressed sensing by spread spectrum

In this section, we first recall the standard recovery conditions of sparse signals randomly sampled in a bounded orthonormal system. These recovery results depend on the mutual coherence μ of the system. Hence, we study the effect of a random pre-modulation on this value and deduce the new recovery conditions for the spread spectrum technique. We finally show that the number of measurements needed to recover sparse signals becomes universal for a family of sensing matrices Φ which includes the Fourier basis.

2.1 Recovery results in a bounded orthonormal system

For the setting presented in Section 1, the theory of compressed sensing already provides sufficient conditions on the number of measurements needed to recover the vector α from the measurements \mathbf{y} by solving the

ℓ_1 -minimization problem (3) [6, 7].

Theorem 1. Let $\mathbf{A} = \Phi^* \Psi \in \mathbb{C}^{N \times N}$, $\mu = \max_{1 \leq i, j \leq N} |\langle \phi_i, \psi_j \rangle|$, $\boldsymbol{\alpha} \in \mathbb{C}^N$ be a s -sparse vector, and $\mathbf{y} = \mathbf{A}_\Omega \boldsymbol{\alpha} \in \mathbb{C}^m$. For some universal constants $C > 0$ and $\gamma > 1$, if

$$m \geq CN\mu^2 s \log^4(N), \quad (5)$$

then $\boldsymbol{\alpha}$ is the unique minimizer of the ℓ_1 -minimization problem (3) with probability at least $1 - N^{-\gamma \log^3(N)}$.

Let us acknowledge that even if the measurements are corrupted by noise or if $\boldsymbol{\alpha}$ is non-exactly sparse, the theory of compressed sensing also shows that the reconstruction obtained by solving the ℓ_1 -minimization problem remains accurate and stable:

Theorem 2. Let $\mathbf{A} = \Phi^* \Psi$, and $\mathcal{T}_s(\boldsymbol{\alpha})$ be the best s -sparse approximation of the (possibly non-sparse) vector $\boldsymbol{\alpha} \in \mathbb{C}^N$. Let the noisy measurements $\mathbf{y} = \mathbf{A}_\Omega \boldsymbol{\alpha} + \mathbf{n} \in \mathbb{C}^m$ be given with $\|\mathbf{n}\|_2 \leq \eta$. Then, if relation (5) is satisfied, the solution $\boldsymbol{\alpha}^*$ of the ℓ_1 -minimization problem

$$\boldsymbol{\alpha}^* = \arg \min_{\bar{\boldsymbol{\alpha}} \in \mathbb{C}^N} \|\bar{\boldsymbol{\alpha}}\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{A}_\Omega \bar{\boldsymbol{\alpha}}\|_2 \leq \eta, \quad (6)$$

satisfies

$$\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^*\|_2 \leq D \frac{\|\boldsymbol{\alpha} - \mathcal{T}_s(\boldsymbol{\alpha})\|_1}{s^{1/2}} + E\eta, \quad (7)$$

with probability at least $1 - N^{-\gamma \log^3(N)}$, where $D, E > 0$ and $\gamma > 1$ are some universal constants.

In the above theorems, the role of the mutual coherence μ is crucial as the number of measurements needed to reconstruct \boldsymbol{x} scales quadratically with its value. In the worst case where Φ and Ψ are identical, $\mu = 1$ and the signal \boldsymbol{x} is probed in a domain where it is also sparse. According to relation (5), the number of measurements necessary to recover \boldsymbol{x} is of order N . This result is actually very intuitive. For an accurate reconstruction of signals sampled in their sparsity domain, all the non-zero entries need to be probed. It becomes highly probable when $m \simeq N$. On the contrary, when Φ and Ψ are as incoherent as possible, i.e. $\mu = N^{-1/2}$, the energy of the sparsity basis vectors spreads equally over the sensing basis vectors. Consequently, whatever the sensing basis vector selected, one always gets information of all the sparsity basis vectors describing the signal \boldsymbol{x} , therefore reducing the need in the number of measurements. This is confirmed by relation (5) which shows that the number of measurements can be reduced to the order of s if $\mu = N^{-1/2}$. To achieve much better performance when the mutual coherence is not optimal, one would naturally try to modify the measurement process to achieve a better global incoherence. We will see in

the next section that a simple random pre-modulation is an efficient and sufficient way to achieve this goal whatever the sparsity matrix Ψ .

2.2 Pre-modulation effect on the mutual coherence

The spread spectrum technique consists of pre-modulating the signal \mathbf{x} by a wide-band signal $\mathbf{c} = (c_l)_{1 \leq l \leq N} \in \mathbb{C}^N$, with $|c_l| = 1$ and random phases, before projecting the resulting signal onto m vectors of the basis Φ . The measurement vector \mathbf{y} then still satisfies relation (1) but with

$$\mathbf{A}_\Omega = \Phi_\Omega^* \mathbf{C} \Psi \in \mathbb{C}^{m \times N}, \quad (8)$$

where the additional matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ stands for the diagonal matrix associated to the sequence \mathbf{c} .

In this setting, the matrix \mathbf{A} is still orthonormal. Therefore, the recovery condition of sparse signals sampled with the use of this matrix depends on the mutual coherence $\mu = \max_{1 \leq i, j \leq N} |\langle \phi_i, \mathbf{C} \psi_j \rangle|$. With a pre-modulation by a random Rademacher or Steinhaus sequence, Lemma 1 shows that the mutual coherence μ is essentially bounded by the modulus-coherence $\beta(\Phi, \Psi)$ defined in equation (4).

Lemma 1. *Let $\mathbf{c} \in \mathbb{C}^N$ be a random Rademacher or Steinhaus sequence and $\mathbf{C} \in \mathbb{C}^{N \times N}$ be the associated diagonal matrix. Then, the mutual coherence $\mu = \max_{1 \leq i, j \leq N} |\langle \phi_i, \mathbf{C} \psi_j \rangle|$ satisfies*

$$\mu \leq \beta(\Phi, \Psi) \sqrt{2 \log(2N^2/\epsilon)}, \quad (9)$$

with probability at least $1 - \epsilon$.

The proof of Lemma 1 relies on a simple application of the Hoeffding's inequality and the union bound.

