

SPARSITY PATTERN RECOVERY USING FRI METHODS

Jon Oñativia[†], Yue M. Lu^{*} and Pier Luigi Dragotti[†]

[†] Communications and Signal Processing Group (CSP), Imperial College London, UK
 { jon.onativia, p.dragotti } @imperial.ac.uk

^{*} Signals, Information, and Networks Group (SING), Harvard University, USA
 yuelu@seas.harvard.edu

ABSTRACT

The problem of finding the sparse representation of a signal has attracted a lot of attention over the past years. In particular, uniqueness conditions and reconstruction algorithms have been established by relaxing a non-convex optimisation problem.

The finite rate of innovation (FRI) theory is an alternative approach that solves the sparsity problem using algebraic methods based around Prony’s algorithm. Recent extensions to this framework have shown that it is possible to recover sparse representations beyond the uniqueness limits, that is, finding all the possible sparse representations that fit the observation for the case of signals which are sparse in the union of Fourier and canonical bases. In this paper, we show the application of such methods to the case of the union of DCT and Haar basis. We present an extension that takes advantage of the even symmetry of the cosine functions to build an algorithm that can operate over the observed vector and in a dual domain. We also analyse the case of the union of frames. Simulation results confirm the validity of this new approach and show that it outperforms state of the art algorithms in a number scenarios.

Index Terms— Sparse representation, union of bases, Prony’s method, finite rate of innovation.

1. INTRODUCTION

Many signals of interest have often a simpler representation in some appropriate overcomplete dictionaries. However, finding this representation is an ill-posed problem since we are dealing with an underdetermined system of equations. Images are a good example of such signals, where, for instance, regular textures have a compact representation in the DCT basis but discontinuities and localised details can be better represented in wavelet dictionaries. In this scenario, we would like to find the combinations of DCT and Haar elements that are able to reproduce the 2D image with the smallest number of elements, *i.e.*, we want to find a sparse representation in an overcomplete dictionary consisting of the union of DCT and Haar atoms. Being able to recover a sparse representation of an observed signal has a wide range of applications going from compression to denoising or signal detection. It is therefore necessary to devise fast and reliable algorithms that can find such representations.

Let $\mathbf{y} \in \mathbb{C}^N$ be a complex-valued finite dimensional signal that has a K -sparse representation in an overcomplete dictionary, that is,

$$\mathbf{y} = \mathbf{D} \mathbf{x}, \quad (1)$$

This work is supported by the European Research Council (ERC) starting investigator award Nr. 277800 (RecoSamp).

where $\mathbf{D} \in \mathbb{C}^{N \times L}$ is the overcomplete dictionary with $L > N$ atoms and $\mathbf{x} = (x[\ell])_{\ell=0}^{L-1}$, with $x[\ell] \in \mathbb{C}$, satisfies $\|\mathbf{x}\|_0 \stackrel{\text{def}}{=} \#\{\ell : |x[\ell]| \neq 0\} = K$. One way of recovering the sparse vector \mathbf{x} from the observation \mathbf{y} is by solving the following problem:

$$(P_0) : \quad \arg \min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{D} \tilde{\mathbf{x}}, \quad (2)$$

which is intractable since the ℓ_0 “norm” is nonconvex. In practice, this problem can be solved via a convex relaxation, which replaces the ℓ_0 -norm by the ℓ_1 -norm [1]–[5]. This technique is known as Basis Pursuit and under some conditions, the ℓ_1 approach is guaranteed to find the solution to (P_0) . It is thus important to establish the uniqueness of the solution to (P_0) . Since the system in (1) is underdetermined, there are an infinite number of vectors \mathbf{x} that lead to the same observation \mathbf{y} , provided that \mathbf{y} belongs to the span of the columns of \mathbf{D} . However, it has been shown that in the case where \mathbf{D} is built from the union of two orthogonal bases, the solution is unique if

