PROSPARSE EXTENSION: PRONY'S BASED SPARSE PATTERN RECOVERY WITH EXTENDED DICTIONARIES

Jun-Jie Huang and Pier Luigi Dragotti

Communications and Signal Processing Group (CSP), Imperial College London, UK

ABSTRACT

ProSparse is a Prony's based method that solves the sparse representation problem of signals in the union of Fourier and canonical bases. By exploiting the structure of the dictionary, ProSparse is able to reconstruct sparse signals beyond the recovery bound of Basis Pursuit. We generalize this framework for a broader class of dictionaries which are still formed from the union of two bases. The proposed algorithm achieves perfect reconstruction over a lower sparsity level than Basis Pursuit in noiseless cases. In the presence of noise, we extend the ProSparse Denoise algorithm to the generalized dictionaries by considering their intrinsic structure. The original ProSparse can be viewed as a special case of our proposed algorithm. From simulation results, our approach outperforms state-of-the-art algorithms.

Index Terms— Sparse representation, Prony's method, union of bases, denoising

1. INTRODUCTION

The sparse representation problem is to estimate a K-sparse vector $\boldsymbol{x} \in \mathbb{C}^M$ from an observation $\boldsymbol{y} \in \mathbb{C}^N$ under dictionary $\boldsymbol{D} \in \mathbb{C}^{N\times M}$ with $M > N$:

$$\boldsymbol{y} = \boldsymbol{D}\boldsymbol{x}, \tag{1}$$

where $||\boldsymbol{x}||_0 \leq K$ with $||\mathbf{x}||_0 \stackrel{\text{def}}{=} \#\{n : |x[n]| \neq 0\}$.

The sparse representation problem is usually solved using convex relaxation methods such as Basis Pursuit (BP) or greedy algorithms such as Orthogonal Matching Pursuit (OMP) methods [1, 2, 3, 4], Compressive Sampling Matching Pursuit (CoSAMP) [5] or Subspace Pursuit (SP) method [6]. And the performance of these methods is pretty well understood [7, 8, 9, 10, 11].

ProSparse (Prony's based sparsity) method [12] considers a classical sparse representation problem where the observed signal is sparse in the union of Fourier and canonical bases. This means that the signal is a combination of exponentials and spikes. ProSparse method recovers a wider range of sparsity patterns than BP [13] in polynomial complexity. Its main idea is to search for clean segments which do not contain spikes and recover the exponentials using Prony's method. The spikes can then be retrieved from the residual. To recover sparse signal from noisy observation, Cadzow denoising [14] is adopted in ProSparse Denoise method [15] to estimate the underlying exponentials. The spikes are iteratively removed from the residual between observation and the estimated Fourier signal. Due to the periodicity property of Fourier transform, a fast version of Cadzow denoising, which is based on fast Fourier Transform (FFT) and non-iterative shrinkage, replaces the traditional iterative Cadzow denoising [14]. Compared with state-of-the-art algorithms [5, 6, 13, 16, 17], ProSparse Denoise offers higher probability of support retrieval and lower complexity. The philosophy behind ProSparse [12, 15, 18, 19] is to fully exploit the structure of the dictionary to achieve low complexity and better performance rather than focusing on the universality of the algorithm. However, the dictionary is still essentially restricted to the union of Fourier and canonical bases. It is essential to enable robust recovery from a broader class of dictionaries with deterministic structure.

In this paper, we extend ProSparse [12, 15] to recover sparse vectors with an extended class of dictionary and achieve perfect reconstruction in noiseless scenario and robust recovery in noisy case. We show that signals which are sparse in a dictionary, which is in the form $\boldsymbol{D} = [\boldsymbol{V}, \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H]$ where $\boldsymbol{V}^H$ is an orthogonal Vandermonde-like matrix, can be recovered using a ProSparse-like algorithm. In noiseless scenario, the exponential components corresponding to the Vandermonde-like matrix is recovered using Prony's method and the rest of sparse signal can be estimated from the residual signal. In noisy case, the exponentials are estimated using Cadzow denoising with a noise pre-whitening step. The exponential frequencies are measured using Prony's method, and amplitudes are recovered with least squares. An iterative spike removal scheme gradually estimate the spikes. The simulation results show our proposed algorithm outperforms state-of-the-art sparse recovery algorithms.

