Practical band-limited extrapolation relying on Slepian series and compressive sampling

Laurent Gosse

Practical band-limited extrapolation relying on Slepian series and compressive sampling

Laurent Gosse

IAC–CNR "Mauro Picone" (sezione di Bari)
Via Amendola 122/I - 70126 Bari, Italy

Abstract

We consider a rather simple algorithm to address the fascinating field of numerical extrapolation of (analytic) band-limited functions. It relies on two main elements: namely, the lower frequencies are treated by projecting the known part of the signal to be extended onto the space generated by "Prolate Spheroidal Wave Functions" (PSWF, as originally proposed by Slepian), whereas the higher ones can be handled by the recent so–called "Compressive Sampling" (CS, proposed by Candès) algorithms which are independent of the largeness of the bandwidth. Slepian functions are recalled and their numerical computation is explained in full detail whereas Compressive Sampling techniques are summarized together with a recent iterative algorithm which has been proved to work efficiently on so–called "compressible signals" which appear to match rather well the class of smooth bandlimited functions. Numerical results are displayed for both numerical techniques and the accuracy of the process consisting in putting them altogether is studied for several test-signals.

Key words: Band-limited extrapolation, Prolate spheroidal wave functions, Slepian series, analytic continuation, finite Fourier transform, compressive sampling, sparse and compressible signals recovery.
1991 MSC: 42A16; 41A58; 65D15; 68P30

1 Introduction

Band-limited functions are lying at the very core of various questions arising in many areas of application, especially signal processing, information and approximation theories. Let us begin by recalling that any function f belonging to $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ is said to be *band-limited*

* Corresponding Author
Email address: l.gosse@area.ba.cnr.it (Laurent Gosse).

if there exists a positive number σ such that:

$$\forall t \in \mathbb{R}, \qquad f(t) = \frac{1}{2\pi} \int_{-\sigma}^{\sigma} \hat{f}(\xi) \exp(it\xi).d\xi,$$

the Fourier transform $\hat{f}$ being defined according to the functional space [1] of f. The number $\sigma \in \mathbb{R}^+$ is called the *bandwidth* of f; it is not an obvious task to decide whether or not real-life signals are endowed with a finite bandwidth, see [30]. The classical Paley-Wiener theorem states that any such function belonging to $L^2(\mathbb{R})$ extends to the whole complex plane as an entire function of exponential type; more precisely, there holds (see *e.g.* [43])

$$\forall z \in \mathbb{C}, \qquad |f(z)| \leq \sup_{t \in \mathbb{R}} |f(t)| \exp(\sigma \Im(z)).$$

As a consequence of analytic continuation theory for functions of complex variable, the knowledge of such a function restricted to any arbitrary interval of $\mathbb{R}$ allows to deduce all the remaining values corresponding to any $z \in \mathbb{C}$.

Despite this very neat theory, numerical extrapolation of band-limited functions remains a delicate practical problem having quite a long history; several survey papers are available, like for instance [16,24,35,43]. Two main approaches coexist relying on different algorithmic methodologies: iterative methods generally follow the classical Gerchberg-Papoulis algorithm [28,12], but recently, a renewed interest grew around non-iterative methods relying on the so–called Slepian series [32,33]. Iterative routines for *scale-limited extrapolation* have been developed around similar ideas and led to a so–called "generalized Gerchberg-Papoulis algorithm": see [22,23,37]. In the present text, we shall not deal with iterative extrapolation routines for band-limited functions; instead we refer to existing papers [12,16,28,35,40] and references therein. It turns out that, for a long time, iterative methods have gained from the fact that extrapolation algorithms based on Slepian's Prolate functions (outlined in the paper [31], page 388) were far beyond practical computing possibilities. However, the situation has changed dramatically nowadays and several papers have been recently published to present feasible algorithms for the computation of Slepian series: see for instance [3,15,17,18,25,36,38]. We aim at taking advantage of them in order to propose a realizable and efficient numerical scheme to extrapolate accurately smooth band-limited functions.

Compressive Sensing differs greatly from the aforementioned algorithms in the sense that it doesn't rely on the straightforward projection of a recorded signal onto the specific basis of some Hilbert functional space; its way of proceeding goes more like *correlating it with a collection of randomly chosen but perfectly known test-signals*. Hence Compressed Sensing is less naturally an extrapolation method because it needs unknown values to be located on the complementary of a random set of points. In most cases, unknown values are located in a connected domain (or two of them, each one around the edges of some observations interval). At this level, what saves the day is that CS can reconstruct signals which are strongly under-sampled: hence in many cases one can manage a random set of observations not seeing the whole interval containing the unknown values. For instance, in case there is a collection of 100 measures and it is to deduce the 10 ones located on the right, one needs for CS to work one particular random sequence of measures taken inside the whole set of 110 values but seeing only the 100 first ones. This can easily happen when one needs to measure only a very small subset of values, the smallness being directly related to the sparse character of (for instance) the Fourier transform of the signal to extrapolate. This is in this sense that we propose here to exploit CS algorithms to handle efficiently extrapolation problems involving high frequencies which typically aren't easy to process by the standard Gerchberg-Papoulis method; to the best of our knowledge, this idea seems to be new.

[1] one could even consider f in a space of tempered distributions $\mathcal{S}'$ on $\mathbb{R}$, like for instance the Zakaï class [41,21,4] but this is out of the scope of the present paper.

The paper is organized as follows. In §2.1, we recall useful theoretical considerations concerning the derivation of Prolate functions and present an extrapolation algorithm following Slepian [31]. An illustration of these functions together with their L^2 norm is given in §2.1.3. §2.2 is devoted to surveying some important results in the theory of Compressed Sensing; especially the so–called Restricted Isometry Property, a crucial feature of certain random matrices, is recalled in §2.2.1 and recovery theorems are stated in §2.2.2. §2.2.3 shows a numerical illustration of the perfect recovery phenomenon for a sparse signal. In §3, we explain in detail how to set up a practical algorithm to conduct extrapolation of any band-limited (preferably low frequency) signal relying on the projection of the Paley-Wiener subspace of $L^2(\mathbb{R})$ generated by Prolate functions. Coefficients are computed using Gauss-Legendre quadrature. Several numerical test-cases are shown in §3.3. Extrapolation of higher frequencies relying on Compressed Sensing techniques are dealt with in §4. In §4.1, we introduce the useful class of "compressible signals" and state recovery results for them. In §4.2, we present a simple but efficient iterative recovery algorithm called ROMP endowed with good stability properties. The issue of finding convenient random sequences of observations is discussed in §4.3 where also a comparison between CS and standard Gerchberg-Papoulis recovery methods is made (see Remark 4). Numerical test-cases are presented in §4. In §5, we display the numerical outcome of an overall extrapolation process where the signal to be treated is first split between a "trend" and a "fluctuation" being processed by Prolate functions and Compressed Sensing respectively; two test-cases are considered. Results are to be checked on Figures 12 and 13. Conclusive remarks are given in §6.

2 "Adiabatic decoupling" between lower and higher frequencies

Part of our methodology relies on the assumption that a quite general class of interesting signals allow for being split between a "trend" corresponding to lower frequencies and a "fluctuation" which accounts for higher frequencies. Moreover, we shall also assume in all what follows that higher frequencies don't have much effect on low ones during a limited amount of time. Put in another way, extrapolating on an interval of moderate length permits to uncouple between low and high frequencies inside the signal. This is quite appealing because modern numerical strategies to handle economically low and high frequencies are quite different and such a decoupling opens the way of taking the best from each of them to treat both parts of the observed signal. We used the terminology "adiabatic decoupling" in a loose sense to underline the fact that high frequencies are separated from the rest of the signal and that they are not supposed to have a sensible effect on the lower ones during a limited amount of time; this is strongly reminiscent of the adiabatic theory as encountered in quantum mechanics, for instance in the context of molecular dynamics where light particles are driven by the movement of heavier ones (consult *e.g.* [34] for details in this direction).

Moreover, this framework is also well suited for other multiscale phenomena being of a given size d, but which are dominated by only the $c \ll d$ trend's Fourier coefficients (see *e.g.* [1,44]). In this paper, we shall limit ourselves to check whether or not such a frequency splitting is appropriate or not based on the elementary signals correlation. More precisely, assume that we know a signal $t \mapsto f(t) \in \mathbb{R}$ inside some observation interval $[a, b]$ and there is a possible decomposition $f = f_{lo} + f_{hi}$ such that the convex hull of the support of $\hat{f}_{lo}$ and the support of $\hat{f}_{hi}$ are disjoint. Then this splitting will be admissible if moreover,

$$\kappa := \frac{\int_a^b f_{lo}(t) f_{hi}(t).dt}{\|f_{lo}\|_{L^2(a,b)} \|f_{hi}\|_{L^2(a,b)}} \tag{1}$$

is lower that some threshold value $\varepsilon_0 \ll 1$ in modulus. By Cauchy-Schwarz inequality, we have that $\kappa \in [-1, 1]$ with $|\kappa| = 1$ if and only if there exists $\alpha \in \mathbb{R}$ such that $f_{lo} = \alpha f_{hi}$.

2.1 The theory of Prolate Spheroidal Wave Functions (PSWF)

We begin by presenting a theoretical framework adapted to the processing of low frequencies.