$$\|\mathbf{x}\|_0 < 1/\mu(\mathbf{D}), \quad (3)$$

where

$$\mu(\mathbf{D}) \stackrel{\text{def}}{=} \max_{k \neq \ell} \frac{|\mathbf{d}_k^* \mathbf{d}_\ell|}{\|\mathbf{d}_k\|_2 \|\mathbf{d}_\ell\|_2} \quad (4)$$

is the mutual coherence of the dictionary and \mathbf{d}_k and \mathbf{d}_ℓ are the columns of matrix \mathbf{D} [3]. Research in this area has also focused on finding the best overcomplete dictionary in order to have sparse representations [6]–[8]. We refer the interested reader to the Elad’s book [9] for further details.

Finding sparse representations is also related to the theory of finite rate of innovation (FRI) signals [10]–[12]. This framework considers the problem of sampling and reconstructing non-bandlimited signals and has been applied to several problems such as image super-resolution [13], ultrasound imaging [14] and many others [15]–[21]. A recently proposed algorithm, *ProSparse* [22], solves the sparsity problem with an approach based on Prony’s method [23]. The authors present the algorithm for the overcomplete dictionary built from the Fourier and identity bases and extend it to more general pairs of bases where one of the bases can be expressed in terms of a Vandermonde matrix and the other basis is local, that is, its elements have very few non-zero elements.

Here, we focus on the DCT and Haar pair of bases. We extend the approach presented in [22] and we study the average performance of this algorithm compared to Basis Pursuit. We also analyse the *more underdetermined* scenario where the dictionary is built from the union of a DCT frame and a Haar basis. In this case, Basis Pursuit is heavily penalised because the coherence of the dictionary increases considerably. In both cases, ProSparse shows competitive

performances compared to Basis Pursuit, and in some situations outperforms the latter. We first give a brief overview of ProSparse and then present the proposed extension to the DCT-Haar case. We prove that the algorithm is guaranteed to find the sparsest solutions when the number of non-zero elements is below a threshold that goes beyond the uniqueness bound. We then validate the approach with simulations.

2. SPARSE RECONSTRUCTION BASED ON PRONY'S METHOD

This section summarises the main contributions of [22]. In this section we focus our discussion on the case of the union of Fourier and canonical bases.

2.1. Fourier and canonical bases

When the dictionary is the union of Fourier and identity bases the atoms can be recovered applying a variation of Prony's method [22]. This new algorithm is called *ProSparse*.

Let $\mathbf{D} = [\Phi, \Psi]$ be an overcomplete dictionary, where Φ is a unitary Fourier matrix of size $N \times N$ and Ψ is the identity matrix. \mathbf{x} can be expressed in terms of the elements that correspond to each basis as follows: $\mathbf{x} = [\mathbf{x}_\Phi, \mathbf{x}_\Psi]^T$, where $\mathbf{x}_\Phi, \mathbf{x}_\Psi \in \mathbb{C}^N$. It follows that $\mathbf{y} = \Phi \mathbf{x}_\Phi + \Psi \mathbf{x}_\Psi$. We assume that the vector \mathbf{y} is composed of K_p Fourier atoms and K_q spikes, that is, \mathbf{x}_Φ is K_p -sparse and \mathbf{x}_Ψ is K_q -sparse. The overall sparsity of the vector \mathbf{x} is thus $K = K_p + K_q$. It follows that each element of \mathbf{y} is given by

$$y[n] = \sum_{k=1}^{K_p} \frac{a_k}{\sqrt{N}} e^{j \frac{2\pi}{N} m_k n} + \sum_{k=1}^{K_q} b_k \delta[n - n_k], \quad 0 \leq n < N, \quad (5)$$

where $a_k, b_k \in \mathbb{C} \setminus \{0\}$, $0 \leq m_1 < \dots < m_{K_p} < N$ and $0 \leq n_1 < \dots < n_{K_q} < N$. The goal is to recover the two sets of parameters $\{(a_k, m_k)\}_{k=1}^{K_p}$ and $\{(b_k, n_k)\}_{k=1}^{K_q}$ from the observation $\mathbf{y} = \mathbf{D} \mathbf{x}$.