The rest of the paper is organized as follows: Section 2 gives a more detailed introduction about ProSparse [12] and ProSparse Denoise [15]. Section 3 introduces a new class of dictionaries and sparse recovery algorithms for both noiseless

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

and noisy observations. Section 4 presents simulation results and Section 5 draws conclusions.

2. RELATED WORKS

When the dictionary is the union of canonical and Fourier bases $\boldsymbol{D} = [\boldsymbol{I}, \boldsymbol{F}]$, the observation is the sum of K_p spikes and K_q exponentials:

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

where $0 \leq n < N$, $0 \leq n_1 < ... < n_{K_p} < N$, $0 \leq m_1 < ... < m_{K_q} < N$, and a_k, b_k are non-zero weights.

With Prony's method, the K_q exponential frequencies can be extracted from the segment between two spikes which has at least $2K_q$ consecutive entries. The insight of Prony's method is that any $K_q + 1$ consecutive terms on the segment without spikes forms a K_q order difference equation. With the coefficients of this difference equation (called annihilating filter), the exponential components can be extracted through the eigenvalues of its companion matrix. The spikes can then be obtained from the residual signal given by the difference between the observation and the estimated Fourier contribution. The ProSparse algorithm [12] searches for all (K_p, K_q) combinations which fulfills $K_p K_q < N/2$ and gives K_q distinctive exponentials. If $K_q > K_p$, it is computationally more efficient to first retrieve K_p spikes from the dual signal $\overline{\boldsymbol{F}^H \boldsymbol{y}} = [\boldsymbol{F}, \boldsymbol{I}]\overline{\boldsymbol{x}}$ where the spikes are transformed into K_p complex exponentials.

In the presence of noise, ProSparse algorithm [12] will fail as the Prony's method is sensitive to noise. The ProSparse Denoise [15] treats the spikes as noise and applies Cadzow denoising [14] to separate exponentials from noise. The spikes are iteratively removed from the residual signal. At the end, the Fourier atoms are estimated from the denoised signal using Prony's method. Cadzow denoising [14] builds a Toeplitz matrix with all elements of $y[n]$ and make it as square as possible. If there is no noise, the rank of this Toeplitz matrix is still K_q. However, it will become full rank if noise is present. Cadzow denoising iteratively tries to find a matrix belonging to the intersection between the non-convex manifold of low rank matrix and linear space of Toeplitz matrices. The periodicity of Fourier atoms enables a fast circulant Cadzow algorithm which is based on FFT and on finding directly a rank-K_q Toeplitz matrix.

3. PROSPARSE EXTENSIONS

ProSparse [12, 15] exploits the structured property of exponentials, which corresponds to the Fourier basis, in the observed signal. The Fourier basis is a special case of Vandermonde matrices whose rows are powers of a vector. We now show that ProSparse algorithm can be applied to a class of more general dictionary.

Definition: Any matrix $\boldsymbol{A} = \boldsymbol{\Gamma W S}$, where $\boldsymbol{\Gamma} \in \mathbb{C}^{N \times N}$ is a diagonal matrix, $\boldsymbol{W} \in \mathbb{C}^{N \times M}$ is a Vandermonde matrix, and $\boldsymbol{S} \in \mathbb{C}^{M \times N}$ has sparse columns with bounded sparsity B, is called an orthogonal Vandermonde-like matrix when $\boldsymbol{A}^H = \boldsymbol{\Delta A}^{-1}$ where $\boldsymbol{\Delta}$ is a diagonal matrix.

We note that the Vandermonde matrix $\boldsymbol{W}$ may not have the periodic property as the Fourier basis. Let define an indicator $\tau = 0$ for $\boldsymbol{W}$ with periodic property and $\tau = 1$ otherwise.