2.1.1 Several equivalent definitions

We now briefly resume the derivation and the main properties of Slepian's Prolate functions, following various survey articles, see [25,19,24,31,36,38]. Originally, these functions were sought as possible real-valued solutions of the maximum concentration problem for bandlimited functions raised by Shannon himself (see [24], first page): for any $T \in \mathbb{R}^+$, consider the highest value of

$$0 \leq \alpha(T)^2 = \frac{\int_{-\frac{T}{2}}^{\frac{T}{2}} |f(t)|^2 dt}{\int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi}.$$

Assuming f to be σ-bandlimited, this expression rewrites as: (* denotes complex conjugate)

$$\begin{aligned}\alpha(T)^2 &= \frac{\int_{-\frac{T}{2}}^{\frac{T}{2}} dt \int_{-\sigma}^{\sigma} \hat{f}(\xi)\exp(it\xi)d\xi \int_{-\sigma}^{\sigma} \hat{f}(\xi')^* \exp(it\xi')d\xi'}{\int_{-\sigma}^{\sigma} |\hat{f}(\xi)|^2 d\xi} \\ &= \frac{\int_{-\sigma}^{\sigma} d\xi \int_{-\sigma}^{\sigma} \frac{\sin T(\xi-\xi')}{\pi(\xi-\xi')} \hat{f}(\xi)\hat{f}(\xi')^* d\xi'}{\int_{-\sigma}^{\sigma} \hat{f}(\xi)\hat{f}(\xi)^* d\xi}.\end{aligned}$$

The last step is completed by integration in the t variable. A maximizing f has to satisfy the following Fredholm equation of the second kind:

$$\forall \xi \in [-\sigma,\sigma], \qquad \int_{-\sigma}^{\sigma} \frac{\sin T(\xi-\xi')}{\pi(\xi-\xi')} \hat{f}(\xi')d\xi' - \alpha(T)^2 \hat{f}(\xi) = 0.$$

After easy changes of variables, this reduces to the well-known eigenvalue problem:

$$\forall \xi \in [-1,1], \qquad \int_{-1}^{1} \frac{\sin c(\xi-\xi')}{\pi(\xi-\xi')} \psi(\xi')d\xi' = \lambda\psi(\xi), \quad \lambda \geq 0,\ c = T\sigma. \tag{2}$$

This presentation clearly shows that Prolate functions are indeed objects defined in the Fourier domain, even if many authors use the generic variables x or t to write them down, possibly suggesting the opposite. As the "sinc" kernel is symmetric positive definite, the integral operator is compact and equation (2) has countably many solutions λ_n, ψ_n, $n \in \mathbb{N}$, which satisfy furthermore:

- $\lambda_n \geq 0$, $\lambda_0 \geq \lambda_1 \geq ... \geq \lambda_n \geq ...$ and $\lambda_n \to 0$ as $n \to +\infty$,
- real-valued eigenfunctions ψ_n constitute a complete orthogonal system of $L^2(-1,1)$.

We stress that besides the index n, eigenvalues and eigenfunctions λ_n, ψ_n do heavily depend on the parameter c. At this point, it has been qualified as a "lucky accident" by Slepian himself that these Prolate functions ψ_n also furnish solutions bounded on $(-1,1)$ of the

eigenvalue problem for the following differential operator:

$$\forall x \in [-1,1], \qquad -\frac{d}{dx}\left((1-x^2)\frac{d\psi}{dx}\right) + c^2x^2\psi = \chi\psi, \quad \chi > 0. \tag{3}$$

This fact is of great utility when it comes to computing numerically the functions ψ_n as discretizing directly the convolution equation (2) yields a matrix which coefficients decay very slowly to zero; hence standard algorithms won't be able to furnish accurately its eigenvalues and eigenvectors. If one plugs $c = 0$ inside (3), one obtains the differential equation leading to the orthogonal polynomial basis of Legendre on $[-1,1]$ with $\chi_n = n(n+1)$. For this reason, Prolate functions with $c > 0$ are sometimes seen as a "bandlimited generalization of Legendre polynomials"; one has to be careful with this terminology as the Legendre basis is not primarily defined in the Fourier domain.

2.1.2 Orthogonal bases of both $L^2(-1,1)$ and $L^2(\mathbb{R})$

Another amazing property is that, contrary to Legendre polynomials, Prolate functions can be easily extended so as to furnish an orthogonal system in $L^2(\mathbb{R})$. Following Slepian, we define for any $n \in \mathbb{N}$:

$$\forall |\xi| > 1, \qquad \psi_n(\xi) = \frac{1}{\lambda_n}\int_{-1}^{1} \frac{\sin c(\xi - \xi')}{\pi(\xi - \xi')}\psi_n(\xi')d\xi'.$$

They can be normalized so as to satisfy:

$$\int_{\mathbb{R}} \psi_m(\xi)\psi_n(\xi)d\xi = \delta_{m,n}, \qquad \int_{-1}^{1} \psi_m(\xi)\psi_n(\xi)d\xi = \lambda_n\delta_{m,n}, \tag{4}$$

with $\delta_{m,n}$ the Kronecker symbol. With such a normalization, eigenvalues can be interpreted as the fraction of the total mass which is concentrated in the interval $[-1,1]$. Eigenfunctions constitute therefore an orthonormal basis of a subset of $L^2(\mathbb{R})$ called the Paley-Wiener space, constituted of c-bandlimited functions endowed with a finite $L^2(\mathbb{R})$ norm. It has been believed for long time that Slepian's functions were the unique ones to possess this surprising property: another example of such a system has been found only recently in [42].

A property which reveals itself of crucial importance for extrapolation goes as follows: any Prolate function ψ_n also satisfies

$$\forall t \in \mathbb{R}, \qquad \int_{-1}^{1} \exp(ict\xi)\psi_n(\xi)d\xi = \mu_n\psi_n(t), \quad \mu_n \in \mathbb{C}. \tag{5}$$

Besides emphasizing the bandlimited character of these functions, this feature expresses that, up to a change of scale, the function ψ_n has the same shape as its Fourier transform. The values of μ_n and λ_n are related through:

$$\lambda_n(c) = \frac{c}{2\pi}|\mu_n(c)|^2. \tag{6}$$

Concerning extrapolation purposes, the expression (5) opens the way to completing the program outlined by Slepian in [31]. Namely, it allows for expressing the value of $\psi_n(t)$, for any $t \in \mathbb{R}$, relying only on the knowledge of this function restricted in the interval $[-1,1]$, restriction that can be computed accurately thanks to the eigenvalue problem for the differential operator given in (3). So, as a consequence of this, since we know that ψ_n constitute an orthogonal system of both $L^2(-1,1)$ and of the subset of $L^2(\mathbb{R})$ made of c-bandlimited functions, a practical extrapolation routine for a c-bandlimited function f known only in $[-1,1]$ will work as follows:

- compute as accurately as possible λ_n and ψ_n restricted to $[-1,1]$,
- express f in the orthogonal basis made of ψ_n through the formula:

$$\forall t \in [-1,1], \qquad f(t) = \sum_{n\in\mathbb{N}} f_n \psi_n(t), \quad f_n = \frac{1}{\lambda_n} \int_{-1}^{1} f(\tau)\psi_n(\tau).d\tau. \tag{7}$$

- extend the definition of ψ_n by means of (5) to any other interval of $\mathbb{R}$,
- the extrapolation $\tilde{f}$ of f is then given by:

$$\forall t \in]-\infty,-1[\cup]1,+\infty[, \qquad \tilde{f}(t) = \sum_{n\in\mathbb{N}} f_n \psi_n(t) = \sum_{n\in\mathbb{N}} \frac{f_n}{\mu_n} \int_{-1}^{1} \exp(ict\xi)\psi_n(\xi)d\xi. \tag{8}$$

The passage from $[-1,1]$ to $\mathbb{R}$ could be also achieved by means of the formula (2) which may look more interesting because in this last case, only the nonnegative λ_n values are needed; however, we found that, relying on the algorithms proposed in [38], the computing cost of the μ_n values was equivalent to the one of the λ_n. Therefore, the aforementioned extrapolation algorithm is completely operational as soon as one has reliable numerical algorithms to compute the Prolate functions ψ_n, the complex values μ_n and the scalar products f_n for any c-bandlimited function f known on $[-1,1]$ only.

2.1.3 Numerical examples

We shall present some reliable algorithms in §3 of the present paper. Before that, we display some numerical simulations to illustrate the behavior of both Prolate functions and their corresponding eigenvalues. On the left of Fig. 1, we display the first even/odd Prolate functions on the interval $[-1,1]$ for the bandwidth parameter $c = 7$. For this moderate value,

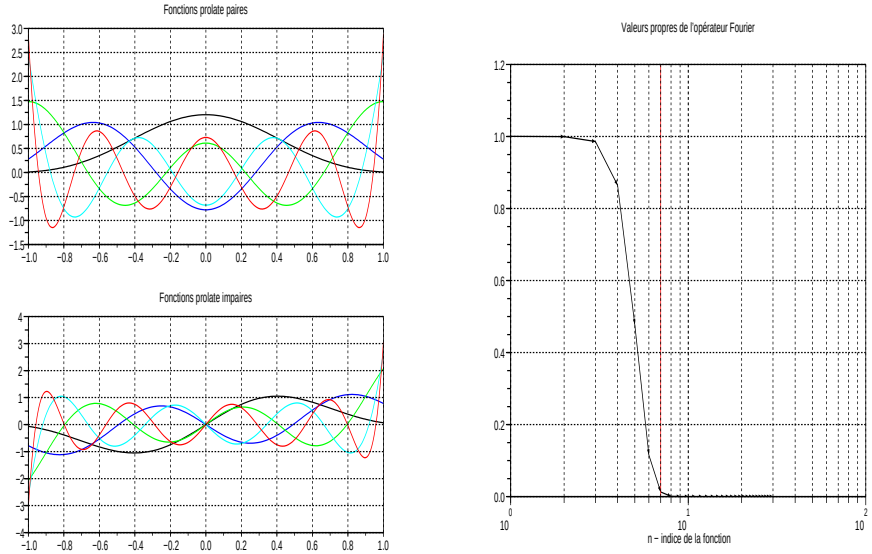

Fig. 1. Slepian functions (left) with $c = 7$ and repartition of eigenvalues λ_n (right).