Proof. We have $\langle \phi_i, \mathbf{C} \psi_j \rangle = \sum_{k=1}^N c_k \phi_{ki}^* \psi_{kj} = \sum_{k=1}^N c_k a_k^{ij}$, where $a_k^{ij} = \phi_{ki}^* \psi_{kj}$. An application of the Hoeffding's inequality shows that

$$\mathbb{P}(|\langle \phi_i, \mathbf{C} \psi_j \rangle| > u) \leq 2 \exp\left(-\frac{u^2}{2\|\mathbf{a}^{ij}\|_2^2}\right),$$

for all $u > 0$ and $1 \leq i, j \leq N$, with $\|\mathbf{a}^{ij}\|_2^2 = \sum_{k=1}^N |a_k^{ij}|^2$. The union bound then yields

$$\begin{aligned} \mathbb{P}(\mu > u) &\leq \sum_{1 \leq i, j \leq N} \mathbb{P}(|\langle \phi_i, \mathbf{c} \cdot \psi_j \rangle| > u) \\ &\leq 2 \sum_{1 \leq i, j \leq N} \exp\left(-\frac{u^2}{2\|\mathbf{a}^{ij}\|_2^2}\right), \end{aligned}$$

for all $u > 0$. As $\beta^2(\Phi, \Psi) = \max_{1 \leq i, j \leq N} \sum_{k=1}^N |a_k^{ij}|^2$ then $\|\mathbf{a}^{ij}\|_2^2 \leq \beta^2(\Phi, \Psi)$ for all $1 \leq i, j \leq N$, and the previous relation becomes

$$\mathbb{P}(\mu > u) \leq 2N^2 \exp\left(-\frac{u^2}{2\beta^2(\Phi, \Psi)}\right),$$

for all $u > 0$. Taking $u = \sqrt{2\beta^2(\Phi, \Psi) \log(2N^2/\epsilon)}$ terminates the proof. \square

2.3 Sparse recovery with the spread spectrum technique

Combining Theorem 1 with the previous estimate on the mutual coherence, we can state the following theorem:

Theorem 3. *Let $\mathbf{c} \in \mathbb{C}^N$ be a random Rademacher or Steinhaus sequence, with $N \geq 2$, and $\mathbf{C} \in \mathbb{C}^{N \times N}$ the associated diagonal matrix. Let $\boldsymbol{\alpha} \in \mathbb{C}^N$ be a s -sparse vector and $\mathbf{y} = \mathbf{A}_\Omega \boldsymbol{\alpha} \in \mathbb{C}^m$, with $\mathbf{A} = \Phi^* \mathbf{C} \Psi$. For $\rho < \log^3(N)$ and $C_\rho > 0$, if*

$$m \geq C_\rho N \beta^2(\Phi, \Psi) s \log^5(N), \quad (10)$$

then $\boldsymbol{\alpha}$ is the unique minimizer of the ℓ_1 -minimization problem (3) with probability at least $1 - \mathcal{O}(N^{-\rho})$.

Proof. It is straightforward to check that $\mathbf{C}^* \mathbf{C} = \mathbf{C} \mathbf{C}^* = \mathbf{I}$ where \mathbf{I} is the identity matrix. The matrix $\mathbf{A} = \Phi^* \mathbf{C} \Psi$ is thus orthonormal and Theorem 1 applies. From Lemma 1, we know that the coherence satisfies relation (9) with probability at least $1 - \epsilon$. Consequently, with probability at least $(1 - N^{-\gamma \log^3(N)}) (1 - \epsilon)$, if

$$m \geq 2C N \beta^2(\Phi, \Psi) s \log(2N^2/\epsilon) \log^4(N),$$

then $\boldsymbol{\alpha}$ is the unique minimizer of the ℓ_1 -minimization problem (3). Taking $\epsilon = N^{-\rho}$, noticing that $(1 - N^{-\gamma \log^3(N)}) (1 - N^{-\rho}) > 1 - \mathcal{O}(N^{-\rho})$, that the previous relation is satisfied if

$$m \geq 2(3 + \rho) C N \beta^2(\Phi, \Psi) s \log^5(N),$$

and taking $C_\rho = 2(3 + \rho)C$ terminates the proof. \square

Note that relation (10) also ensure the stability of the spread spectrum technique relative to noise and compressibility by combination of Theorem 2 and Lemma 1.

2.4 Universal sensing bases with ideal modulus-coherence

Theorem 3 shows that the performance of the spread spectrum technique is driven by the modulus-coherence $\beta(\Phi, \Psi)$. In general the spread spectrum technique is not universal and the number of measurements required for accurate reconstructions depends on the value of this parameter.

Definition 1. (*Universal sensing basis*) An orthonormal basis $\Phi \in \mathbb{C}^{N \times N}$ is called a universal sensing basis if all its entries ϕ_{ki} , $1 \leq k, i \leq N$, are of equal complex magnitude.

For universal sensing bases, e.g. the Fourier transform or the Hadamard transform, we have $|\phi_{ki}| = N^{-1/2}$ for all $1 \leq k, i \leq N$. It follows that $\beta(\Phi, \Psi) = N^{-1/2}$ and $\mu \simeq N^{-1/2}$, i.e. its optimal value up to a logarithmic factor, whatever the sparsity matrix considered! For such sensing matrices, the spread spectrum technique is thus a simple and efficient way to render a system incoherent independently of the sparsity matrix.

Corollary 1. (*Spread spectrum universality*) Let $\mathbf{c} \in \mathbb{C}^N$, with $N \geq 2$, be a random Rademacher or Steinhaus sequence and $\mathbf{C} \in \mathbb{C}^{N \times N}$ the associated diagonal matrix. Let $\boldsymbol{\alpha} \in \mathbb{C}^N$ be a s -sparse vector and $\mathbf{y} = \mathbf{A}_\Omega \boldsymbol{\alpha} \in \mathbb{C}^m$, with $\mathbf{A} = \Phi^* \mathbf{C} \Psi$. For $\rho < \log^3(N)$, $C_\rho > 0$, and universal sensing bases $\Phi \in \mathbb{C}^{N \times N}$, if

$$m \geq C_\rho s \log^5(N), \quad (11)$$

then $\boldsymbol{\alpha}$ is the unique minimizer of the ℓ_1 -minimization problem (3) with probability at least $1 - \mathcal{O}(N^{-\rho})$.

For universal sensing bases, the spread spectrum technique is thus *universal*: the recovery condition does not depend on the sparsity basis and the number of measurements needed to reconstruct sparse signals is optimal in the sense that it is reduced to the sparsity level s . The technique is also *efficient* as the pre-modulation only requires a sample-by-sample multiplication between \mathbf{x} and \mathbf{c} . Furthermore, fast multiplication matrix algorithms are available for several universal sensing bases such as the Fourier or Hadamard bases.

In light of Corollary 1, one can notice that sampling sparse signals in the Fourier basis is a universal encoding strategy whatever the sparsity basis Ψ - even if the original signal is itself sparse in the Fourier basis! We will confirm these results experimentally in Section 3.

2.5 Related work

Let us acknowledge that the techniques proposed in [19–23] can be related to the spread spectrum technique. In [19], the benefit of a random pre-modulation in the measurement system is already briefly suggested.

In [20], the author proposes to convolve the signal \mathbf{x} with a random waveform and randomly under-sample the result in time-domain. The random convolution is performed through a random pre-modulation in the Fourier domain and the signal thus spreads in the time-domain. In our setting, this method actually corresponds to taking Φ as the Fourier matrix and Ψ as the composition of the Fourier matrix and the initial sparsity matrix. In [21], the authors propose a technique to sample signals sparse in the Fourier domain. They first pre-modulate the signal by a random sequence, then apply a low-pass antialiasing filter, and finally sample it at low rate. Finally, random pre-modulation is also used in [22] and [23] but for dimension reduction and low dimensional embedding.