Proposition 1: Assume $\boldsymbol{D} = [\boldsymbol{V}, \boldsymbol{V \Lambda V}^H]$, where $\boldsymbol{V}^H \in \mathbb{C}^{N \times N}$ is an arbitrary orthogonal Vandermonde-like matrix and $\boldsymbol{\Lambda} \in \mathbb{R}^{N \times N}$ is a diagonal matrix with non-zero diagonals. Let $\boldsymbol{y} \in \mathbb{C}^N$ be an arbitrary signal. A ProSparse-like algorithm is able to find all (K_p, K_q)-sparse signal $\boldsymbol{x} \in \mathbb{C}^{2N}$ with a worst-case complexity of $\mathcal{O}(N^3)$ such that:

$$\begin{aligned} &\boldsymbol{y} = \boldsymbol{D x} \\ \text{and} \quad &BK_q(K_p + \tau) \leq (N - 1 + \tau)/2. \end{aligned} \qquad (3)$$

Proof: We have $\boldsymbol{V}^H = \boldsymbol{\Delta V}^{-1}$ and the dictionary can be factorized as a union of a diagonal matrix $(\boldsymbol{\Gamma \Lambda})^{-1}$ and a matrix $\boldsymbol{WS}$:

$$(\boldsymbol{\Delta \Gamma \Lambda})^{-1} \boldsymbol{V}^H \boldsymbol{D} = [(\boldsymbol{\Gamma \Lambda})^{-1}, \boldsymbol{WS}]. \qquad (4)$$

Given $\boldsymbol{y} = \boldsymbol{Dx}$ and based on Eqn.(4), we have that:

$$(\boldsymbol{\Delta \Gamma \Lambda})^{-1} \boldsymbol{V}^H \boldsymbol{y} = (\boldsymbol{\Gamma \Lambda})^{-1} \boldsymbol{x}_p + \boldsymbol{WS x}_q, \qquad (5)$$

where $\boldsymbol{x} = [\boldsymbol{x}_p^T, \boldsymbol{x}_q^T]^T$, $\boldsymbol{x}_p, \boldsymbol{x}_q \in \mathbb{C}^N$, $||\boldsymbol{x}_p||_0 = K_p$, and $||\boldsymbol{x}_q||_0 = K_q$.

As $\boldsymbol{S x}_q$ is the sum of K_q sparse vectors from $\boldsymbol{S}$, its sparsity is bounded by $||\boldsymbol{S x}_q||_0 \leq BK_q$. Thus the observation is a combination of K_p spikes $(\boldsymbol{\Gamma \Lambda})^{-1} \boldsymbol{x}_p$ and BK_q exponentials $\boldsymbol{WS x}_q$. Prony's method requires a clean window of size at least $2BK_q$ to retrieve BK_q exponentials. Once these are retrieved, the residual consists of K_p non-zero spikes weighted by $(\boldsymbol{\Gamma \Lambda})^{-1}$. Their amplitudes can be recovered by multiplying by $\boldsymbol{\Gamma \Lambda}$.

Due to the possible periodicity of $\boldsymbol{W}$, the K_p spikes split the observed signal into $K_p + \tau$ segments. Let $\mathcal{N}_{\boldsymbol{x}_p}$ be the number of length $2BK_q$ intervals which are clean and d_i the length of i^{th} segment which is the one not corrupted by spikes with $1 \leq i \leq K_p + \tau$. The total length of all clean segments is the difference between the length of the observed signal and the number of spikes and equals to $N - K_p$. The number of length $2BK_q$ intervals within the i^{th} segment is at most $d_i - 2BK_q + 1$ (if $d_i < 2BK_q$, it cannot contain a length $2BK_q$ interval). Thus, $\mathcal{N}_{\boldsymbol{x}_p}$ can be computed as:

$$\begin{aligned} \mathcal{N}_{\boldsymbol{x}_p} \quad &= \sum_{1 \leq i \leq K_p + \tau} max(0, d_i - 2BK_q + 1) \\ &\geq \sum_{1 \leq i \leq K_p + \tau} d_i - 2BK_q + 1 \end{aligned}$$

$$= \left(\sum_{1 \le i \le K_p + \tau} d_i \right) - (K_p + \tau)(2BK_q - 1)$$
$$= N - 2K_qK_p + \tau(1 - 2BK_q). \quad (6)$$