we observe that the first eigenfunctions, besides looking rather similar to ordinary Legendre polynomials, are quite spread across the interval $[-1,1]$. Their absolute values grow strongly when approaching the edges ± 1. The corresponding eigenvalues λ_n are arranged like a step function (which explains why trying to solve directly a discretized version of (2) cannot be considered a good way of deriving Slepian functions) with a sharp transition zone between zero and one. It is a well-known fact that this transition occurs around the value n^* such that

$\lambda_{n^*} \simeq c$; accordingly, this value of c has been drawn with a vertical red line on both figures 1 and 2. Figure 2 deals with the higher value $c = 21$. In this case, the bandwidth is larger and

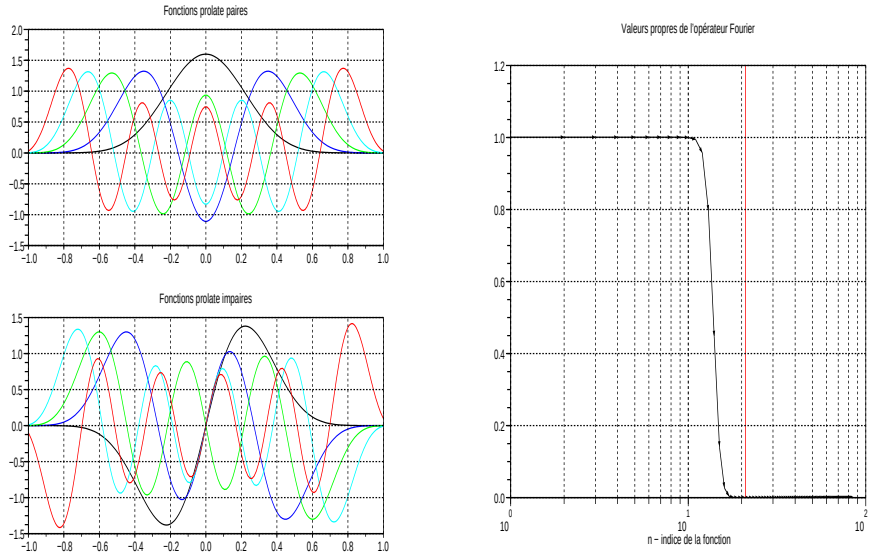

Fig. 2. Slepian functions (left) with $c = 21$ and repartition of eigenvalues λ_n (right).

consequently, Prolate functions are more peaked and more concentrated inside the interval $[-1, 1]$. For moderate values of $n \in \mathbb{N}$, their moduli remain bounded in the vicinity of the edges ± 1, which also differs from the preceding case. Eigenvalues λ_n still display their "step function" profile, but the numerical transition zone is now located before the value c, which reflects some difficulties in the task of computing them accurately as c grows (see *e.g.* [25]).

2.2 Some important results of Compressed Sampling (CS) theory

Compressive Sampling (or Compressed Sensing) consists now in a set of firmly founded mathematical algorithms which allow to reconstruct a so–called sparse signal out of a surprisingly restricted number of its observations, regardless to its possibly high frequencies. Roughly speaking, the required number of observed values is proportional only to *sparsity*, that is to say the amount of nonzero coefficients one needs in order to describe the signal in a supposedly well-suited base; of course, a signal can be sparse in a given basis and not in another one. For an introductory text, see `http://compressedsampling.com/`.

2.2.1 The "Restricted Isometry Property" (RIP)

Here, we shall not attempt to give the more advanced theoretical results of Compressed Sampling; rather we shall try to convey the main ideas underlying the algorithms and state some of the results having far-reaching consequences for our purposes. Details are to be found in the original papers [5–9].

A signal $f \in L^2(\mathbb{R})$ (or belonging to any Hilbert functional space) is said to be *s-sparse* in the basis $e_i(t)$, $i \in \mathbb{N}$ if the synthesis of $f(t)$ for any $t \in \mathbb{R}$ involves only a small number $s \in \mathbb{N}$ of elements in this basis. That is,

$$\forall t \in \mathbb{R}, \qquad f(t) = \sum_{i=0}^{s} \alpha_i e_i(t), \quad s \text{ small.}$$

In the discrete case, given a suitable basis $\vec{e}_i$, $i \in \{1, 2, ..., d\}$ of $\mathbb{R}^d$, we shall deal with *s-sparse vectors* $\vec{v}$ which satisfy $\vec{v} = \sum_{i=0}^{s} \alpha_i \vec{e}_i$ and $s \ll d$. For any $\vec{x} \in \mathbb{R}^d$, we also define its s-sparse approximation $\vec{x}_s$ as the vector made of the s largest entries of $\vec{x}$. Of course, any vector $\vec{v} \in \mathbb{R}^d$ being fixed, one can always consider an *ad hoc* basis such that its coordinates read $(1, 0, ...0)$; here we work with the unitary Inverse Discrete Fourier Transform (IDFT) reading $(\exp(i(k-1)(j-1)/d)/\sqrt{d})_{k=1,...,d,\ j=1,...,d}$ or the canonical basis of $\mathbb{R}^d$.

This being said, it makes sense to state the following definition:

Definition 1 *(Candès [6]) For any integer $d \geq s \in \mathbb{N}$, the restricted isometry constant δ_s of a (possibly rectangular) matrix Φ is defined as the smallest number such that it holds for every s-sparse vector $\vec{x}$: (we recall that for $1 \leq p < +\infty$, $\|\vec{x}\|_{\ell^p} = (\sum_{i=0}^{d} |x_i|^p)^{\frac{1}{p}}$)*

$$(1 - \delta_s)\|\vec{x}\|_{\ell^2}^2 \leq \|\Phi\vec{x}\|_{\ell^2}^2 \leq (1 + \delta_s)\|\vec{x}\|_{\ell^2}^2. \tag{9}$$

A vector $\vec{x} \in \mathbb{R}^d$ is said to be s-sparse if it has at most $s \leq d$ non-zero entries.

A matrix endowed with this Restricted Isometry Property for a certain integer s is constituted such that every set of of its d columns with cardinality less than s approximately behaves like an orthonormal system. This means in particular that if $\delta_s \ll 1$, then all the submatrices of Φ made of s columns are well conditioned; moreover, (9) is a frame inequality [10] restricted to s-sparse vectors for the frame defined as the set of d columns of Φ.

2.2.2 Recovery theorems: sparse and noisy cases

From this, Candès and his collaborators established the following recovery theorems:

Theorem 1 *(Candès [6]) Consider $\vec{y} \in \mathbb{R}^N$, $N \leq d$, a matrix Φ which satisfies $\delta_{2s} < \sqrt{2}-1$, and $\vec{x}_{opt} \in \mathbb{R}^d$ the solution of the following ℓ^1 minimization process:*

$$\inf_{\vec{x} \in \mathbb{R}^d} \|\vec{x}\|_{\ell^1}, \qquad \Phi\vec{x} = \vec{y} \in \mathbb{R}^d. \tag{10}$$

Then the following error estimates hold:

$$\|\vec{x}_{opt} - \vec{x}\|_{\ell^1} \leq C\|\vec{x}_s - \vec{x}\|_{\ell^1}, \qquad \|\vec{x}_{opt} - \vec{x}\|_{\ell^2} \leq \frac{C}{\sqrt{s}}\|\vec{x}_s - \vec{x}\|_{\ell^1}.$$

In particular, in case $\vec{x} \in \mathbb{R}^d$ is s-sparse, the recovery is exact.

A stability result is also available in case the observed vector $\vec{y}$ is corrupted by noise:

Theorem 2 *(Candès [6]) Consider $\vec{y} \in \mathbb{R}^N$, $N \leq d$, a matrix Φ which satisfies $\delta_{2s} < \sqrt{2}-1$, a noise vector $\vec{z} \in \mathbb{R}^N$ such that $\Phi\vec{x} = \vec{y} + \vec{z}$, $\|\vec{z}\|_{\ell^2} \leq \varepsilon$ and $\vec{x}_{opt} \in \mathbb{R}^d$ the solution of the convex minimization process:*

$$\inf_{\vec{x} \in \mathbb{R}^d} \|\vec{x}\|_{\ell^1}, \qquad \|\vec{y} - \Phi\vec{x}\|_{\ell^2} \leq \varepsilon. \tag{11}$$

Then the following error estimate holds with constants C, C' being explicitly computable:

$$\|\vec{x}_{opt} - \vec{x}\|_{\ell^2} \leq \frac{C}{\sqrt{s}}\|\vec{x}_s - \vec{x}\|_{\ell^1} + C'\varepsilon.$$

One issue remaining with both these results is that the user has to know in advance whether or not the observed vector $\vec{y} \in \mathbb{R}^N$ is noisy because depending on the answer to this question, one of the two programs (10) or (11) will be selected.

Remark 1 *The RIP (9) as advocated in both Theorems 1 and 2 has a particular meaning when Φ is any $N \times d$ rectangular matrix extracted from the $d \times d$ unitary IDFT matrix. The requirement $\delta_{2s} < \sqrt{2}-1$ means that for any $2s$-sparse vector of Fourier coefficients $\hat{z} \in \mathbb{C}^d$ of a (generally non-sparse) vector $z \in \mathbb{C}^d$, there holds:*

$$(2-\sqrt{2})\|z\|^2_{\ell^2(\mathbb{C}^d)} < \|\Phi\hat{z}\|^2_{\ell^2(\mathbb{C}^N)} < \sqrt{2}\|z\|^2_{\ell^2(\mathbb{C}^d)},$$

where $\|\hat{z}\|_{\ell^2(\mathbb{C}^d)} = \|z\|_{\ell^2(\mathbb{C}^d)}$ by Parseval's equality and $\Phi\hat{z} = \tilde{z}$ with $\tilde{z} \in \mathbb{C}^N$ is obtained by retaining only the coordinates of z corresponding to the N rows of Φ. So, this is a way of expressing that the original signal z cannot be strongly concentrated on the set of $d-N$ remaining points; on the contrary, it has to be rather uniformly spread on the whole set of d points, which is a consequence of $\hat{z}$ being sparse since the Fourier basis, having infinite support in the time variable, is strongly delocalized. This idea is even reinforced in case the set of N rows is "generic", in the sense that it is chosen equally probably as any other set of N rows among the available d ones: this is somewhat the reason behind exact recovery.