Each Fourier atom contributes to the entire vector \mathbf{y} , but since the spikes are localised, only K_q samples are influenced by them. We can thus search for *clean* intervals of consecutive samples that are only due to the Fourier atoms. For such intervals, the samples of vector \mathbf{y} are given by

$$y[n] = \sum_{k=1}^{K_p} \frac{a_k}{\sqrt{N}} e^{j \frac{2\pi}{N} m_k n}, \quad (6)$$

where we restrict the index n to the clean window. If the clean window contains at least $2K_p$ samples, parameters a_k and m_k can be perfectly recovered from samples $y[n]$ by applying Prony's method [24]–[26]. Prony's method is based on finding the coefficients $h[n]$ of a filter that annihilates the sequence $y[n]$, that is,

$$y[n + K_p] + h[1]y[n + K_p - 1] + \dots + h[K_p]y[n] = 0. \quad (7)$$

The z -transform of the sequence $h[n]$ is given by $H(z) = 1 + \sum_{n=1}^{K_p} h[n]z^{-n} = \prod_{k=1}^{K_p} (1 - u_k z^{-1})$, where $u_k = e^{j2\pi m_k/N}$ are the zeros of $H(z)$. If the clean window has at least $2K_p$ samples, the coefficients $h[n]$ can be obtained by solving a linear system of equations given by (7). The locations of the non-zero elements, m_k , are then obtained from the roots of $H(z)$. The amplitudes a_k can then be obtained from (6), since the parameters m_k are now known. Therefore, with Prony's method we can reconstruct \mathbf{x}_Φ . The spikes' locations and amplitudes are then obtained from the residual $\mathbf{y} - \Phi \mathbf{x}_\Phi$.

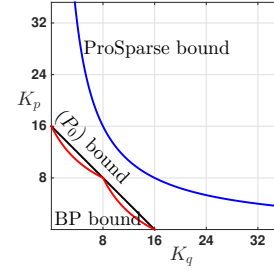


Fig. 1: Comparison of bounds for sparse signal representation in union of Fourier and canonical bases for $N = 256$.

Note that the Fourier atoms in \mathbf{y} are periodic with period N samples, and therefore the search of the clean window can be performed modulo N , that is, for a window of length L located at index n we access the elements of vector \mathbf{y} as: $(y[(n + \ell) \bmod (N)])_{\ell=0}^{L-1}$. Moreover, if the reconstruction algorithm fails to find a clean section of at least $2K_p$ samples in \mathbf{y} , we can also apply the same strategy to the discrete Fourier transform of \mathbf{y} : $\hat{\mathbf{y}} = \text{DFT}_N \{\mathbf{y}\}$. In this dual case, the spikes become complex exponentials and thus \mathbf{x}_Ψ is reconstructed by applying Prony's method to a window of at least $2K_q$ samples.

2.2. ProSparse

For the sake of completeness, we state here the key proposition related to ProSparse.

Proposition 1. [22] Assume $\mathbf{D} = [\mathbf{F}, \mathbf{I}]$ and let $\mathbf{y} \in \mathbb{C}^N$ be an arbitrary signal. There exists an algorithm, with a worst-case complexity of $\mathcal{O}(N^3)$, that finds all (K_p, K_q) -sparse signals \mathbf{x} such that $\mathbf{y} = \mathbf{D} \mathbf{x}$ and $K_p K_q < N/2$.

Note that this bound is much weaker than the uniqueness bound (3) and the ℓ_1 bound which for the case of Fourier and identity is approximately given by $K_p + K_q < 0.91 \sqrt{N}$ [3]. These bounds are illustrated in Figure 1.