The sufficient and necessary condition for $\mathcal{N}_{\boldsymbol{x}_p} \ge 1$ (i.e. the perfect reconstruction of the sparse signal) is:

$$BK_q(K_p + \tau) \le (N - 1 + \tau)/2. \quad (7)$$

The sufficiency is straightforward. The necessity can be shown by an example with K_p spikes which divide the signal into $K_p + \tau$ segments with equal length. Assuming N can be divided by $K_p+\tau$, $d_i = N/(K_p+\tau)-1$ for $1 \le i \le K_p+\tau$. If $BK_q(K_p + \tau) > (N - 1 + \tau)/2$, it leads to $d_i < 2BK_q$ which implies $\mathcal{N}_{\boldsymbol{x}_p} = 0$. ■

Let us show two examples of the orthogonal Vandermonde-like matrix. When $\boldsymbol{W}$ is the transpose of the discrete Fourier matrix $\boldsymbol{F}^H$ and $\boldsymbol{\Gamma}$, $\boldsymbol{S}$ are identity matrices, $\boldsymbol{V\Lambda V}^H = \boldsymbol{F\Lambda F}^H$ is a circulant matrix. Further, if $\boldsymbol{\Lambda}$ is also an identity matrix, the union of bases $[\boldsymbol{F}, \boldsymbol{I}]$ analyzed in ProSparse [12, 15] can be regarded as a special case of $[\boldsymbol{F}, \boldsymbol{F\Lambda F}^H]$. The DCT matrix can also be written in the form $\boldsymbol{\Gamma W S}$ with $M = 2N$. The diagonal matrix $\boldsymbol{\Gamma} = \frac{1}{\sqrt{2N}} diag\{\frac{1}{\sqrt{2}}, 1, 1, ...1\}$. The rows of the Vandermonde matrix $\boldsymbol{W}$ are powers of vector $\boldsymbol{p} = [p_0, p_1 ..., p_{2N-1}]^T$ with $p_n = e^{-j\pi(n+0.5)/N}$ for $0 \le n < N-1$ and $p_n = e^{j\pi(n+0.5)/N}$ for $N \le n < 2N-1$. The sparse matrix $\boldsymbol{S} = [1, 1]^T \otimes \boldsymbol{I}$, where $\otimes$ is the Kronecker product.

3.1. Noisy Case

For dictionary $\boldsymbol{D} = [\boldsymbol{V}, \boldsymbol{V\Lambda V}^H]$, we propose an algorithm similar to the ProSparse Denoise [15] to recover the sparse signal $\boldsymbol{x} \in \mathbb{C}^{2N}$ from noisy observation $\boldsymbol{y} \in \mathbb{C}^N$:

$$\boldsymbol{y} = \boldsymbol{Dx} + \boldsymbol{\epsilon}, \quad (8)$$

where $\epsilon[n]$ are independent and identical distributed random variables from a normal distribution, for $0 \le n < N$.

Multiplying the noisy observation with $(\boldsymbol{\Delta\Gamma\Lambda})^{-1}\boldsymbol{V}^H$ makes the white noise $\boldsymbol{\epsilon}$ coloured if the diagonal values on $(\boldsymbol{\Delta\Gamma\Lambda})^{-1}$ are not the same. The obtained signal consists of spikes weighted by $(\boldsymbol{\Gamma\Lambda})^{-1}$, sum of exponentials $\boldsymbol{WSx}_q$, and coloured noise $(\boldsymbol{\Delta\Gamma\Lambda})^{-1}\boldsymbol{V}^H\boldsymbol{\epsilon}$:

$$(\boldsymbol{\Delta\Gamma\Lambda})^{-1}\boldsymbol{V}^H\boldsymbol{y} = (\boldsymbol{\Gamma\Lambda})^{-1}(\boldsymbol{x}_p + \boldsymbol{\Delta}^{-1}\boldsymbol{V}^H\boldsymbol{\epsilon}) + \boldsymbol{WSx}_q, \quad (9)$$

where $\boldsymbol{x}_p, \boldsymbol{x}_q \in \mathbb{C}^N$, $||\boldsymbol{x}_p||_0 = K_p$, and $||\boldsymbol{x}_q||_0 = K_q$.