2.2.3 Numerical examples

The Restricted Isometry Property (9), which is actually crucial when it comes to underdetermined recovery, is verified by few classes of matrices only. Well-known examples (see [5]) are as follows:

- Gaussian measurements: Entries of $N \times d$ matrix Φ are independently sampled from the normal distribution $(0, \frac{1}{N})$. Then if $N \geq Cs\log(d/N)$, Φ meets the requirements of Theorems 1 and 2 with probability $1 - C\exp(-\gamma d)$, $\gamma > 0$.
- Fourier measurements: Rows of $N \times d$ matrix Φ are randomly chosen among the ones of the unitary IDFT matrix. The resulting columns are normalized so as to have norm 1 in ℓ^2 (and the entries of the resulting vector $\vec{x}_{opt}$ may have to be modified accordingly after the minimization process in case a FFT is necessary afterward). It is recalled in [26] that for $N \geq Cs\log(d)^\gamma$ ($\gamma = 6$ or $\gamma = 4$), Theorem 1 holds with overwhelming probability; it is conjectured that $\gamma = 1$ suffices in [5].

As an illustration, we display on Fig. 3 the recovery of a random signal in $\mathbb{R}^{256}$ involving 20 spikes (so $s = 20$ and $d = 256$ so this signal is sparse in the canonical basis of $\mathbb{R}^{256}$). We selected a random Gaussian measurement matrix with N is taken as the integer part of $s\log(d)$: the reconstruction can be considered as perfect. Comparison with a standard ℓ^2 minimization is also shown for which the numerical outcome is not sparse at all.

3 Extrapolation of low frequencies with Prolate functions

3.1 Numerical computation of Prolate functions inside $[-1, 1]$

The numerical computation of Slepian functions relies on the solving of the differential Sturm-Liouville equation (3) which is a perturbation of the one admitting the Legendre polynomials P_r as eigenfunctions. Relying on this last feature, it sounds natural to seek ψ_n as a Legendre series of the form:

$$\psi_n(t) = \sum_{r\in\mathbb{N}} \beta_{n,r}\bar{P}_r(t), \qquad \bar{P}_r(t) = P_r(t)\sqrt{r+1/2}. \tag{12}$$

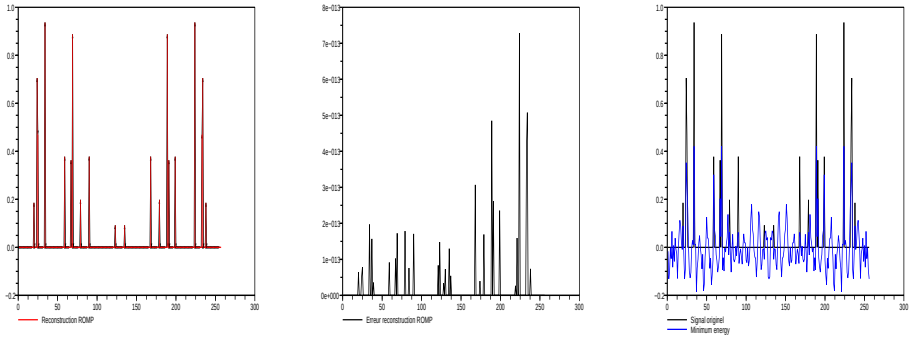

Fig. 3. Sparse spectrum recovery (left), absolute errors (middle) and ℓ^2 minimization (right).

The polynomials $\bar{P}_r$ are called *normalized Legendre polynomials* and satisfy $\int_{-1}^{1} |\bar{P}_r(t)|^2 dt = 1$. Plugging (12) inside (3) and rearranging leads to an eigenvalue problem for a symmetric matrix A; more precisely, we have the following result.

Theorem 3 *(Xiao et al. [38]) For any value of $c > 0$, the numbers χ_n and $\vec{\beta}_{n,.}$ are the eigenvalues and eigenvectors of the endomorphism of ℓ^2 represented by the infinite matrix A whose entries are zero except:*

$$A_{r,r} = r(r+1) + \frac{2r(r+1)-1}{(2r+3)(2r-1)}c^2, \;\; A_{r,r+2} = A_{r+2,r} = \frac{(r+2)(r+1)}{(2r+3)\sqrt{(2r-1)(2r+5)}}c^2.$$

Thanks to its particular structure, the matrix A can be split into two submatrices A_{even} and A_{odd} containing only even-numbered and odd-numbered rows and columns respectively. However, truncation of these infinite matrices isn't straightforward because their entries don't decay quickly with increasing rows and columns indexes. But thanks to the rapid decay in r of the scalar products $|\int_{-1}^{1} \psi_n(t) P_r(t).dt|$, see Theorem 3.4 in [38], the eigenvectors of interest can be obtained considering only the leading rows and columns of A_{even} and A_{odd}.

Hence a feasible algorithm to compute ψ_n through the Legendre series representation (12) goes as follows: solving the eigenvalue problems following Theorem 3:

$$(A_{even} - \chi_{2n}.Id)\vec{\beta}_{2n,.} = \vec{0}, \qquad (A_{odd} - \chi_{2n+1}.Id)\vec{\beta}_{2n+1,.} = \vec{0}, \qquad n \in \mathbb{N},$$

and then generating efficiently the values of the normalized Legendre polynomials $\bar{P}_r(t)$ at particular abscissas. Especially, it is interesting to derive values of $\bar{P}_r$ at points t_k, $k = 1, 2, ..., K$ corresponding to Legendre-Gauss quadrature [20]. For computing the nodes t_k and weights ω_k of Gaussian quadrature rules, the fundamental tool is the three-term recurrence relation satisfied by the set of orthogonal polynomials associated to the corresponding weight function. In case P_r is the monic orthogonal polynomial of degree r, it is a well-known fact that such orthogonal polynomials are related through a three-terms recurrence relation

$$\forall t \in [-1, 1], \qquad P_{r+1}(t) + a_r P_r(t) + b_r P_{r-1}(t) = t P_r(t), \quad P_0 \equiv 1, \quad P_{-1} \equiv 0.$$

One introduces therefore the so–called Jacobi $K \times K$ matrices reading,

$$J_1 = \begin{pmatrix} a_0 & 1 & 0 & \cdots & \cdots & 0 \\ b_1 & a_1 & 1 & 0 & \cdots & 0 \\ 0 & b_2 & a_2 & 1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \qquad J_2 = \begin{pmatrix} a_0 & \sqrt{b_1} & 0 & \cdots & \cdots & 0 \\ \sqrt{b_1} & a_1 & \sqrt{b_2} & 0 & \cdots & 0 \\ 0 & \sqrt{b_2} & a_2 & \sqrt{b_3} & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

and if $\vec{P} := (P_0(t_k), P_1(t_k), P_2(t_k), ...P_{K-1}(t_k))^T$, then $J_1\vec{P} = t_k\vec{P}$. Following [13] (see also [20]), the nodes of a Gaussian quadrature can actually be computed as the eigenvalues of the tridiagonal symmetric matrix J_2, and also the weights through the corresponding eigenvectors. This is the approach proposed by Gubner [15] in order to compute efficiently

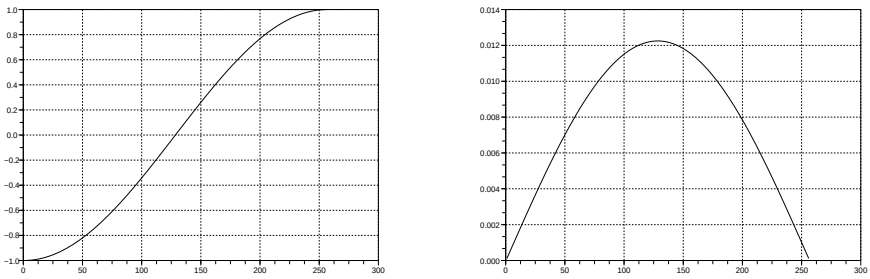

Fig. 4. Nodes (left) and weights (right) for Gauss-Legendre quadrature and $K = 256$.

any scalar product of the form which is exactly what is required for (7),

$$\int_{-1}^{1} f(t)\psi_n(t).dt \simeq \sum_{k=1}^{K} \omega_k f(t_k) \sum_{r=0,1}^{2N,2N+1} \beta_{n,r}\bar{P}_r(t_k), \qquad N \in \mathbb{N}, \tag{13}$$

with f any smooth function defined on $[-1, 1]$ and $(\omega_k, t_k)_{k=1,...,K}$ the weights and nodes of a Legendre-Gauss quadrature rule (see Fig. 4). The summation on index r is carried out according to the parity of n; MATLAB codes are also furnished in the papers [15] and [20]. Because of (12) and the convention which normalizes eigenvectors with unit ℓ^2 norm, both [38] and [15] present algorithms which involve Prolate functions normalized so as to be of unit norm in the interval $[-1, 1]$, on the contrary of (4).