We recover similar results, albeit in a different way. We also have a more general interpretation. In particular, we proved that changing the sensing matrix from the Fourier basis to the Hadamard does not change the recovery condition (11).

3 Numerical simulations

In this section, we confirm our theoretical predictions by showing, through a numerical analysis of the phase transition of the ℓ_1 -minimization problem, that the spread spectrum technique is universal for the Fourier and Hadamard sensing bases.

3.1 Settings

For the first set of simulations, we consider the Dirac, Fourier and Haar wavelet bases as sparsity matrices Ψ and choose the Fourier basis as the sensing matrix Φ . We generate complex s -sparse signals of size $N = 1024$ with $s \in \{1, \dots, N\}$. The positions of the non-zero coefficients are chosen uniformly at random in $\{1, \dots, N\}$, their phases are set by generating a Steinhaus sequence, and their amplitudes follow a uniform distribution over $[0, 1]$. The signals are then probed according to relation (1) in the conditions of Theorem 1 or 3 and reconstructed from different number of measurements $m \in \{s, \dots, 10s\}$ by solving the ℓ_1 -minimization problem (3) with the SPGL1 toolbox¹ [24]. For each pair (m, s) , we compute the probability of recovery² over 100 simulations.

For the second set of simulations, the same protocol is applied with the same sparsity matrices but with the Hadamard basis as the sensing matrix Φ .

¹available at <http://www.cs.ubc.ca/labs/scl/spgl1>

²perfect recovery is considered if the ℓ_2 norm between the original signal \mathbf{x} and the reconstructed signal \mathbf{x}^* satisfy: $\|\mathbf{x} - \mathbf{x}^*\|_2 \leq 10^{-3}\|\mathbf{x}\|_2$

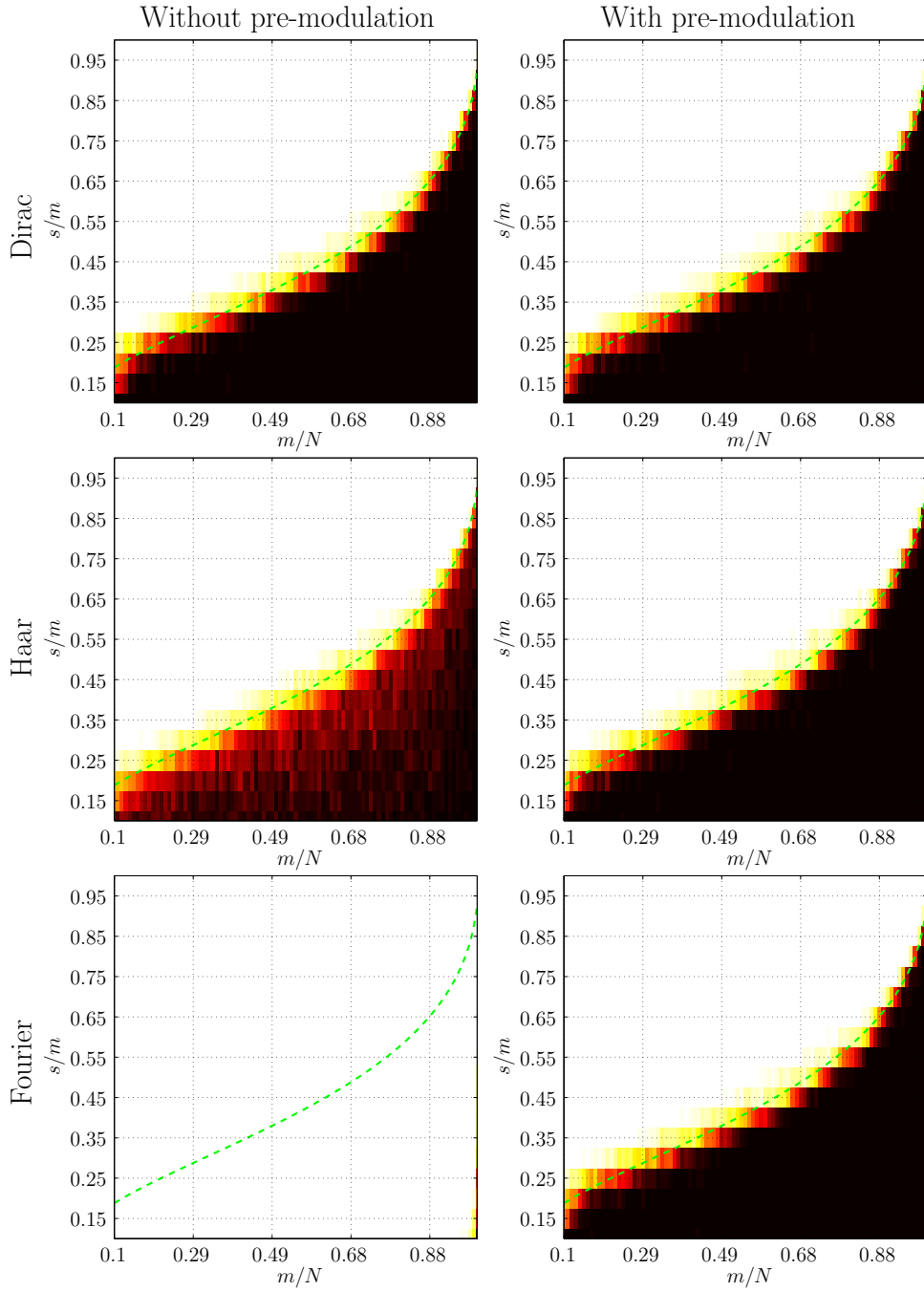


Figure 1: Phase transition of the ℓ_1 -minimization problem for different sparsity bases and random selection of **Fourier** measurements without (left panels) and with (right panels) random modulation. The sparsity bases considered are the Dirac basis (top), the Haar wavelet basis (center), and the Fourier basis (bottom). The dashed green line indicates the phase transition of Donoho-Tanner [5]. The color bar goes from white to black indicating a probability of recovery from 0 to 1.