3. EXTENSION TO THE UNION OF DCT AND HAAR BASES

ProSparse can be extended to other pairs of bases. For instance, it can also be applied to the case where Φ is a DCT matrix and Ψ is a *local* basis such as a Haar matrix with one decomposition level. The ℓ -th column of the DCT matrix of size $N \times N$ is given by

$$\phi_\ell[n] = \begin{cases} 1/\sqrt{N} & \text{for } \ell = 0, \\ \sqrt{\frac{2}{N}} \cos[\frac{\pi}{N}(n + 0.5)\ell] & \text{for } 1 \leq \ell < N, \end{cases} \quad (8)$$

where n corresponds to the row index and goes from 0 to $N-1$. Note that $\{\phi_\ell\}_{\ell=0}^{N-1}$ form an orthonormal basis of the Euclidean space \mathbb{R}^N and therefore the matrix Φ is orthogonal. The Haar orthogonal matrix of size $N \times N$ with one decomposition level is given by

$$\Psi = \frac{1}{\sqrt{2}} \left[\mathbf{I}_{N/2} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{I}_{N/2} \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right] \quad (9)$$

where $\mathbf{I}_{N/2}$ is the identity matrix of size $N/2$ and \otimes denotes the Kronecker product.

The DCT atoms can be expressed as the sum of two complex exponentials and thus be recovered with Prony's method. However, in order to reconstruct the entire sparse vector $\mathbf{x} = [\mathbf{x}_{\Phi}^T, \mathbf{x}_{\Psi}^T]^T$ we also require Ψ to be local. This guarantees that we will be able to find clean windows where we can apply Prony's method provided that the sparsity of \mathbf{x} is not too large. By local we mean that a non-zero element in \mathbf{x}_{Ψ} influences a small number of elements in vector \mathbf{y} . The ideal local basis is the identity or canonical basis, where each atom contributes to a single element in \mathbf{y} . For the Haar matrix given as in (9), each atom has effect on two consecutive samples of \mathbf{y} . In what follows we propose an extension of Prony's method for DCT atoms and improve ProSparse's bound for the union of DCT and Haar bases.

3.1. Prony's with complex conjugate roots on the unit circle

The DCT atoms can be expressed as two complex exponentials¹:

$$\phi_{\ell}[n] = \sqrt{\frac{1}{2N}} \left(e^{j\frac{\pi\ell}{2N}} e^{j\frac{2\pi}{2N}\ell n} + e^{-j\frac{\pi\ell}{2N}} e^{-j\frac{2\pi}{2N}\ell n} \right). \quad (10)$$

In this case, a naive application of Prony's method can estimate K_p DCT atoms from $2K_p$ complex exponentials, and therefore would require $4K_p$ consecutive clean samples. However, since the exponentials are in complex conjugate pairs this number can be reduced.

Lemma 1. *Let $y[n] = \sum_{k=1}^K (a_k e^{j\alpha_k n} + b_k e^{-j\alpha_k n})$. The α_k parameters are uniquely determined from $3K$ consecutive samples.*

Proof. The filter that annihilates this sequence has $2K + 1$ taps: $H(z) = \prod_{k=1}^K (1 - e^{j\alpha_k} z^{-1})(1 - e^{-j\alpha_k} z^{-1})$. Since the roots of $H(z)$ appear in complex conjugate pairs, this can also be written as $H(z) = \prod_{k=1}^K (1 - 2 \cos \alpha_k z^{-1} + z^{-2})$. The coefficients of $H(z)$ satisfy: $h_l = h_{2K-l}$, for $l = 0, 1, \dots, K-1$. Moreover, $h_0 = h_{2K} = 1$. Thus, we only need to estimate K coefficients. This can be done by solving the following system:

$$\begin{bmatrix} y_{2K} + y_0 & \cdots & y_{K+1} + y_{K-1} & y_K \\ y_{2K+1} + y_1 & \cdots & y_{K+2} + y_K & y_{K+1} \\ \vdots & \ddots & \vdots & \vdots \\ y_{3K-1} + y_{K-1} & \cdots & y_{2K} + y_{2K-2} & y_{2K-1} \end{bmatrix} \begin{bmatrix} 1 \\ h_1 \\ \vdots \\ h_K \end{bmatrix} = 0.$$