3.2 Eigenvalues: how to pass from $[-1, 1]$ *to any abscissa of* $\mathbb{R}$

We saw before in §2.1.2 that one of the key points to complete the extrapolation process is the efficient computation of eigenvalues appearing in either (2) or (5); following [38] and taking advantage of our knowledge of the functions ψ_n restricted to the interval $[-1, 1]$, we present now a concrete way to complete this step. According to formula (5), it suffices to discretize the Fourier integral to produce the value of $\psi_n(x)$ for any $x \in \mathbb{R}$. This can be done accurately exploiting the Gaussian quadrature rule derived in the preceding section; namely, we write that:

$$\forall x \in \mathbb{R}, \qquad \psi_n(x) \simeq \sum_{k=1}^{K} \omega_k \exp(ic\ x\ t_k) \sum_{r=0,1}^{2N,2N+1} \beta_{n,r}\bar{P}_r(t_k),$$

which is just a special case of (13). All the problem of setting up (8) reduces therefore in extracting accurately the eigenvalues μ_n in (5). Following [38], we recall an useful result:

Theorem 4 *(Xiao et al. [38]) For any value of $c > 0$ and $m, n \in \mathbb{N}^2$ such that $m \neq n$ (mod 2), there holds:*

$$\frac{|\mu_m|^2}{|\mu_n|^2} = \frac{\int_{-1}^{1} \psi_n'(t)\psi_m(t).dt}{\int_{-1}^{1} \psi_n(t)\psi_m'(t).dt}.$$

It becomes possible to compute the eigenvalues by induction: first, λ_0 can be derived through

$$\mu_0\psi_0(t=0) = \int_{-1}^{1} \psi_0(\xi)d\xi.$$

Then, all the successive values of $|\mu_n|$, $n \geq 1$ can be obtained relying on Theorem 4 by using again the Gaussian quadrature derived in §3.1. At last, we recall from [38] that $\mu_n = i^n|\mu_n|$ and it remains to use (6). The figures 1 and 2 have been produced this way.

Remark 2 *Another algorithm to derive eigenvalues μ_n follows from the primary knowledge of Prolate functions ψ_n obtained from solving the diagonalization problem raised by Theorem 3: from the numerical values $\psi_n(t_k)$ known on the Gaussian quadrature points, one can compute through (13) the Fourier integral in (5) by the same quadrature rule. Thus it suffices to divide and compute an average to derive each corresponding eigenvalue. Numerical results are similar to those coming from the algorithm by Xiao et al.*

3.3 Numerical results

3.3.1 A simple test-case

We aim at completing a 15% extrapolation of the following simple and $C^\infty(\mathbb{R})$ (but not rigorously bandlimited) function,

$$f_{pro}(t) = \frac{1}{2}\exp(-2t^2)\cos(5\pi t), \qquad t \in [-1, 1], \tag{14}$$

that is to say, finding its numerical values on the interval $]1, 1.3]$ by means of its Prolate expansion following equations (7), (8) and the numerical procedures explained in the preceding sections. The result is shown in Fig. 5 with the choice $c = 30$ for the bandwidth. One

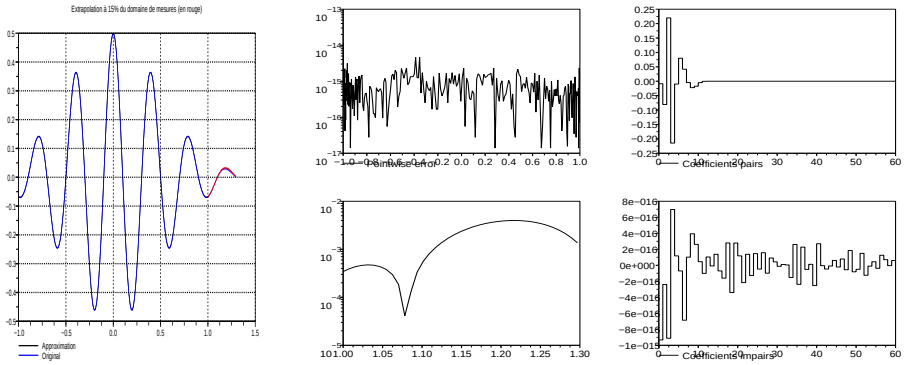

Fig. 5. Extrapolation (left), absolute errors (middle) and Prolate coefficients (right).

can observe that for this value, the compression is very good since roughly 10 coefficients

acting on even Prolate functions (f_{pro} is clearly even) suffice to get a very good approximation: absolute errors inside the interval $[-1,1]$ are of the order of 10^{-14}: this agrees with the values indicated in Table 1 and Theorem 3.11 of [29]. Coefficients for odd Prolate functions are of the order of 10^{-16}. When it comes to the extrapolation, errors grow no more than values around 10^{-3}, which is still satisfying. Gaussian quadrature was used with 256 nodes.

3.3.2 First composite problem

We are now interesting in extrapolating more involved signals, for which we shall split between lower and higher frequencies, each one being treated by means of a specific algorithm. The examination of absolute errors made on the global problem will be deferred to §5. First, we consider the following smooth function for $t \in [-1,1]$:

$$f(t) = \underbrace{\exp\left(\frac{t}{2}\right)\sin(\pi t)}_{f_{lo}(t)} + \underbrace{0.15[\sin(13\pi t) - 0.5\cos(23\pi t) + 0.3\cos(7\pi t)]}_{f_{hi}(t)}. \tag{15}$$

In order to check whether or not the splitting between low and high frequencies is admissible,

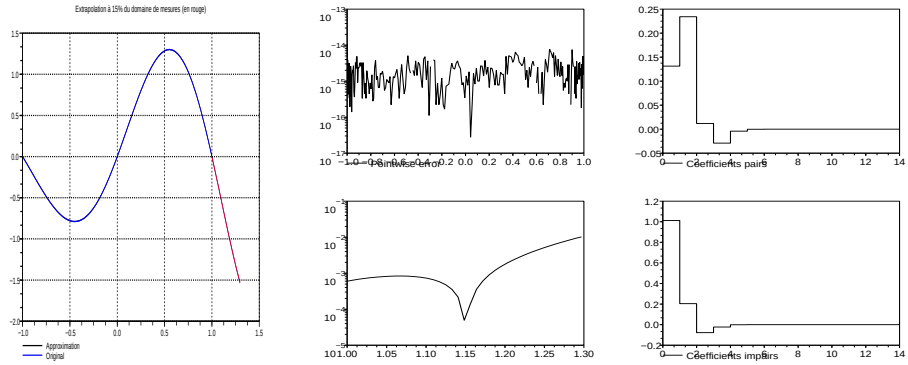

Fig. 6. Extrapolation (left), absolute errors (middle) and Prolate coefficients (right).

we computed the empirical correlation between these two parts following (1) where integrals are obtained with the Gaussian quadrature like (13): we found that for $a = -1$, $b = 1$, $\kappa \simeq 0.0014$ which can be considered as satisfying. The first part is to be treated by the extrapolation algorithm based on Prolate functions with the bandwidth parameter $c = 7$. The numerical outcome for an extrapolation of 15% (up to $t = 1.3$) is shown in Fig. 6. Prolate coefficients are equally spread among even and odd PSWF. The reconstruction error inside the interval $[-1,1]$ is very low (of the order of 10^{-16}); the extrapolation error is bigger, but remains at acceptable levels (of the order of 10^{-3}). These results can be considered as being satisfying.

3.3.3 Second composite problem

As a second test-case, we consider the quite intricate smooth function:

$$\forall t \in [-1,1], \quad f(t) = \underbrace{\exp(-2t^2)\exp(3t)\sin(\pi t)\cos(3\pi t)}_{f_{lo}(t)} + \underbrace{0.5[\sin(5\pi t) - \cos(7\pi t)]}_{f_{hi}(t)}. \tag{16}$$

Similarly as in the preceding example, we checked (1): with $a = -1$, $b = 1$, $\kappa \simeq -0.019$. Here, we aim at extrapolating the first Gaussian-modulated part from the observed interval $[-1,1]$ to $]1,1.3]$ by means of PSWF with the bandwidth parameter $c = 20$. Numerical

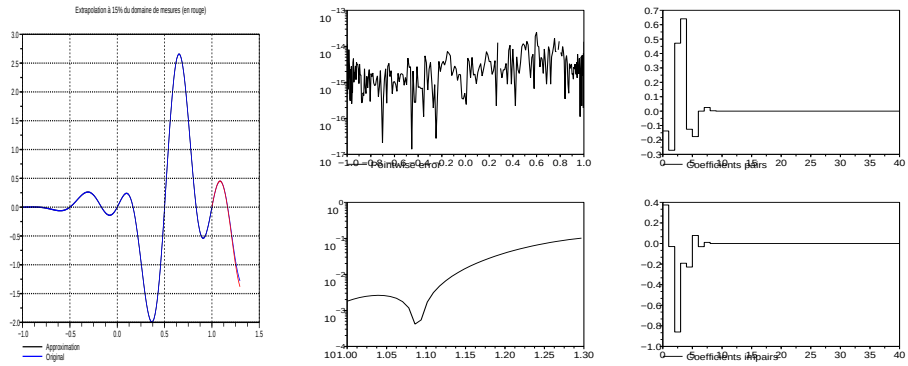

Fig. 7. Extrapolation (left), absolute errors (middle) and Prolate coefficients (right).

results are displayed on Fig. 7. One can see that reconstruction errors inside $[-1, 1]$ are still very small, but extrapolation errors for $1 < t \leq 1.3$ are bigger compared to the preceding example since they are now of the order of 10^{-2}. Few Prolate coefficients are needed to represent the function (7) despite the fact it is not rigorously band-limited because of the Gaussian function involved. Results are still satisfying though.