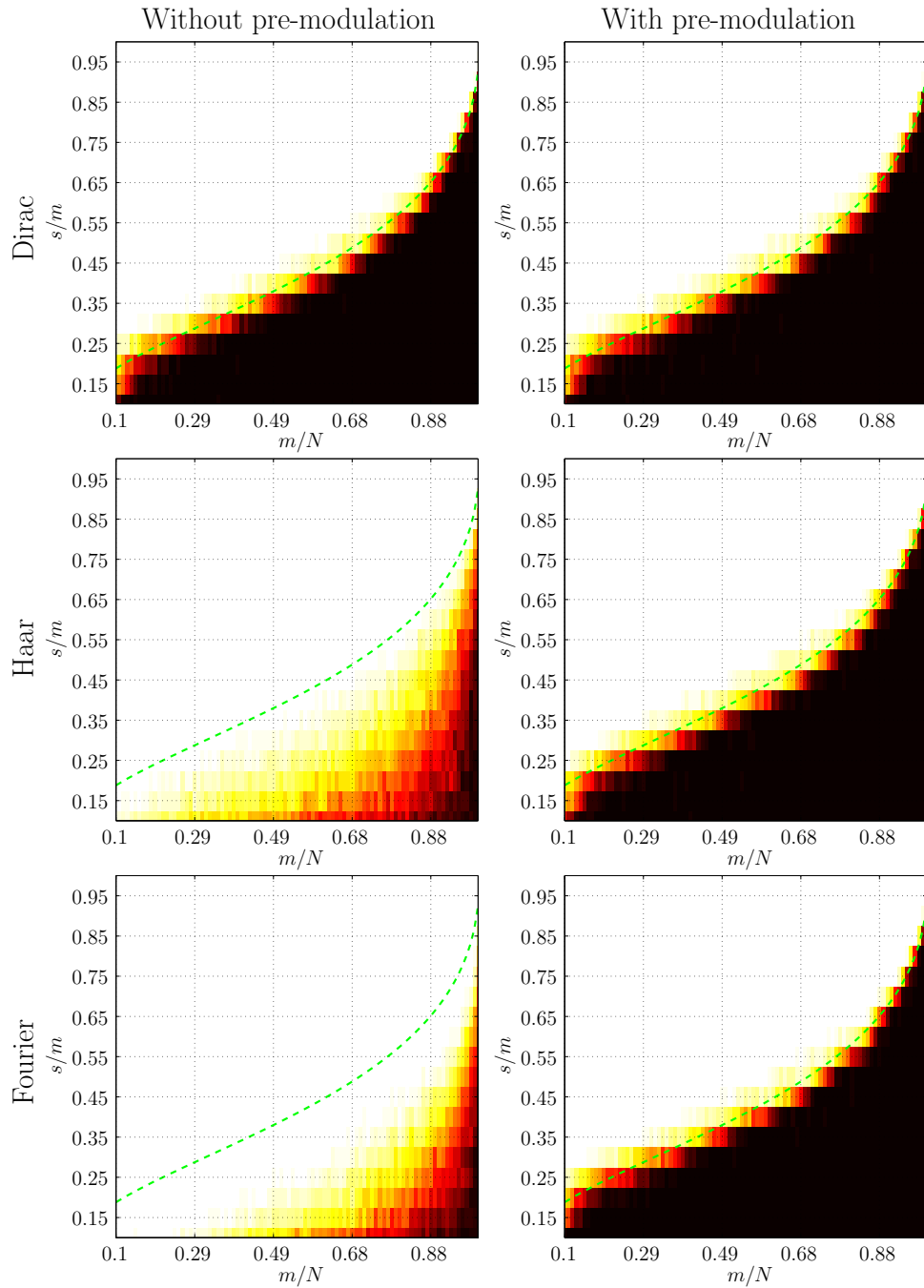


Figure 2: Phase transition of the ℓ_1 -minimization problem for different sparsity bases and random selection of **Hadamard** measurements without (left panels) and with (right panels) random modulation. The sparsity bases considered are the Dirac basis (top), the Haar wavelet basis (center), and the Fourier basis (bottom). The dashed green line indicates the phase transition of Donoho-Tanner [5]. The color bar goes from white to black indicating a probability of recovery from 0 to 1.

3.2 Results

Figure 1 shows the phase transitions of the ℓ_1 -minimization problem obtained for sparse signals in the Dirac, Haar and Fourier sparsity bases and probed in the Fourier basis with and without random pre-modulation. Figure 2 shows the same graphs but with measurements performed in the Hadamard basis.

In the absence of pre-modulation, one can note that the phase transitions depend on the mutual coherence of the system as predicted by Theorem 1. For the pairs Fourier-Dirac and Hadamard-Dirac, the mutual coherence is optimal and the experimental phase transitions match the one of Donoho-Tanner (dashed green line) [5]. For all the other cases, the coherence is not optimal and the region where the signals are recovered is much smaller. The worst case is obtained for the pair Fourier-Fourier for which $\mu = 1$.

In the presence of pre-modulation, Corollary 1 predicts that the performance should not depend on the sparsity basis and become optimal. It is confirmed by the phases transition showed on Figures 1 and 2 as they all match the phase transition of Donoho-Tanner, even for the pair Fourier-Fourier!

4 Application to realistic Fourier imaging

In this section, we discuss the application of the spread spectrum technique to realistic analog Fourier imaging such as radio interferometric imaging or MRI. Firstly, we introduce the exact sensing matrix needed to account for the analog nature of the imaging problem. Secondly, while our original theoretical results strictly hold only in a digital setting, we derive explicit performance guarantees for the analog version of the spread spectrum technique. We also confirm on the basis of simulations that the spread spectrum technique drastically enhances the quality of reconstructed signals.

4.1 Sensing model

In Fourier imaging techniques such as radio interferometry or MRI, standard measurements take the form of Fourier coefficients of the original signal probed. In light of the results of Section 2, we are in a perfect framework to apply the spread spectrum technique, apart from the analog nature of the corresponding imaging problems. In the quoted applications, the random pre-modulation is replaced by a linear chirp pre-modulation [12, 13, 16–18]. In radio interferometry, this modulation is inherently part of the acquisition process [16]. In MRI, it is easily implemented through the use of dedicated coils or RF pulses [13]. For two-dimensional signals, the linear chirp with chirp rate $w \in \mathbb{R}$ reads as a complex-valued function $c : \tau \mapsto e^{i\pi w\tau^2}$ of the spatial variable $\tau \in \mathbb{R}^2$. Note that, for high chirp rates w , this chirp shares the following important properties with the random modulation: it is a wide-band signal which does not change the norm of the

signal x , as $|c(\tau)| = 1$ whatever $\tau \in \mathbb{R}^2$.

In this setting, the complete linear relationship between the signal and the measurements is given by equation (1) with

$$\mathbf{A}_\Omega = \mathbf{F}_\Omega^* \mathbf{C} \mathbf{U} \Psi \in \mathbb{C}^{m \times N}. \quad (12)$$

In the above equation, the matrix \mathbf{U} represents an up-sampling operator needed to avoid any aliasing of the modulated signal due to a lack of sampling resolution in a digital description of the originally analog problem. The convolution in Fourier space induced by the analog modulation implies, in contrast with the digital setting studied before, that the band limit of the modulated signal is the sum of the individual band limits of the original signal and of the chirp c . In radio interferometry or MRI, signals are usually approximated by band limited functions with band limit B on their finite field of view L : $N = 2LB$. On this field of view L , the linear chirp c may be approximated by a band limited function of band limit identified by its maximum instantaneous frequency $|w|L/2$. This band limit can also be parametrized in terms of a discrete chirp rate $\bar{w} = wL^2/N$ and thus $|w|L/2 = |\bar{w}|B$. Therefore, an up-sampled grid with at least $N_{\bar{w}} = (1 + |\bar{w}|)N$ points needs to be considered and the modulated signal is correctly obtained by applying the chirp modulation on the signal after up-sampling on the $N_{\bar{w}}$ points grid³. The up-sampling operator \mathbf{U} , implemented in Fourier space by zero padding, is of size $\mathbb{C}^{N_{\bar{w}} \times N}$ and satisfies $\mathbf{U}^* \mathbf{U} = \mathbf{I} \in \mathbb{C}^{N \times N}$. Finally, the matrix $\mathbf{C} \in \mathbb{C}^{N_{\bar{w}} \times N_{\bar{w}}}$ is the diagonal matrix implementing the chirp modulation on this up-sampled grid and the matrix $\mathbf{F} = (\mathbf{f}_i)_{1 \leq i \leq N_{\bar{w}}} \in \mathbb{C}^{N_{\bar{w}} \times N_{\bar{w}}}$ stands for the discrete Fourier basis on the same grid. The indices $\Omega = \{l_1, \dots, l_m\}$ of the Fourier vectors selected to probe the signal are chosen independently and uniformly at random from $\{1, \dots, N_{\bar{w}}\}$.