If all α_k are distinct, this system has a unique solution. \square

3.2. Making the problem circular

The underlying periodicity of the DCT atoms is $2N$ samples, but vector \mathbf{y} is of length N . However, we can take advantage of the even symmetry of the cosine function to extend the vector \mathbf{y} to a vector of $2N$ samples. From (10) it follows that $\phi_{\ell}[n] = \phi_{\ell}[2N - 1 - n]$. We can therefore build the following extended vector of size $2N$:

$$y_{\text{ext}}[n] = \begin{cases} y[n] & \text{for } 0 \leq n < N, \\ y[2N - 1 - n] & \text{for } N \leq n < 2N, \end{cases} \quad (11)$$

where each atom that comes from the DCT elements in \mathbf{x}_{Φ} can be reconstructed by recovering the two exponentials that appear in complex conjugate pairs. Since the common period of these complex exponentials is $2N$, we can perform the search of the clean window by

¹For $\ell = 0$ the atom $\phi_0[n]$ has a constant value of $1/\sqrt{2N}$, this matches the decomposition in (10) with an amplitude correction of a factor of $1/\sqrt{2}$.

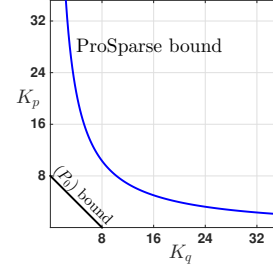


Fig. 2. Comparison of bounds for sparse signal representation in union of DCT and Haar bases for $N = 256$.

accessing the elements of \mathbf{y}_{ext} modulo $2N$. In matricial form we can express the extended vector as follows:

$$\mathbf{y}_{\text{ext}} = \begin{bmatrix} \mathbf{I} \\ \mathbf{J} \end{bmatrix} \mathbf{y} \quad (12)$$

where \mathbf{I} is the identity matrix of size N , and \mathbf{J} is the identity with rows flipped in the up/down direction. Note that with this extension the elements that are due to \mathbf{x}_{Ψ} will influence four elements, two in the first N elements of \mathbf{y}_{ext} and another two in the second half of the extended vector.

3.3. Building the dual vector

Similarly to the Fourier-Identity case, we can build a dual vector where we can also apply Prony's method to reconstruct the sparse vector \mathbf{x} . Since the complex exponential elements in vector \mathbf{y}_{ext} have a common discrete period of $2N$ samples, their DFT are given by Dirac deltas: $\text{DFT}_{2N} \left\{ e^{j\frac{\pi\ell}{2N}} e^{j\frac{2\pi}{2N}\ell n} \right\} = e^{j\frac{\pi\ell}{2N}} \delta[m - \ell]$ and $\text{DFT}_{2N} \left\{ e^{-j\frac{\pi\ell}{2N}} e^{-j\frac{2\pi}{2N}\ell n} \right\} = e^{-j\frac{\pi\ell}{2N}} \delta[m - (2N - \ell)]$. The DCT atoms are thus the *local* atoms in the Fourier transform of \mathbf{y}_{ext} given by $\hat{\mathbf{y}}_{\text{ext}} = \text{DFT}_{2N} \{ \mathbf{y}_{\text{ext}} \}$. As previously described, the Haar atoms in the extended vector \mathbf{y}_{ext} influence four elements. Therefore, when taking the DFT of \mathbf{y}_{ext} , each Haar atom will lead to four complex exponentials. It follows that we need a clean window of at least $8K_q$ samples in order to recover the Haar atoms from $\hat{\mathbf{y}}_{\text{ext}}$.