4 Extrapolation of higher frequencies with Compressed Sampling

4.1 The class of compressible signals

Roughly speaking, compressible signals correspond to a subclass for which there exists a convenient Hilbert basis allowing to represent each of them with a very limited number of coefficients; in this case, one says that these signals are "well compressed" in this basis. This notion is clearly dependent on the choice of the aforementioned basis since, in the case one considers Fourier or wavelet bases, smoothness in the time variable implies compressibility in both these bases. With the Fourier transform, it is possible to give a precise meaning to the subclass of "signals being compressible in the Fourier basis":

Proposition 1 *Let $f \in C^p(\mathbb{R})$ and suppose that $f^{(i)}$ belongs to $L^1(\mathbb{R})$ or to $L^2(\mathbb{R})$, then for all $\xi \in \mathbb{R}$, $|\hat{f}(\xi)| \leq C_p(1 + |\xi|)^{-p}$ for a suitable constant $C_p \in \mathbb{R}^+$.*

Consequently, the class of compressible signals for which the approximate recovery theorems hold contains all the signals having a certain number of derivatives when expressed within the temporal variable. A property of the same flavor would also hold for wavelet bases. Hence we can introduce a discrete analogue of this notion:

Definition 2 *Let $\vec{x} \in \mathbb{R}^d$; we say that it is a compressible vector in a given basis if its coefficients decay according to a power law. More precisely, there exists $p \geq 1$ such that for any $k \leq d$, the modulus of the k^{th} largest coefficient of $\vec{x}$ is bounded by $C_p k^{-p}$.*

Being correctly oversampled, bandlimited functions clearly yield discrete compressible signals in the Fourier basis since their Fourier transform has compact support. The first result concerning recovery of randomly sampled but compressible signals has been given in [9]:

Theorem 5 *(Optimal recovery of weak-ℓ^p, [9]) Let $\vec{x} \in \mathbb{R}^d$ a compressible vector in the sense of Definition 2 with $p > 1$ or $\|\vec{x}\|_{\ell^1} \leq C_1$ for $p = 1$, and let $\alpha > 0$ be a sufficiently small number. Assume we are given N random measurements $\Phi\vec{x}$, then with probability 1, the minimizer $\vec{x}_{opt}$ of (10) is unique. Furthermore, with probability at least $1 - O(d^{-\frac{\rho}{\alpha}})$, the*

following error bound holds:

$$\|\vec{x} - \vec{x}_{opt}\|_{\ell^2} \leq C_{p,\alpha} C_p \left(\frac{\log d}{N} \right)^{p-\frac{1}{2}}. \tag{17}$$

Remark 3 *Theorem 5 doesn't refer to the RIP despite an earlier but similar concept appears in its proof. Its conclusion is probabilistic, as were the ones of the results proved in [7]: it says that given $O(s \log d)$ measurements, one is able to retrieve through ℓ^1 minimization an approximation which is as good as if the s largest entries of $\vec{x}$ were known in advance. Moreover, the recovery algorithm (10) doesn't need to know the quantities p and C_p.*

4.2 Regularized Orthogonal Matching Pursuit (ROMP)

Both theorems stated in §2.2.2 rely on the numerical solving of a ℓ^1 minimization program. Despite the fact powerful methods now exist for completing this task, we prefer to turn to lighter iterative solvers, like the so–called ROMP routine introduced in [26,27]. In this section, we recall two recovery theorems for ROMP, adapted to the kind of problems we want to process; the measurement $N \times d$ matrix is still denoted by Φ.

Theorem 6 *(Needell et al. [26]) Consider $\vec{y} \in \mathbb{R}^N$, $N \leq d$ and the matrix Φ which satisfies $\delta_{2s} \leq 0.03/\sqrt{\log s}$. For any $\vec{x} \in \mathbb{R}^d$ with $\vec{y} = \Phi\vec{x}$, ROMP outputs a set I such that $supp(\vec{x}) \subset I$, $|I| \leq 2s$ in at most s iterations.*

This case corresponds to exact sparse recovery; we note that Theorem 2.1 in [26] gives rigorous bounds in terms of RIP for widely used random measurement matrices. Concerning the stability when measurements or the original (supposedly sparse) signal is perturbed, we state the result in the context of compressible signals for which no algorithmic changes are required (contrary to (10) and (11)):

Theorem 7 *(Needell et al. [27]) Consider $\vec{y} \in \mathbb{R}^N$, $N \leq d$ and the matrix Φ which satisfies $\delta_{8s} \leq 0.01/\sqrt{\log s}$. For any compressible vector $\vec{x} \in \mathbb{R}^d$ in the sense of Definition 2 with $\vec{y} = \Phi\vec{x} + \varepsilon$, ROMP produces an approximation to $\vec{x}$ satisfying:*

$$\|\vec{x}_{romp} - \vec{x}\|_{\ell^2} \leq \sqrt{\log s} \left(\frac{\tilde{C}_p}{s^{p-\frac{1}{2}}} + \tilde{C}\|\varepsilon\|_{\ell^2} \right). \tag{18}$$

We used the same letter p in both Proposition 1 and Definition 2 on purpose since the discrete Fourier transform converts smoothness into fast coefficients decay. Hence, assuming we deal only with smooth signals, and under a strong RIP restriction, Theorem 7 gives us an error control which compares with (17) onto the output of the ROMP algorithm *in a case where the original vector $\vec{x}$ isn't sparse*. The authors of [27] conjecture that the logarithmic factor in front of ε is unnecessary, and that a bound like the one of Theorem 2 actually holds. We point out that a MATLAB code for ROMP is available at the following location: `http://www.math.ucdavis.edu/ dneedell/romp.m`

4.3 The choice of convenient random matrices

The first articles [7–9] insist on the fact that the spectacular results of CS stem from the combined use of random sensing on the one hand and sparse recovery through ℓ^1 minimization on the other hand. In order to achieve good numerical quality, the randomness of Φ is crucial; however, in the context of extrapolation, it is not completely obvious how to meet this requirement. Since Φ is obtained by randomly selecting $N = O(s \log d)$ rows inside the

unitary $d \times d$ IDFT matrix and further renormalizing its columns so as to be of unit ℓ^2 norm, one can easily understand that not all random sequences can be used. Actually, we only know a fraction of $d_{obs} < d$ values; moreover, the $d - d_{obs}$ unknown values are all packed at the end of the observations interval in a strongly non-random manner. Hence, in order to set up correctly the CS algorithms, we must work with sensing matrices Φ obtained with those particular random sequences of N rows among d *which don't perceive the unknown* $d - d_{obs}$ *values* located on the right of the observations interval. Clearly, this puts a limitation on the size of the ratio $\alpha = d/d_{obs} \geq 1$ as it will become more and more difficult to find such sequences for increasing α's while keeping constant the number $N = O(s \log d)$. We stress that according to the results of [7,9], the simple idea of selecting a random sequence of N rows among the d_{obs} first ones (corresponding to the observed values) isn't convenient; indeed, the selected chosen measurement ensemble of N values has really to be equally likely as any other one (see [7], page 492 and Remark 1).

Remark 4 *It is interesting to observe that compared to the classical extrapolation techniques based on variants of Gerchberg-Papoulis (GP) algorithm [12,28,16,35], the use of CS techniques allows to handle cases which would be intractable otherwise. In [11], a bound is given for uniqueness and perfect recovery using GP; within our notation, it reads:*

$$N > d\left(1 - \frac{1}{2s}\right).$$

This is highly more restrictive than the restriction $N \geq O(s \log d)$ which suffices to apply the theorems in [7–9] holding for ℓ^1 minimization and random Fourier measurements.

Usual random sequences possibly contain the abscissas to be extrapolated, which is a problem because they correspond to unknown values of the signal (these values we precisely want to discover through the extrapolation process). Hence it may be tempting to restrict random sequences to the set of abscissas corresponding to the observed part only, which is the most straightforward way to proceed. However, such a choice leads to numerical results showing a sharp transition in terms of absolute errors when the edge of the observations interval is crossed. It resulted more efficient to generate global random sequences and, as a post-processing, to discard the subset of which perceiving something inside the unknown part. This seemingly differs in terms of numerical accuracy because the transition zone between observed and extrapolated parts of the signal is less noticeable; see for instance the figure 10 for which the signal is extrapolated in the area $t \in]1, 1.3]$.

Finally, a simple extrapolation algorithm based on Compressed Sensing techniques reads:

- Evaluate the sparsity s of the Fourier transform of the signal $\vec{x} \in \mathbb{R}^d$ to be reconstructed, for which there are d_{obs} known values and $d - d_{obs}$ values to be extrapolated (generally located in one or two connected sets).
- Compute the length $N = O(s \log d)$ of the random sequence to be used to under-sample $\vec{x}$ in order to produce $\vec{y}$; decreasing the ratio $\alpha = d/d_{obs}$ in case no convenient sequence can be found may be necessary.
- Build up the sensing matrix $\Phi = \frac{1}{\sqrt{N}}(\exp(i(k-1)(j-1)/d))_{k \ random \ of \ length \ N, \ j=1,\ldots,d}$ and $\vec{y} = (x_k)_{k \ random \ of \ length \ N}$.
- Compute the optimal $\vec{x}_{opt}$ by ℓ^1 minimization or $\vec{x}_{romp}$ through the ROMP algorithm.
- In general, one has to multiply this vector by $\sqrt{N}$ in order to produce the extrapolated signal by carrying out its FFT afterward.

4.4.1 The simple test-case

We repeat the simple test-case of producing a 15% (*i.e.* $\alpha = 1.15$) extrapolation of the function (14) observed in the interval $[-1, 1]$ using the ROMP algorithm. The sparsity level has been fixed as $s = 8$ and a Fourier measurement matrix was used with N the integer part of $s\log(K)$ where $K = 256$ is the number of discretization points in $[-1, 1]$. Results are shown in Fig. 8 which can be directly compared to Fig. 5 which involved an extrapolation based on Prolate functions. The outcome is less satisfying, especially inside the observation

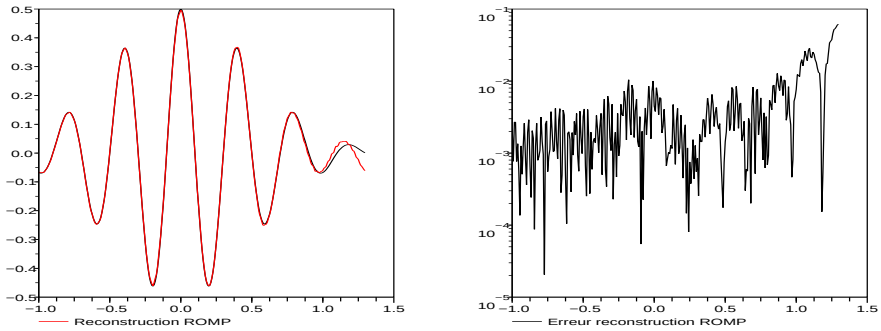

Fig. 8. Extrapolation with ROMP (left) and absolute errors (right).

interval since the absolute errors in this region are of the order of 10^{-3}; the absolute errors in the extrapolation region are around 10^{-2}. We stress that the Fourier basis we used isn't optimal for such a signal which isn't periodic; even if the difference $|f_{pro}(1.3) - f_{pro}(-1)|$ isn't big, the numerical Fourier transform won't be rigorously sparse. Hence we believe that part of the extrapolation error comes from this fact because the ROMP algorithm seeks an extrapolation endowed with a s-sparse representation in the Fourier basis. This phenomenon should decrease when considering signals of higher frequency since the accuracy of Compressed Sensing doesn't depend on it (on the contrary of Prolate functions which heavily depends on the bandwidth). This experiment confirms our choice of extrapolating lower frequencies by means of Prolate functions; compare Fig. 8 with Fig. 5.