4.2 Illustration

Up to the introduction of the matrix \mathbf{U} and the substitution of the linear chirp modulation for the random modulation, we are in the same setting as the one studied in Section 2. To illustrate the effectiveness of the spread spectrum technique, we consider the image of size $N = 256 \times 256$ showed in Figure 3. This image shows the radio emission associated with the encounter of a galaxy with its northern neighbor. It was acquired with the Very Large Array⁴ in New Mexico. To simulate an acquisition by a radio interferometer, this image is probed according to relation (1) with the measurement matrix (12) in the absence ($\bar{w} = 0$)

³In a full generality, natural signals are not necessarily band-limited. The spread spectrum technique can easily be adapted to this case. The sensing model should simply be modified to account for the fact that, if measurements are performed at frequencies up to a band limit B , they unavoidably contain energy of the signal up to band limit $(1 + \bar{w})B$.

⁴<http://www.vla.nrao.edu/>

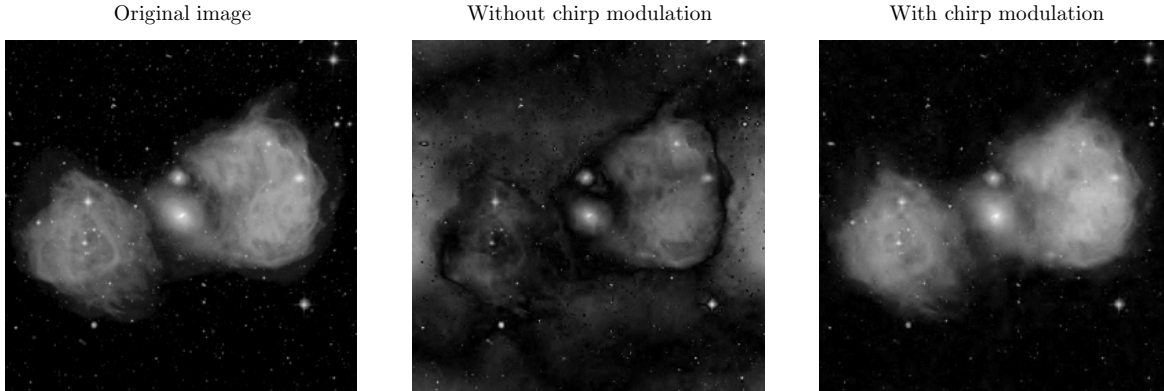


Figure 3: Left panel: Image of the giant elliptical galaxy NGC1316 (center of the image) *devouring* its small northern neighbor. The image shows the radio emission associated with this encounter superimposed on an optical image. The radio emission was imaged using the Very Large Array in New Mexico (Image courtesy of NRAO/AUI and J. M. Uson). The image size is $N = 256 \times 256$. Middle panel: Reconstructed image from $m = 0.5N$ number of measurements in the absence of chirp modulation. Right panel: Reconstructed image from $m = 0.5N$ number of measurements in the presence of a linear chirp modulation with chirp rate $\bar{\omega} = 0.1$.

and presence ($\bar{\omega} = 0.1$) of a linear chirp modulation. Independent and identically distributed Gaussian noise with zero-mean is also added to the measurements. The variance σ^2 of the noise is defined such that the input $\text{snr} = 10 \log_{10} (\sigma^2 / \Sigma_{\mathbf{x}}^2)$ is 30 dB ($\Sigma_{\mathbf{x}}$ stands for the sampled standard deviation of \mathbf{x}). It is then reconstructed from $m = 0.5N$ number of measurements by solving the ℓ_1 -minimization problem (6). The sparsity basis Ψ considered is a Daubechies-6 wavelet basis.

In the absence of linear chirp modulation, the quality of the reconstructed image is very low. However, one can already note that the fine scale structures are much better reconstructed than those at large scales. The fine details live at the small scales of the wavelet decomposition whereas the large structures live at larger scales. The small scale wavelets being more incoherent with the Fourier basis than the larger wavelets, the high frequency details are naturally better recovered.

In the presence of the linear chirp modulation, all the wavelets in Ψ become optimally incoherent with the Fourier basis thanks to the universality of the spread spectrum technique. Consequently, as one can observe on Figure 3, the low and high frequency details are well reconstructed and the image quality is drastically enhanced.

4.3 Modified recovery condition

Because of the modifications introduced in (12) to account for the analog nature of the problem, the digital theory associated the measurement matrix (8) does not explicitly apply. Nevertheless, the previous illustration shows that the spread spectrum technique is indeed still very effective in this analog setting. Actually, some explicit performance guarantees may be obtained in this setting. Let us denote $\alpha_s \in \mathbb{C}^s$ the restriction of α to its non-zero entries and assume that the phase of these entries are random. One can then prove the following *non-uniform* recovery condition:

Theorem 4. *Let $\alpha \in \mathbb{C}^N$ be a s -sparse vector, $\mathbf{y} = \mathbf{A}_\Omega \alpha \in \mathbb{C}^m$ with $\mathbf{A} = \mathbf{F}^* \mathbf{C} \mathbf{U} \Psi \in \mathbb{C}^{m \times N}$, and $\mu_{\bar{w}} = \max_{1 \leq i, j \leq N} |\langle \mathbf{f}_i, \mathbf{C} \mathbf{U} \psi_j \rangle|$. Let further assume that the phase of the non-zero entries α_s of α define a random Rademacher or Steinhaus sequence. For a universal constant $C > 0$, if*

$$m \geq C N_{\bar{w}} \mu_{\bar{w}}^2 s \log^2(6N/\epsilon), \quad (13)$$

then α is the unique minimizer of the ℓ_1 -minimization problem (3) with probability at least $1 - \epsilon$.

The proof of the theorem can be found in appendix A. Note that this *non-uniform* recovery result is weaker than the so-called *uniform* recovery conditions presented so far. The uniform recovery results of the previous section state that all sparse vectors can be recovered with high probability once a measurement matrix \mathbf{A}_Ω is chosen. This non-uniform recovery result states that sparse vectors, once fixed, can be recovered with high probability from a random choice of m sensing basis vectors. In other words, the different choices of sensing basis vectors for which the recovery fails is dependent on the original signal.