3.4. ProSparse in the DCT-Haar case

We now denote by \mathbf{U} the DCT matrix given as in (8), \mathbf{H} , the Haar matrix defined as in (9), \mathbf{F} the Fourier matrix and \mathbf{I} and \mathbf{J} the identity and flipped identity as used in (12). Subscripts denote the dimension of these square matrices. Based on the previous derivations we can state the following result:

Proposition 2. *Assume $\mathbf{D} = [\mathbf{U}_N, \mathbf{H}_N]$ and let $\mathbf{y} \in \mathbb{C}^N$ be an arbitrary signal. There exists an algorithm, with a worst-case complexity of $\mathcal{O}(N^3)$, that finds all (K_p, K_q) -sparse signals \mathbf{x} such that*

$$\mathbf{y} = \mathbf{D} \mathbf{x} \quad \text{and} \quad 3K_p K_q + K_q < N. \quad (13)$$

Proof. From the observed signal \mathbf{y} we can build the extended vector \mathbf{y}_{ext} as in (12). We have that

$$\mathbf{y}_{\text{ext}} = \begin{bmatrix} \mathbf{U}_N & \mathbf{H}_N \\ \mathbf{J}_N \mathbf{U}_N & \mathbf{J}_N \mathbf{H}_N \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{U}_N & \mathbf{H}_N \\ \mathbf{J}_N \mathbf{U}_N & \mathbf{J}_N \mathbf{H}_N \end{bmatrix} \begin{bmatrix} \mathbf{x}_U \\ \mathbf{x}_H \end{bmatrix},$$

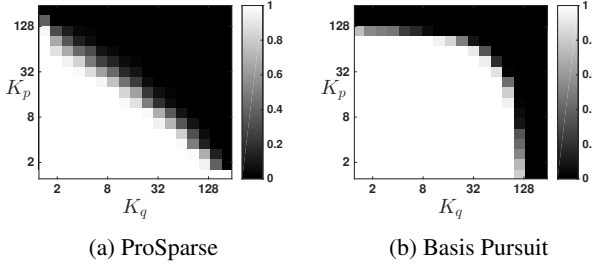


Fig. 3: Probability of success of ProSparse and Basis Pursuit (white represents probability of reconstruction of the sparse vector equal to one and black probability equal to zero). The overcomplete dictionary is given by the union of the DCT and Haar bases with $N = 256$. For each pair (K_p, K_q) , 100 realisations of the sparse vector \mathbf{x} have been generated with the locations of the atoms uniformly distributed and the amplitudes drawn from $\mathcal{N}(0, 1)$.

where \mathbf{x}_U is K_p -sparse and \mathbf{x}_H is K_q -sparse. This is equivalent to

$$\mathbf{y}_{\text{ext}} = [\mathbf{F}_{2N} \mathbf{A}, \mathbf{H}_{2N} \mathbf{B}] \mathbf{x} = [\mathbf{F}_{2N}, \mathbf{H}_{2N}] \begin{bmatrix} \mathbf{x}_F \\ \mathbf{x}_G \end{bmatrix} \quad (14)$$

where $\mathbf{x}_F \stackrel{\text{def}}{=} \mathbf{A} \mathbf{x}_U$, $\mathbf{x}_G \stackrel{\text{def}}{=} \mathbf{B} \mathbf{x}_H$,

$$\mathbf{A} \stackrel{\text{def}}{=} \begin{bmatrix} \sqrt{2} & 0 & \dots & 0 \\ 0 & e^{j \frac{\pi}{2N}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{j \frac{\pi(N-1)}{2N}} \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & e^{-j \frac{\pi(N-1)}{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & e^{-j \frac{\pi}{2N}} & \dots & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{B} \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{I}_{N/2} & \mathbf{0}_{N/2} \\ \mathbf{J}_{N/2} & \mathbf{0}_{N/2} \\ \mathbf{0}_{N/2} & \mathbf{I}_{N/2} \\ \mathbf{0}_{N/2} & \mathbf{J}_{N/2} \end{bmatrix}.$$

It is easy to verify that if \mathbf{x}_U and \mathbf{x}_H are K_p and K_q sparse respectively, then \mathbf{x}_F and \mathbf{x}_G are $2K_p$ and $2K_q$ sparse, with the $2K_p$ Fourier atoms appearing in complex conjugate pairs. The result follows Proposition 3 from [22] and Lemma 1 applied to (14). Note that the size of this equivalent problem is $2N$. \square

Remark 1. This condition is less restrictive than the uniqueness bound (3) for the DCT-Haar case since $\lim_{N \rightarrow \infty} \mu(\mathbf{D}) = 2/\sqrt{N}$ (with $\mu([\mathbf{U}, \mathbf{H}]) \leq 2/\sqrt{N}$). This is illustrated in Figure 2.