Remark 5 *There is a special case for which the Fourier basis is especially well adapted, namely the one involving signals being periodic with a period equal to the Lebesgue measure of the whole interval (in our context $|(-1, 1.3)| = 2.3$). For instance, we briefly consider $f_{per}(t) = \cos(8\pi(t+1)/2.3) + \sin(14\pi(t+1)/2.3)$; from observations of f inside the interval $[-1, 1]$, we aim at extrapolating in $]1, 1.3]$ using the ROMP algorithm with $N = 1.5s\log d$ random measures where $s = 6$, $d = 295$. Many convenient random sequences exist and the numerical results are displayed in Fig. 9. The salient feature is that for this particular class of signals, the transition between the absolute errors inside the observations interval and the ones in the extrapolation domain is hardly noticeable. This particular case shows how crucial is the choice of the basis for the quality of recovery and extrapolation.*

4.4.2 First composite problem

We now deal with the oscillatory part of (15) which reads:

$$\forall t \in [-1, 1], \qquad f_{hi}(t) = 0.15[\sin(13\pi t) - 0.5\cos(23\pi t) + 0.3\cos(7\pi t)].$$

We carry out a 15% extrapolation of this strongly band-limited function by means of the

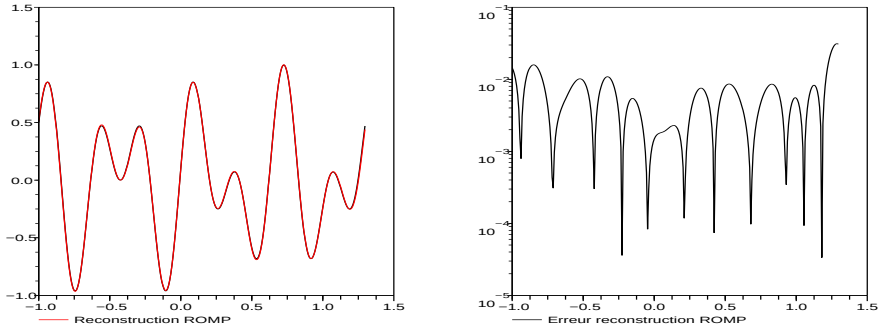

Fig. 9. Periodic extrapolation with ROMP (left) and absolute errors (right).

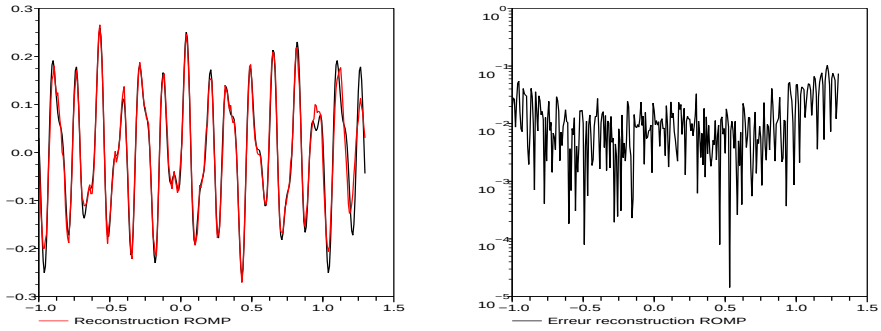

Fig. 10. Extrapolation with ROMP (left) and absolute errors (right).

ROMP algorithm with parameters $s = 12$ and the choice $N = 1.1s\log(d)$ where $d = 295$. In particular, finding convenient random sequences as explained in §4.3 is possible only for $N < 1.2s\log(d) \simeq 82$. The results displayed in Fig. 10 show that these CS techniques handle such an extrapolation problem quite well; the absolute recovery errors are spread rather uniformly inside the interval $[-1, 1.3]$ with a slight increase close to its edges. When comparing with Fig. 8, one sees that the Fourier basis isn't really adapted to the present case as well; however, the results are still rather good because the salient feature of the signal is its frequencies. And since there's no limitation in Compressed Sensing related to the bandwidth (only the sparsity of the Fourier transform matters), higher frequencies allows to get a slightly oversampled signal as an input.

4.4.3 Second composite problem

The oscillatory part of (16) has lower frequencies involved and reads:

$$\forall t \in [-1, 1], \qquad f_{hi}(t) = \frac{1}{2}[\sin(5\pi t) - \cos(7\pi t)].$$

We treat its extrapolation from $[-1, 1]$ to $]1, 1.3]$ by means of ROMP with $s = 7$ and a slightly lower number of random observations, namely $N = s\log(295) \simeq 40$. Results are shown in Fig. 11. This case contains frequencies lower than the preceding one, but its structure is a bit simpler as only 2 trigonometric functions are involved. Fig. 11 shows that the reconstruction is rather good uniformly in the computational domain, but there is a slight increase of absolute errors close to the edges of the interval $[-1, 1]$ of observations. Starting from $t \simeq 0.8$, there are some small high-frequencies oscillations which don't perturb

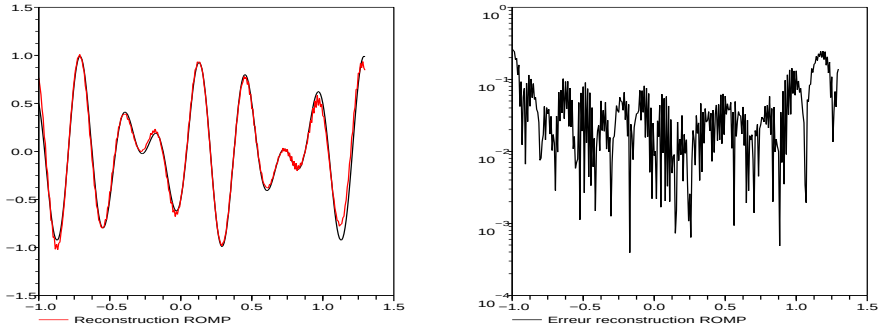

Fig. 11. Extrapolation with ROMP (left) and absolute errors (right).

very much the overall quality of the extrapolated signal.

5 Putting the pieces altogether

By doing a byproduct of Theorem 3.11 in [29] (see also Theorem 3 in [2]) and Theorem 7, it would be possible to derive an error estimate for the recovery of the original signal f by means of both algorithms put together *inside the observations interval only*. Indeed, it seems that no rigorous error estimates exist concerning the extrapolation of a band-limited signal through a truncated Prolate series of the type (8), even if in our context, (17) and (18) provide a reliable indication of the discrepancy coming from the Compressed Sensing part. This may constitute a direction for future research.

5.1 First composite problem

We now present the results of combined extrapolation algorithms on problem (15) in Fig. 12 together with the corresponding absolute errors on the whole interval $[-1, 1.3]$. We observe

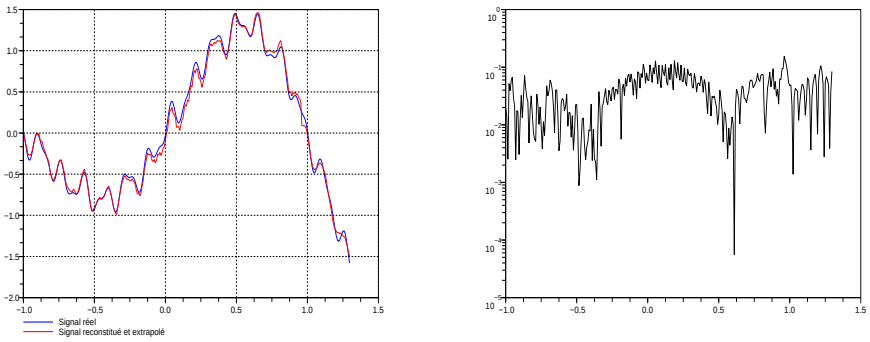

Fig. 12. First case: global extrapolation (left) and absolute errors (right).

that the overall quality is quite good from both viewpoints of recovery (inside the interval $[-1, 1]$) and extrapolation (inside $]1, 1.3]$). Absolute errors are uniformly distributed inside the whole interval $[-1, 1.3]$ and their amplitudes vary roughly between 10^{-3} and 10^{-1}; especially, there is no sharp transition when passing the edge $t = 1$. From Fig. 6, we know that absolute errors inside $[-1, 1]$ come in very large part from the ROMP algorithm since

the reconstruction errors for Prolate functions are negligible. In $[-1, 1.3]$, both algorithms are endowed with roughly the same amount of numerical errors, and they accumulate when adding lower and higher parts of the signal. However, the numerical outcome can be considered as globally satisfying.