In view of this theorem, one can notice that the number of measurements needed for accurate reconstructions of sparse signals is proportional to the sparsity s times the product $N_{\bar{w}} \mu_{\bar{w}}^2$, with factors depending on the chirp rate \bar{w} . In the analog framework, two effects are actually competing. On the one hand, the mutual coherence $\mu_{\bar{w}}^2$ of the system is decreasing with the spread spectrum phenomenon, but, on the other hand, the number of accessible frequencies $N_{\bar{w}}$ that bear information is increasing linearly with the chirp

Sparsity basis	Dirac	Fourier
$N_{\bar{w}} \mu_{\bar{w}}^2$ at $\bar{w} = 0$	1.00	$1.02 \cdot 10^3$
$N_{\bar{w}} \mu_{\bar{w}}^2$ at $\bar{w} = 0.10$	2.58	$1.54 \cdot 10^1$
$N_{\bar{w}} \mu_{\bar{w}}^2$ at $\bar{w} = 0.25$	3.15	6.95
$N_{\bar{w}} \mu_{\bar{w}}^2$ at $\bar{w} = 0.50$	3.46	4.13

Table 1: Influence of a chirp modulation on $N_{\bar{w}} \mu_{\bar{w}}^2$.

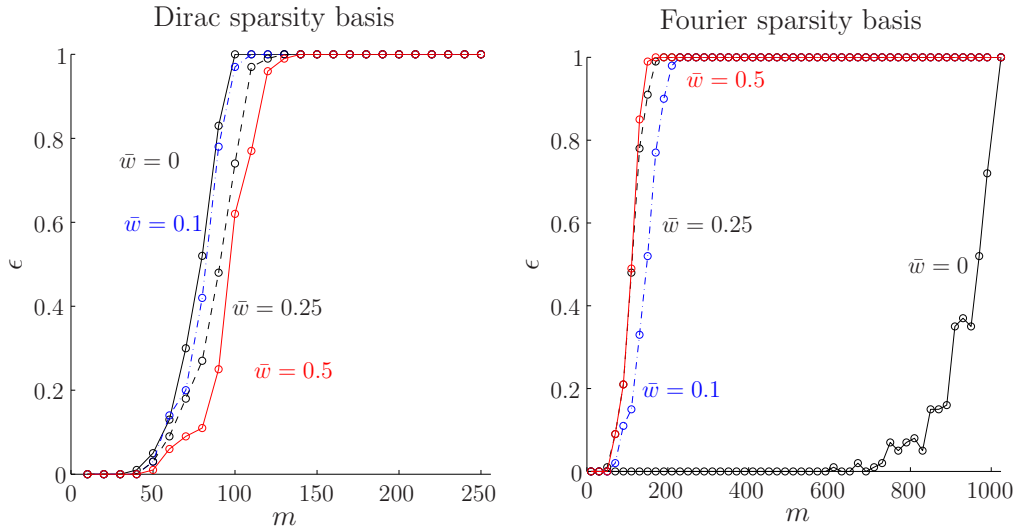


Figure 4: Probability of recovery ϵ of 10-sparse signals as a function of the number measurement m obtained with the measurement matrix (12) for two different sparsity basis: the Dirac basis (left) and the Fourier basis (right). The continuous black curve corresponds to the probability of recovery for $\bar{w} = 0$. The dot-dashed blue curve corresponds to the probability of recovery for $\bar{w} = 0.1$. The dashed black curve corresponds to the probability of recovery for $\bar{w} = 0.25$. The continuous red curve corresponds to the probability of recovery for $\bar{w} = 0.5$.

rate. The optimal recovery conditions are reached for a chirp rate \bar{w} that ensures that the product $N_{\bar{w}} \mu_{\bar{w}}^2$ is at its minimum. This minimum might change depending on the sparsity matrix, so the universality of the recovery is formally lost.

To illustrate this effect, Table 1 shows values of the product $N_{\bar{w}} \mu_{\bar{w}}^2$ for different chirp rates \bar{w} and two different sparsity matrices: the Fourier and the Dirac bases. The values are computed numerically for a size of signal $N = 1024$. In the case of the Dirac basis, one can notice that the product $N_{\bar{w}} \mu_{\bar{w}}^2$ slightly increases with the chirp rate, thus predicting that the performance should even slightly diminish in the presence of a chirp. On the contrary, the product is drastically reduced for the Fourier basis as \bar{w} increases, predicting a significant performance improvement in the presence of chirp modulation.

4.4 Experiments

To confirm the theoretical predictions of the previous section, we consider the Dirac and Fourier bases as sparsity matrices Ψ . We then generate complex s -sparse signals of size $N = 1024$ with $s = 10$. The positions of the non-zero coefficients are chosen uniformly at random in $\{1, \dots, N\}$, their signs are set by generating a Steinhaus sequence, and their amplitudes follow a uniform distribution in $[0, 1]$. The signals

are then probed according to relation (1) with the measurement matrix (12). They are then reconstructed from different number of measurements $m \in \{s, \dots, N\}$ by solving the ℓ_1 -minimization problem (3) with the SPGL1 toolbox. For each pair (m, s) , we compute the probability of recovery over 100 simulations for different chirp rate $\bar{w} \in \{0, 0.1, 0.25, 0.5\}$. Figure 4 shows the probability of recovery ϵ as a function of the number of measurements.

Firstly, in the case where Ψ is the Dirac basis, one can notice that the number of measurements needed to reach a probability of recovery of 1 slightly increases with the chirp rate \bar{w} . This is in line with the value in Table 1 and Theorem 4.

Secondly, in the case where Ψ is the Fourier basis, the performance becomes much better in the presence of a chirp. As predicted by the value in Table 1, the improvement is drastic when \bar{w} goes from 0 to 0.1 and then starts to saturate between 0.1 and 0.5.

Thirdly, according to Table 1, the product $N_{\bar{w}} \mu_{\bar{w}}^2$ is equal to 4.13 for the Fourier basis when $\bar{w} = 0.5$. This is nearly the value obtained with the Dirac basis for the same chirp rate, suggesting the same probability of recovery for the same number of measurements. Indeed, one can notice on Figure 4 that the number of measurement needed to reach a probability of recovery of 1 is around 100 in both cases.

Finally, these results also suggest that the spread spectrum technique in the modified setting is almost universal in practice. Indeed, for the perfectly incoherent pair Fourier-Dirac of sensing-sparsity bases, the number of measurements needed for perfect recovery is around 100 and this number is almost not changed in presence of the linear chirp modulation. Furthermore, for the pair Fourier-Fourier, the spread spectrum technique allows to reduce the number of measurements for perfect recovery close to this optimal value.