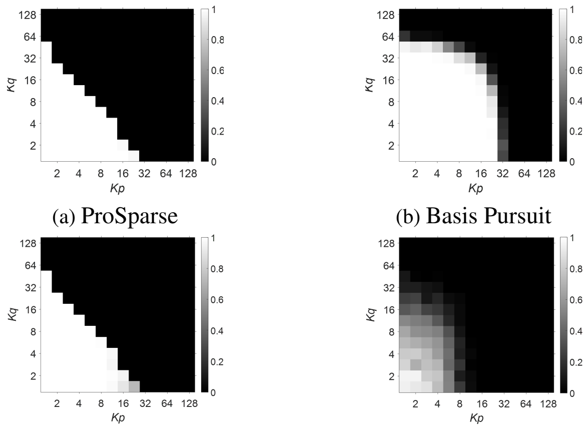
Our strategy is to first separate the structured component, which is the sum of exponentials, from the unstructured one, which consists of spikes and noise. Cadzow denoising [14], which is the noise resilient version of Prony's method, is utilized in [15] to retrieve the exponentials for signal separation. As Cadzow denoising assumes white noise, a pre-whitening process is necessary before exponentials extraction. Cadzow denoising removes noise as well as part of the exponential energy. This energy leakage will be magnified by the diagonal matrix $\boldsymbol{\Gamma\Lambda}$ and affects the spikes estimation. Prony's method is applied to measure the exponential frequencies and the amplitudes of the exponentials are recovered with least squares. A subspace swap [20] would happen when noise level is high. To increase the robustness, Cadzow denoising finds a rank $\widetilde{K}_q$ Toeplitz matrix (with $\widetilde{K}_q > BK_q$) rather than a rank BK_q. And, assuming that the power of the spikes is larger than that of noise and we have a good approximation of the exponentials, the spike locations can be estimated from the residual signal between the observation and the recovered exponentials. The spikes are iteratively removed from the observation and estimated from the residual signal which should be reweighted with $\boldsymbol{\Gamma\Lambda}$. With the obtained K_p spikes and BK_q exponential frequencies, the sparse signal is recovered using least squares. The algorithm for noisy signal recovery is summarized in Algorithm 1.

Algorithm 1 ProSparse Denoise Extension

Input: Noisy observation $\boldsymbol{y}$ and sparsity (K_p, K_q).
Output: (K_p, K_q)-sparsity vector $\widetilde{\boldsymbol{x}} = [\boldsymbol{x}_p^T, \boldsymbol{x}_q^T]^T$.

```
Initialize x_p = 0, indices Ω = {0, 1, ..., N − 1}
Let y_E = (ΔΓΛ)^{-1} V^H y and K̃_q = max(16, 2BK_q).
Denoise y'_E = Cadzow_QSVD(y_E, K̃_q).
Estimate x_q = Prony(y'_E, K̃_q).
for i=1 to K_p do
    Compute residual r = ΓΛ(y_E − y'_E).
    Find spike location n = argmax_{n∈Ω} {|r[n]|}.
    Update Ω ← Ω/{n}.
    Store spike x_p[n] = r[n].
    Remove spike y'_E = y_E − (ΓΛ)^{-1} x_p.
    Denoise y'_E = Cadzow_QSVD(y'_E, K̃_q K_q).
    Estimate x_q = Prony(y'_E, K̃_q).
end for
Retain BK_q atoms with largest amplitudes on x_q.
Obtain x̃ with least squares.
```