5.2 Second composite problem

The extrapolation of function (16) is shown in Fig. 13. This test-case is slightly more

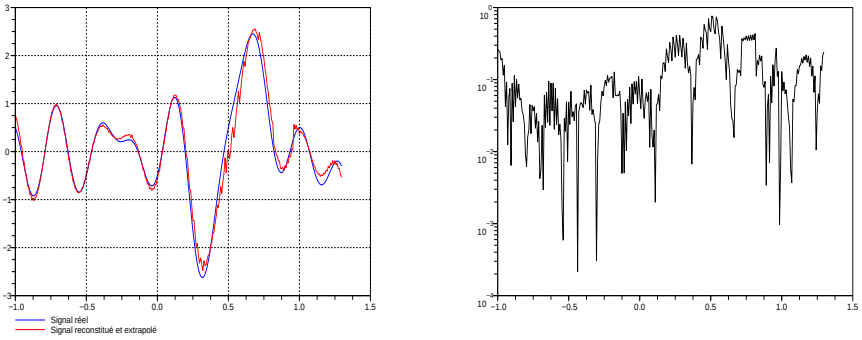

Fig. 13. Second case: global extrapolation (left) and absolute errors (right).

difficult because the correlation between f_{lo} and f_{hi} is stronger, which means that the difference between the lower and higher frequencies parts in the signal f is less sharp. As a consequence, it is less straightforward to make a clean separation between its "trend" and its "fluctuation" (compare with Figures 6, 10 and 12). There are spurious oscillations of high frequency around the location $t = 0.5$ and this comes from the recovery of ROMP since the reconstruction errors of Prolate functions are once again negligible (see Fig. 7). Absolute errors oscillate between 10^{-2} and 10^{-1} which amounts to a relative error around 10% except in this instability area where it comes close to $\frac{1}{2}$. It is quite satisfying that, similarly as the preceding example, no visible transition appears when crossing the abscissa $t = 1$ separating measures from extrapolation.

6 Conclusion and outlook

We have presented in this paper a practical methodology to carry out the extrapolation of smooth and compressible signals relying on modern numerical techniques, namely the computation of Slepian's Prolate functions which are well suited for processing the reasonably low frequency signals and the Compressive Sampling techniques recently developed by Candès et al. which are independent of the size of the bandwidth hence are endowed with enough robustness to treat higher frequencies, especially when they can be expressed by means of a sparse Fourier representation. Each method has been tested independently and then, the absolute errors obtained by gluing together both algorithms have been studied. It came out that in order for this numerical method to deliver good results, several restrictions should be met: the correlation (1) between between low and high frequencies should be weak, the high frequencies should have a sparse representation in the Fourier basis (see §2.2.1 and Definition 2), and the size of the extrapolation domain should be reasonable (see §4.3). When these three points are met, numerical results are far better than those which could be obtained by standard methods, like *e.g.* Gerchberg-Papoulis methods at least because they wouldn't be able to treat such high frequencies (see Remark 4).

References

[1] A. Bensoussan, J.L. Lions, G. Papanicolaou, **Asymptotic analysis for periodic structures**, Amsterdam; North-Holland 1978.

[2] John P. Boyd, *Approximation of an analytic function on a finite real interval by a bandlimited function and conjectures on properties of prolate spheroidal functions*, Appl. Comp. Harm. Anal. **15** (2003) 168–176.

[3] John P. Boyd, *Prolate spheroidal wavefunctions as an alternative to Chebyshev and Legendre polynomials for spectral element and pseudospectral algorithms*, J. Comp. Phys. **199** (2004) 688–716

[4] S. Cambanis and E. Masry, *Zakai's class of bandlimited functions and processes: its characterization and properties* SIAM J. Appl. Math **30** (1976) 10–21.

[5] E.J. Candès, *Compressive sampling*, Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006.

[6] E.J. Candès, *The restricted isometry property and its implications for compressed sensing* C.R. Acad. Sci. Paris, Serie I, **346** 589–592.

[7] E.J. Candès, J. Romberg and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. **59** 1207–1223.

[8] E.J. Candès, J. Romberg and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory **52** 489–509.

[9] E.J. Candès and T. Tao, *Near-optimal signal recovery from random projections: universal encoding strategies*, IEEE Trans. Inform. Theory, **52** 5406-5425.

[10] Ole Christensen, **An introduction to frames and Riesz bases**, Birkhäuser 2003

[11] D.L. Donoho, P.B. Stark, *Uncertainty principles and signal recovery*, SIAM J. Appl. Math. **49** (1989) 906–931.

[12] R.W. Gerchberg, *Super-resolution through error energy reduction*, Opt. Acta **21**, 709-720 (1974)

[13] G.H. Golub, J.H. Welsch, *Calculation of Gauss quadrature rules*, Math. Comp. **23**, 221–230 (1969)

[14] L. Gosse, *Analysis and short-time extrapolation of stock market indexes through projection onto discrete wavelet subspaces*, submitted to Signal Processing (2008).

[15] John A. Gubner, *A Simple Method for Computing Projections onto Subspaces of Prolate Spheroidal Wave Functions*, sumbitted to IEEE Trans. Comm., see `http://eceserv0.ece.wisc.edu/ gubner/pswfProjIpSubmitted2col.pdf`

[16] A.K. Jain, S. Ranganath, *Extrapolation algorithm for discrete signal with application in spectral estimation*, IEEE Trans. Acoust. Speech Signal Process. **29** (1981) 830–845.

[17] K. Khare, N. George, *Sampling theory approach to prolate spheroidal wave functions*, J. Phys. A: Math. Gen. **36** (2003), 10011–10021.

[18] P. Kirby, *Calculation of spheroidal wave functions*, Computer Physics Communications Volume 175, Issue 7, 1 October 2006, Pages 465-472

[19] B. Larsson, T. Levitina, E.J. Brändas, *On prolate spheroidal wave functions for signal processing*, Intern. J. Quantum Chemistry **85** (2001) 392–397.

[20] Dirk P. Laurie, *Computation of Gauss-type quadrature formulas*, J. Comp. Appl. Math. **127** (2001) 201–217.

[21] Alan Lee, *Approximate Interpolation and the Sampling Theorem* SIAM J. Appl. Math **32** (1977) 731–744.

[22] L.-C. Lin, C.-C. Jay Kuo, *On theory and regularization of scale-limited extrapolation*, Signal Processing **54** (1996) 225–237.

[23] L.-C. Lin, X.-G. Xia, C.-C. Jay Kuo, *On the convergence of wavelet-based iterative signal extrapolation algorithms*, Signal Processing **48** (1996) 51–65.

[24] I.C. Moore, M. Cada, *Prolate Spheroidal wave functions, an introduction to the Slepian series and its properties*, Appl. Comput. Harmonic Anal. **16** (2004) 208–230.

[25] Abderrazek Karoui and Tahar Moumni, *New efficient methods of computing the prolate spheroidal wave functions and their corresponding eigenvalues*, Appl. Comp. Harm. Anal. **24** (2008) 269–289.

[26] D. Needell and R. Vershynin, *Uniform Uncertainty Principle and signal recovery via Regularized Orthogonal Matching Pursuit*, Foundations of Computational Mathematics, DOI: 10.1007/s10208-008-9031-3.

[27] D. Needell, R. Vershynin, *Signal recovery from incomplete and inaccurate measurements via Regularized Orthogonal Matching Pursuit*, sumbitted.

[28] A. Papoulis, *A New Algorithm in Spectral Analysis and Band-Limited Extrapolation*, IEEE Trans. Cir. & Sys., **22** (1975) 735–742.

[29] Yoel Shkolniskya, Mark Tygertb, Vladimir Rokhlin, *Approximation of bandlimited functions*, Appl. Comp. Harm. Anal. **21** (2006) 413–420.

[30] D. Slepian, *On bandwidth*, Proceedings of the IEEE **64** (March 1976) 292–300.

[31] D. Slepian, *Some comments on Fourier analysis, uncertainty and modeling*, SIAM Rev. **25** (1983) 379–393.

[32] Slepian D., Pollak H.O., *Prolate spheroidal wave functions, Fourier analysis and uncertainty. I.*, Bell System Technical Journal **40** (1961) 43–63.

[33] Slepian D., *Prolate spheroidal wave functions, Fourier analysis and uncertainty, IV*, Bell System Technical Journal **43** (1964) 3009–3057.

[34] Herbert Spohn, Stefan Teufel, *Adiabatic decoupling and time-dependent Born-Oppenheimer theory*, Comm. Math. Phys. **224** (2001), 113132

[35] T. Strohmer, *On discrete band-limited signal extrapolation*, Contemp. Math. 190:323-337, 1995

[36] G. Walter and T. Soleski, *A new friendly method of computing prolate spheroidal wave functions and wavelets*, Appl. Comput. Harmon. Anal. 19 (2005),

[37] X.-G. Xia, C.-C. Jay Kuo, Z. Zhang, *Signal extrapolation in wavelet subspaces*, SIAM J. Sci. Comp. **16** (1995) 50–73.

[38] H. Xiao, V. Rokhlin, N. Yarvin, *Prolate spheroidal wave functions, quadrature and interpolation*, Inverse Problems **17** (2001) 805–838.

[39] K. Yao, *Applications of reproducing kernel Hilbert spaces–Band-limited signals models*, Information & Control **11** (1967) 429–444.

[40] D.C. Youla, H. Webb, *Image Restoration by the Method of Convex Projections: Part 1 Theory*, IEEE Trans. Med. Im., vol. MI-1, No. 2, 1982.

[41] Moshe Zakai, *Band-Limited Functions and the Sampling Theorem* Information and Control **8** (1965) 143–158.

[42] A. Zayed, *A generalization of the prolate spheroidal wave functions*, Proc. Amer. Math. Soc. **135** (2007) 2193–2203.

[43] A. Zayed, *On the notion of bandlimitedness and its generalizations*, Rev. Union Matematica Argentina **49** (2008) 99–109.

[44] J. Zhou, A. Gilbert, M. Strauss, I. Daubechies, *Theoretical and experimental analysis of randomized algorithm for sparse Fourier transform analysis*, J. Comp. Phys. **211** (2006) 572–595.