5 Conclusion

We have presented a compressed sensing strategy that consists of a wide bandwidth pre-modulation of the signal of interest before projection onto randomly selected vectors of an orthonormal basis. In a digital setting with a random pre-modulation, the technique was proved to be universal for sensing bases such as the Fourier or Hadamard bases, where it may be implemented efficiently. Our results were confirmed through a numerical analysis of the phase transition of the ℓ_1 -minimization problem for different pairs of sensing and sparsity bases.

The spread spectrum technique was also shown to be of great interest for realistic analog Fourier imaging. In applications such as radio interferometry and MRI, the originally digital random pre-modulation may be mimicked by an analog linear chirp. We have derived explicit performance guarantees for the analog version of

the technique with a chirp modulation, showing that it is still enhancing recovery results, though universality does not strictly hold anymore. Numerical simulations have shown that the quality of reconstructed signals is drastically enhanced in this more realistic setting, also for pairs of sensing-sparsity bases initially highly coherent, such as the Fourier-Fourier pair.

A

Let $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_N) = \mathbf{F}\mathbf{C}\mathbf{U}\Psi \in \mathbb{C}^{N_{\bar{w}} \times N}$ be the matrix defined in Section 4 and S be the support of the signal $\boldsymbol{\alpha}$: $S = \{i : |\alpha_i| > 0\}$. We denote $\mathbf{A}_{\Omega S} \in \mathbb{C}^{m \times s}$ the matrix restricted to the columns of \mathbf{A}_{Ω} indexed by S .

From Propositions 7.1 and 7.2 in [7], which are general non-uniform recovery results that still apply in our case, one can notice that we only need to provide an estimate on the norm of the operator $\|\mathbf{A}_{\Omega S}^* \mathbf{A}_{\Omega S} - \mathbf{I}\|$ and on the internal coherence $\nu = \max_{k \neq l} |\langle \mathbf{a}_k, \mathbf{a}_l \rangle|$ of \mathbf{A}_{Ω} to prove Theorem 4. The following lemma gives an estimate of $\|\mathbf{A}_{\Omega S}^* \mathbf{A}_{\Omega S} - \mathbf{I}\|$.

Lemma 2. *Let $\mathbf{A} = \mathbf{F}^* \mathbf{C} \mathbf{U} \Psi \in \mathbb{C}^{N_{\bar{w}} \times N}$ and $\mu_{\bar{w}} = \max_{ij} |a_{ij}| = \max_{ij} |\langle \mathbf{f}_i, \mathbf{C} \mathbf{U} \psi_j \rangle|$. Let the indices $\Omega = \{l_1, \dots, l_m\}$ of the selected measurement vectors be chosen independently and uniformly at random from $\{1, \dots, N_{\bar{w}}\}$. Let $\delta \in (0, 1/2]$ and suppose that $s \geq 2$. Then with probability at least*

$$1 - 2^{3/4} s \exp\left(-\frac{m\delta^2}{CN_{\bar{w}}\mu_{\bar{w}}^2 s}\right), \quad (14)$$

where $C = 9 + \sqrt{17}$, the normalized matrix $\tilde{\mathbf{A}}_{\Omega} = \sqrt{\frac{N_{\bar{w}}}{m}} \mathbf{A}_{\Omega}$ satisfies $\|\tilde{\mathbf{A}}_{\Omega S}^* \tilde{\mathbf{A}}_{\Omega S} - \mathbf{I}\| \leq \delta$.

Proof. The proof is based on the methods used to show Theorem 7.3 in [7]. We have $\mathbf{A} = \mathbf{F}^* \mathbf{C} \mathbf{U} \Psi \in \mathbb{C}^{N_{\bar{w}} \times N}$ where \mathbf{F} , \mathbf{C} and Ψ are orthonormal, and $\mathbf{U}^* \mathbf{U} = \mathbf{I} \in \mathbb{C}^{N \times N}$. We denote $\mathbf{Y} = \tilde{\mathbf{A}}_{\Omega S}^* \tilde{\mathbf{A}}_{\Omega S} - \mathbf{I} = \sum_{l=1}^m \tilde{\mathbf{r}}_l^* \tilde{\mathbf{r}}_l - \mathbf{I} \in \mathbb{C}^{s \times s}$ where $\tilde{\mathbf{r}}_l \in \mathbb{C}^{1 \times s}$ is a row-vector of $\tilde{\mathbf{A}}_{\Omega S}$. By independence of the selected indices, the row-vectors $\tilde{\mathbf{r}}_l$ are i.i.d. random vectors and we have

$$\begin{aligned} \mathbb{E} [\tilde{\mathbf{r}}_l^* \tilde{\mathbf{r}}_l]_{j,k} &= \frac{N_{\bar{w}}}{m} \mathbb{E} [(\overline{\mathbf{f}_l^* \mathbf{C} \mathbf{U} \psi_j}) (\mathbf{f}_l \mathbf{C} \mathbf{U} \psi_k)] \\ &= \frac{N_{\bar{w}}}{m} \sum_{i=1}^{N_{\bar{w}}} \left[\frac{1}{N_{\bar{w}}} (\overline{\mathbf{f}_i^* \mathbf{C} \mathbf{U} \psi_j}) (\mathbf{f}_i^* \mathbf{C} \mathbf{U} \psi_k) \right] \\ &= \frac{1}{m} \langle \mathbf{F}^* \mathbf{C} \mathbf{U} \psi_j, \mathbf{F}^* \mathbf{C} \mathbf{U} \psi_k \rangle \\ &= \frac{1}{m} \psi_j^* \mathbf{U}^* \mathbf{C}^* \mathbf{F} \mathbf{F}^* \mathbf{C} \mathbf{U} \psi_k = \frac{1}{m} \delta_{jk}. \end{aligned}$$

Then $\mathbb{E} [\tilde{\mathbf{r}}_l^* \tilde{\mathbf{r}}_l] = m^{-1} \mathbf{I}$. As in [7], we can now use a symmetrization technique to bound the expected value

of the norm of \mathbf{Y} (Lemma 6.7, [7]). Let $(\beta_1, \dots, \beta_l)$ be a Rademacher sequence and $p \geq 2$, then:

$$\begin{aligned} E_p &:= \mathbb{E} \|\mathbf{Y}\|^p = \mathbb{E} \left\| \sum_{l=1}^m [\tilde{\mathbf{r}}_l^* \tilde{\mathbf{r}}_l - \mathbb{E}(\tilde{\mathbf{r}}_l^* \tilde{\mathbf{r}}_l)] \right\|^p \\ &\leq 2^p \mathbb{E} \left\| \sum_{l=1}^m \beta_l \tilde{\mathbf{r}}_l^* \tilde{\mathbf{r}}_l \right\|^p. \end{aligned}$$

Noticing that $\tilde{\mathbf{A}}_{\Omega S}$ has at most rank s , using Fubini's theorem and Rudelson's lemma (see Lemma 6.18, [7]) conditional on $(\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_m)$ yields

$$\begin{aligned} E_p &\leq 2^{3/4+p} s \left(\frac{p}{e}\right)^{p/2} \mathbb{E} \left[\|\tilde{\mathbf{A}}_{\Omega S}\|^p \max_{1 \leq l \leq m} \|\tilde{\mathbf{r}}_l\|_2^p \right] \\ &\leq 2^{3/4+p} s \left(\frac{p}{e}\right)^{p/2} \sqrt{\mathbb{E} \|\tilde{\mathbf{A}}_{\Omega S}^* \tilde{\mathbf{A}}_{\Omega S}\|^p \mathbb{E} \left[\max_{1 \leq l \leq m} \|\tilde{\mathbf{r}}_l\|_2^{2p} \right]}. \end{aligned}$$

The previous equation is identical to equation (7.6) in [7]. Using the bound $\mathbb{E} \left[\max_{1 \leq l \leq m} \|\tilde{\mathbf{r}}_l\|_2^{2p} \right] \leq (N_{\bar{w}}/m)^p \mu_{\bar{w}}^{2p} s^p$ and following the same remaining steps than for the proof of Proposition 7.3. in [7] terminates ours. \square

We can now give an estimate on the internal coherence ν using the same method than for the proof of corollary 7.4 in [7].