4. UNION OF FRAMES

ProSparse can also be extended to the reconstruction of sparse vectors expressed in a union of frames. We now consider the case where the overcomplete dictionary $\mathbf{D} = [\Phi, \Psi]$ is built from a frame Φ and a basis Ψ . Specifically, we consider the case where Φ is given by the first N rows of a DCT matrix of size $2N \times 2N$ and Ψ is built as in (9).

In this setup we cannot apply circularity or duality concepts because the observed vector \mathbf{y} is made of truncated versions of DCT atoms. However, we can still recover these atoms since the frame Ψ is localised and the truncated DCT atoms can still be expressed as the sum of two complex exponentials.

5. SIMULATION RESULTS

We have performed a series of Monte Carlo simulations to measure empirically the probability of success of ProSparse in the case of

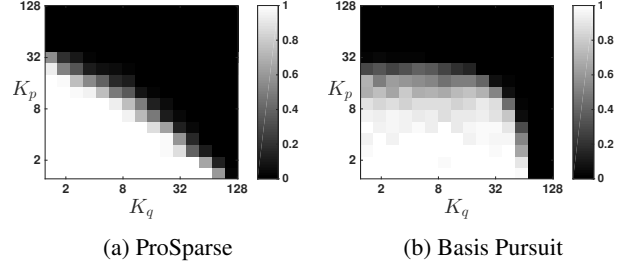


Fig. 4: Probability of success of ProSparse and Basis Pursuit. The overcomplete dictionary is formed by a DCT frame of size $N \times M$ and the Haar frame of size $N \times M$, where $N = 128$ and $M = 256$. For each pair (K_p, K_q) , 100 realisations of the sparse vector \mathbf{x} have been generated with the locations of the atoms uniformly distributed and the amplitudes drawn from $\mathcal{N}(0, 1)$.

union of bases and union of frames and compared its performance to Basis Pursuit. Figure 3 shows the results for the union of bases case and Figure 4 shows the results for the union of frames case. In both cases we have randomly generated 100 different sparse vectors $\mathbf{x} = [\mathbf{x}_\Phi^T, \mathbf{x}_\Psi^T]^T$ for each pair of possible sparsity levels (K_p, K_q) . The x -axis corresponds to the number of non-zero elements in \mathbf{x}_Ψ and the y -axis the number of non-zero elements in \mathbf{x}_Φ . The locations of these non-zero elements are randomly generated at each iteration by picking the first K_p or K_q elements of random permutations of the integers from 0 to $N - 1$ in the case of the union of bases and from 0 to $2N - 1$ in the case of the DCT frame. The amplitudes are then independently drawn from a Gaussian distribution of zero mean and unit variance $\mathcal{N}(0, 1)$.

In the case of the union of bases, ProSparse succeeds when there is a clean window of at least $3K_p$ samples in \mathbf{y}_{ext} or when there is a clean window of at least $8K_q$ samples in $\hat{\mathbf{y}}_{\text{ext}}$. Note that the existence of these clean windows is performed by checking the spike locations via the modulo $2N$ operator. In the case of the union of frames we cannot apply circularity nor duality, and therefore the existence of the clean gap is only checked in \mathbf{y} (not \mathbf{y}_{ext} nor $\hat{\mathbf{y}}_{\text{ext}}$).