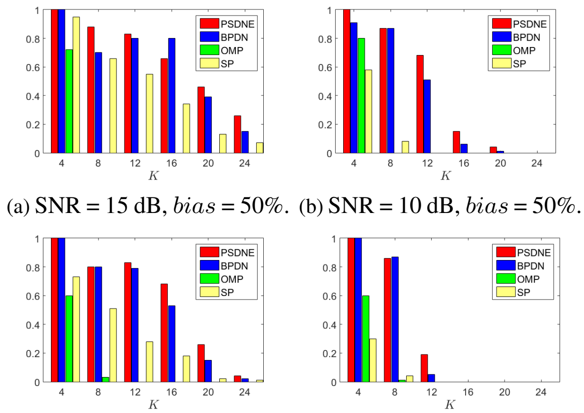
In Cadzow denoising, a Toeplitz matrix $\widetilde{\boldsymbol{S}} = \boldsymbol{S} + \boldsymbol{E}$ is built from noisy observations where $\boldsymbol{S}$ is the desired structured low rank matrix and $\boldsymbol{E}$ corresponds to the noise $(\boldsymbol{\Delta\Gamma\Lambda})^{-1}\boldsymbol{V}^H\boldsymbol{\epsilon}$. As discussed above, the coloured noise leads to $\mathbf{R_E} := \mathrm{E}[\boldsymbol{E}^H\boldsymbol{E}] \ne a\mathbf{I}$. The pre-whiten method used in [21, 22] is applied to enhance the algorithm's robustness to noise. The pre-whitening is realized by replacing the truncated SVD with the truncated quotient singular value decomposition (QSVD) of the pair $(\widetilde{\boldsymbol{S}}, \mathbf{W}_{\mathrm{E}}^{-1})$ during Cadzow denoising where $\mathbf{W}_{\mathrm{E}} := \mathbf{R}_E^{\dagger/2}$. This pre-whitening approach



(a) ProSparse (b) Basis Pursuit

(c) ProSparse (d) Basis Pursuit

Fig. 1: Probability of success for ProSparse extension and Basis Pursuit for $N = 128$ (white corresponds to a high probability of signal recovery, and black corresponds to a low probability). In case one (results shown in (a) and (b)), the Vandermonde-like matrix $\boldsymbol{V}^H$ is set to be the transpose of the discrete Fourier matrix $\boldsymbol{F}^H$ and the diagonal values of $\boldsymbol{\Lambda}$ are non-zero and drawn from a normal distribution $\mathcal{N}(0, 1)$ with zero mean and unit variance. In the second case (results shown in ((c) and (d)), the Vandermonde-like matrix is the same as in case one, while the circulant matrix is obtained by performing circulant shifts of the first row which is a Gaussian signal with variance $\sigma^2 = 8$.



(a) SNR = 15 dB, $bias = 50\%$. (b) SNR = 10 dB, $bias = 50\%$.

(c) SNR = 15 dB, $bias = 25\%$. (d) SNR = 10 dB, $bias = 25\%$.

Fig. 2: Probability of support retrieval for union of Fourier basis and a circulant matrix, each of size $N \times N$ with $N = 256$. The sparsity (K_p, K_q) is specified as $K_p = bias \cdot K$ and $K_q = (1 - bias) \cdot K$. 100 realizations per sparsity level are generated.

leads to more accurate results compared with multiplying directly $\widetilde{\boldsymbol{S}}$ with pre-whiten matrix $\mathbf{W}_{\mathrm{E}}$.

4. SIMULATION RESULTS

For noiseless scenarios, Monte Carlo simulations have been applied to measure the probability of success for the ProSparse algorithm and compared against BP as shown in Figure 1. The Vandermonde-like matrix $\boldsymbol{V}^H$ was selected as $\boldsymbol{F}^H$. Thus, the dictionary $[\boldsymbol{V}, \boldsymbol{V\Lambda V}^H]$ is the union of the Fourier basis and a circulant matrix which is defined either using a random diagonal matrix $\boldsymbol{\Lambda}$ or by defining its first row and then computing its circulant shifts. For each sparsity level (K_p, K_q), 100 sparse vectors have been randomly generated whose non-zero amplitudes are independently drawn from a normal distribution $\mathcal{N}(0, 1)$. If the diagonal matrix is randomly generated, BP outperforms ProSparse extension in the region with similar K_p and K_q sparsity level. In the second case, ProSparse achieves high probability of exact recovery within the determined bound, while BP performs poorly.