Corollary 2. *Let $A = \mathbf{F}^* \mathbf{C} \mathbf{U} \Psi \in \mathbb{C}^{N_{\bar{w}} \times N}$ and $\mu = \max_{ij} |a_{ij}| \geq N_{\bar{w}}^{-1/2}$. Let the indices $\Omega = \{l_1, \dots, l_m\}$ of the selected measurement vectors be chosen independently and uniformly at random from $\{1, \dots, N_{\bar{w}}\}$. Then the internal coherence ν of the normalized matrix $\tilde{\mathbf{A}}_{\Omega} = \sqrt{\frac{N_{\bar{w}}}{m}} \mathbf{A}_{\Omega}$ satisfies*

$$\nu \leq \sqrt{\frac{2C N_{\bar{w}} \mu_{\bar{w}}^2 \log(2^{3/4} N^2 / \epsilon)}{m}},$$

with probability at least $1 - \epsilon$, provided that the right hand side is at most $1/2$ ($C = 9 + \sqrt{17}$).

Finally Theorem 4 is obtained by applying the exact same method than in Section 7.3 of [7] with the use of the last lemma and corollary.

References

1. Candès EJ, Romberg J, Tao T: **Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information.** *IEEE Trans. Inf. Theory* 2006, **52**:489–509.
2. Donoho DL: **Compressed sensing.** *IEEE Trans. Inf. Theory* 2006, **52**:1289–1306.
3. Baraniuk R: **Compressive Sensing.** *IEEE Signal Process. Mag.* 2007, **24**:118–121.

4. Baraniuk R: **Near optimal signal recovery from random projections: universal encoding strategies.** *IEEE Trans. Inf. Theory* 2006, **52**:5406–5425.
5. Donoho DL, Tanner J: **Counting faces of randomly-projected polytopes when the projection radically lowers dimension.** *J. Amer. Math. Soc.* 2009, **22**:1–53.
6. Candès EJ, Romberg J: **Sparsity and Incoherence in Compressive Sampling.** *Inverse Problems* 2007, **23**:969–985.
7. Rauhut H: **Compressive Sensing and Structured Random Matrices.** In *Theoretical Foundations and Numerical Methods for Sparse Recovery*. Edited by Fornasier M, De Gruyter 2010:1–92. [Radon Series on Computational and Applied Mathematics, vol. 9].
8. Högbom JA: **Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines.** *Astron. Astrophys. Suppl.* 1974, **15**:417–426.
9. Cornwell TJ, Evans KF: **A simple maximum entropy deconvolution algorithm.** *Astron. Astrophys.* 1985, **143**:77–83.
10. Wiaux Y, Jacques L, Puy G, Scaife AMM, Vandergheynst P: **Compressed sensing imaging techniques for radio interferometry.** *Mon. Not. R. Astron. Soc.* 2009, **395**:1733–1742.
11. Lustig M, Donoho D, Pauly JM: **Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging.** *Magn. Reson. Med.* 2007, **58**:1182–1195.
12. Wiaux Y, Puy G, Gruetter R, Thiran JP, de Ville DV, Vandergheynst P: **Spread spectrum for compressed sensing techniques in magnetic resonance imaging.** In *Proc. IEEE Int. Sym. on Biomed. Imaging, Volume CFP10BIS-CDR* 2010:756–759.
13. Puy G, Marques J, Gruetter R, Thiran JP, de Ville DV, Vandergheynst P, Wiaux Y: **Accelerated MR imaging with spread spectrum encoding.** In *Proc. Intl. Soc. Mag. Reson. Med., Volume 19* 2011:2808.
14. Larson PEZ, Hu S, Lustig M, Kerr AB, Nelson SJ, Kurhanewicz J, Pauly JM, Vigneron DB: **Fast dynamic 3D MR spectroscopic imaging with compressed sensing and multiband excitation pulses for hyperpolarized (^{13}C) studies.** *Magn. Reson. Med.* 2011, **65**:610–619.

15. Naini FM, Gribonval R, Jacques L, Vandergheynst P: **Compressive sampling of pulse trains: Spread the spectrum!** In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Process.* 2009:2877–2880.
16. Wiaux Y, Puy G, Boursier Y, Vandergheynst P: **Spread spectrum for imaging techniques in radio interferometry.** *Mon. Not. R. Astron. Soc.* 2009, **400**:1029–1038.
17. Wiaux Y, Puy G, Boursier Y, Vandergheynst P: **Compressed sensing for radio interferometry: spread spectrum imaging techniques.** In *Proc. SPIE Conf. WAVELET XIII, Volume 7446* 2009:7446J.
18. Puy G, Wiaux Y, Gruetter R, Thiran JP, de Ville DV, Vandergheynst P: **Spread spectrum for interferometric and magnetic resonance imaging.** In *Proc. IEEE Int. Conf. on Acoustic, Speech and Signal Process.* 2010:2802–2805.
19. Do T, Tran T, Gan L: **Fast compressive sampling with structurally random matrices.** In *Proc. IEEE Int. Conf. on Acoustic, Speech and Signal Process.* 2008:3369 – 3372.
20. Romberg J: **Compressive sensing by random convolution.** *SIAM J. Imaging Sciences* 2009, **02**:1098–1128.
21. Tropp JA, Laska JN, Duarte MF, Romberg JK, Barianuk RG: **Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals.** *IEEE Trans. Inf. Theory* 2010, **56**:520–544.
22. Krahmer F, Ward R: **New and Improved JohnsonLindenstrauss Embeddings via the Restricted Isometry Property.** *SIAM J. on Mathematical Analysis* 2011, **43**:1269–1281.
23. Tropp JA: **Improved analysis of the subsampled randomized Hadamard transform.** *Adv. Adapt. Data Anal.* 2011. [Special issue, “Sparse Representation of Data and Images,” to appear].
24. van den Berg E, Friedlander MP: **Probing the Pareto frontier for basis pursuit solutions.** *SIAM Journal on Scientific Computing* 2008, **31**:890–912.