The probability of success of Basis Pursuit is measured by recovering the sparse vector \mathbf{x} from the measured vector \mathbf{y} by solving the following optimisation problem: $\arg \min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_1$ s.t. $\mathbf{y} = \mathbf{D} \tilde{\mathbf{x}}$. The optimisation problem has been solved using the convex optimisation toolbox CVX in MATLAB. The success is measured by computing the quantity $\|\tilde{\mathbf{x}} - \mathbf{x}\|_2^2 / \|\mathbf{x}\|_2^2$, where \mathbf{x} is the true sparse vector, and checking that this quantity is below a negligible value of 10^{-5} .

The results show that in the case of the union of frames, Basis Pursuit is heavily penalised and fails to recover the sparse vector for the majority of sparsity levels while ProSparse succeeds in a wider region. In the case of the union of bases, Basis Pursuit performs better when K_p and K_q have similar values, but ProSparse does better in the unbalanced scenarios.

6. CONCLUSIONS

We have presented an extension of a polynomial complexity algorithm, *ProSparse*, to the problem of finding sparse representations in the union of DCT and Haar dictionaries. We have shown that the deterministic bound of this algorithm goes beyond the uniqueness bound of the ℓ_0 optimisation problem and therefore beyond the deterministic bound of Basis Pursuit. We have also performed a numerical average performance analysis that shows that this approach does better than Basis Pursuit in a number of situations.

References

- [1] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by Basis Pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [2] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.
- [3] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2558–2567, Sep. 2002.
- [4] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [5] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [6] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.
- [7] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4311–4322, Nov. 2006.
- [8] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [9] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, 1st ed. Springer, Aug. 2010.
- [10] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with Finite Rate of Innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.
- [11] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, May 2007.
- [12] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, Mar. 2008.
- [13] L. Baboulaz and P. L. Dragotti, "Exact feature extraction using finite rate of innovation principles with an application to image super-resolution," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 281–298, Feb. 2009.
- [14] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1827–1842, Apr. 2011.
- [15] V. Y. F. Tan and V. K. Goyal, "Estimating signals with finite rate of innovation from noisy samples: a stochastic algorithm," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5135–5146, Oct. 2008.
- [16] C. S. Seelamantula and M. Unser, "A generalized sampling method for finite-rate-of-innovation-signal reconstruction," *IEEE Signal Processing Letters*, vol. 15, pp. 813–816, Nov. 2008.
- [17] J. Oñativia, S. Schultz, and P. L. Dragotti, "A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging," *Journal of Neural Engineering*, vol. 10, no. 4, pp. 1–14, Jul. 2013.
- [18] J. Oñativia, J. A. Urigüen, and P. L. Dragotti, "Sequential local FRI sampling of infinite streams of Diracs," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013)*, May 2013, pp. 5440–5444.
- [19] J. A. Urigüen, T. Blu, and P. L. Dragotti, "FRI sampling with arbitrary kernels," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5310–5323, Nov. 2013.
- [20] C. Gilliam and T. Blu, "Fitting instead of annihilation: improved recovery of noisy FRI signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 51–55.
- [21] S. Mulleti and C. S. Seelamantula, "Ellipse fitting using finite rate of innovation principles," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 5824–5828.
- [22] P. L. Dragotti and Y. M. Lu, "On sparse representation in Fourier and local bases," *IEEE Transactions on Information Theory*, In Press (arXiv:1310.6011), 2014.
- [23] P. Stoica and R. Moses, *Spectral Analysis of Signals*, 1st ed. Prentice Hall, Mar. 2005.
- [24] R. Prony, "Essai expérimental et analytique sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures," *Journal de l'École Polytechnique*, vol. 1, no. 22, pp. 24–76, 1795.
- [25] P. Milanfar, G. C. Verghese, W. C. Karl, and A. S. Willsky, "Reconstructing polygons from moments with connections to array processing," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 432–443, Feb. 1995.
- [26] M. Elad, P. Milanfar, and G. H. Golub, "Shape from moments — An estimation theory perspective," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1814–1829, Jul. 2004.