For noisy scenarios, our proposed ProSparse Denoise Extension (PSDNE) algorithm has been compared with state-of-the-art algorithms including BPDN, OMP, and SP. We applied CVX package for BPDN realization; OMP has been implemented and the source code of SP was downloaded from the authors' website. We consider a difficult setting where the dictionary is the union of Fourier base $\boldsymbol{F}$ and a circulant matrix $\boldsymbol{C}$. The circulant matrix is defined as $\boldsymbol{C} = \boldsymbol{G} + \boldsymbol{I}$ so that the diagonal matrix does not have zero diagonal values where $\boldsymbol{G}$ is a Gaussian circulant matrix with variance $\sigma^2 = 8$, and $\boldsymbol{I}$ is an identity matrix. Figure 2 shows the probability of support retrieval for all algorithms in 4 different scenarios where the SNR is 15 dB and 10 dB, and the bias (i.e. the ratio of K_p/K_q) is 50% and 25%, respectively. From the simulation results, PSDNE algorithm outperforms other state-of-the-art algorithms in most cases and its probability of exact support recovery gradually drops when K grows. BPDN method provides overall the second best performance. OMP method works at low sparsity level and experiences a sudden drop when K is higher than 4. SP method only performs well at higher SNR.

5. CONCLUSIONS

We extended the ProSparse framework for sparse recovery in both noiseless and noisy scenarios. The generalized ProSparse algorithm can work with a broader class of dictionaries with which the observed signal can be factorized as the sum of exponentials and spikes. In the noiseless case, the exponential frequencies are estimated from a retrieved clean segment using Prony's method. The spikes can further be recovered from the residual signal. In the noisy case, the proposed ProSparse Denoise Extension algorithm handles the coloured noise with the Cadzow denoising based on QSVD. In general, the ProSparse extension algorithm outperforms state-of-the-art sparse recovery algorithms in reconstruction accuracy as the structure of the dictionary has been fully explored.

6. REFERENCES

[1] Stéphane G Mallat and Zhifeng Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on signal processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[2] Deanna Needell and Roman Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Foundations of computational mathematics*, vol. 9, no. 3, pp. 317–334, 2009.

[3] Joel A Tropp and Anna C Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on information theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[4] David L Donoho, Yaakov Tsaig, Iddo Drori, and Jean-Luc Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1094–1121, 2012.

[5] Deanna Needell and Joel A Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[6] Wei Dai and Olgica Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[7] David L Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[8] Emmanuel J Candès, Justin Romberg, and Terence Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.

[9] Emmanuel J Candès and Michael B Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[10] David L Donoho and Xiaoming Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[11] Michael Elad, "Sparse and redundant representations: From theory to applications in signal and image processing," 2010.

[12] Pier Luigi Dragotti and Yue M Lu, "On sparse representation in Fourier and local bases," *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7888–7899, 2014.

[13] Scott Shaobing Chen, David L Donoho, and Michael A Saunders, "Atomic decomposition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.

[14] James A Cadzow, "Signal enhancement-a composite property mapping algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 1, pp. 49–62, 1988.

[15] Jon Onativia, Yue M Lu, and Pier Luigi Dragotti, "Prosparse denoise: Prony's based sparse pattern recovery in the presence of noise," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 4084–4088.

[16] Robert Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[17] Yagyensh Chandra Pati, Ramin Rezaiifar, and PS Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, 1993, pp. 40–44.

[18] Jon Onativia, Yue M Lu, and Pier Luigi Dragotti, "Sparsity pattern recovery using FRI methods," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 5967–5971.

[19] Jon Onativia, Pier Luigi Dragotti, and Yue M Lu, "Sparsity according to prony, average performance analysis," *def*, vol. 50, pp. 0, 2015.

[20] Xiaoyao Wei and Pier Luigi Dragotti, "Guaranteed performance in the FRI setting," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1661–1665, 2015.

[21] Jose Antonio Urigüen, Thierry Blu, and Pier Luigi Dragotti, "FRI sampling with arbitrary kernels," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5310–5323, 2013.

[22] Søren Holdt Jensen, Per Christian Hansen, Steffen Duus Hansen, and John Aasted Sorensen, "Reduction of broad-band noise in speech by truncated QSVD," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 6, pp. 439–448, 1995.