DETECTION OF SPARSE SIGNALS UNDER FINITE-ALPHABET CONSTRAINTS

*Zhi Tian** *Geert Leus*† *Vincenzo Lottici*‡

*Dept. of Electrical & Computer Engineering, Michigan Technological University, Houghton, MI, USA
†Faculty of Electrical Engineering, Delft University of Technology, Delft, The Netherlands
‡Department of Information Engineering, University of Pisa, Pisa, Italy

ABSTRACT

In this paper, we solve the problem of detecting the entries of a sparse finite-alphabet signal from a limited amount of data, for instance obtained by compressive sampling. While existing methods either rely on the sparsity property, the finite-alphabet property, or none of those properties to solve the under-determined system of linear equations, we capitalize on both the sparsity and the finite-alphabet features of the signal. The problem is first formulated in a Bayesian framework to incorporate the prior knowledge of sparsity, which is then shown to be solvable using sphere decoding (SD) or semi-definite relaxation (SDR) for efficient Boolean programming. A few toy simulations show how our method can outperform existing works.

***Index Terms*—** compressed sensing, sparsity, finite alphabet, sphere decoding (SD)

1. INTRODUCTION

Currently, there is a great interest in a range of detection problems where only a reduced set of data samples. e.g., obtained from compressive sampling, is available to detect every entry of a sparse finite-alphabet signal, e.g., a signal containing entries in $\{0, 1\}$. Such problems appear in many fields such as localization of multiple emitters/targets in a geographical area, spectrum sensing of active users in a wide spectrum band, object detection in imaging, and binary symbol detection in digital communications. Note that the considered problem differs from existing work on the detection of sparse signals, where the reduced set of samples is merely used to find out whether a sparse signal is present or not [1].

One way to deal with the considered problem is to view it as an estimation problem and employ for instance a standard minimum mean square error (MMSE) estimator followed by a decision device. However, since the system of linear equations is under-determined, such a method generally does not perform very well.

To cope with the limited amount of data samples, we have to capitalize on the sparsity and/or the finite-alphabet property of the signal. Traditional approaches only rely on a single one of those properties. l_1-norm minimization for instance only exploits the sparsity of the signal [2], whereas generalized sphere decoding is developed to solve an under-determined integer least squares problem for non-sparse constant modulus signals [4]. The novel contribution of this work is that we build on both the sparsity and the finite-alphabet features to detect the entries of the signal. The new problem formulation is deduced from the Bayesian framework, where a priori knowledge of the sparsity property can be incorporated into the objective function. It can then be cast into either a sphere decoding (SD) problem or a semi-definite relaxation (SDR) problem, both of which yield near-optimal estimates of finite-alphabet signals at affordable polynomial complexity. Some toy simulations illustrate the improved performance of the proposed method over existing techniques.

2. SIGNAL MODEL

Consider a system where a sparse finite-alphabet signal $\mathbf{s} \in \{0, 1\}^N$ with $M \ll N$ nonzero elements is transformed into a real-valued output vector $\mathbf{x} \in \mathcal{R}^N$ as

$$\mathbf{x} = \mathbf{\Psi s} + \mathbf{v} \tag{1}$$

where $\mathbf{\Psi} \in \mathcal{R}^{N \times N}$ is the transformation matrix and $\mathbf{v} \in \mathcal{R}^N$ is a random noise vector. Further, we assume that only $K \ll N$ linear observations of $\mathbf{x}$ are available, e.g., obtained through compressive sampling:

$$\mathbf{y} = \mathbf{\Phi x} = \mathbf{\Phi \Psi s} + \mathbf{\Phi v} = \mathbf{Hs} + \mathbf{w} \tag{2}$$

where $\mathbf{\Phi} \in \mathcal{R}^{K \times N}$ is the compressive sampling matrix. Complex-valued linear systems can be expressed by (2) after standard transformation. Now it is clear that if $K < M \ll N$, it will be difficult to reconstruct $\mathbf{s}$; on the other hand, if $M \leq K \ll N$, (2) is an under-determined system of linear equations that might be solvable if both the sparsity and finite-alphabet constraints are exploited.

The above system model appears in several signal processing applications, including signal "on-off" state detection and binary symbol demodulation in digital communications. These examples will be elaborated in Section 5.

Zhi Tian is supported in part by the US NSF grant #CCF-0238174. Geert Leus is supported in part by NWO-STW under the VIDI program (DTC.6577).

3. SIGNAL DETECTION TECHNIQUES

In this section, we give a brief overview of a few existing approaches that could be adopted to extract $\mathbf{s}$ from $\mathbf{y}$ in (2).

3.1. Minimum Mean Square Error (MMSE) Estimator

Suppose that $\mathbf{w}$ is white Gaussian noise with zero mean and covariance $\sigma_w^2\mathbf{I}$. In that case, we can solve (2) by an MMSE estimator followed by a thresholding decision device:

$$\tilde{\mathbf{s}} = (\mathbf{H}^T\mathbf{H} + \sigma_w^2\mathbf{I})^{-1}\mathbf{H}^T\mathbf{y} \quad (3a)$$

$$\hat{\mathbf{s}} = (\tilde{\mathbf{s}} \geq \zeta); \qquad \zeta = 0.5. \quad (3b)$$

Here, ζ is a decision threshold set to a nominal value 0.5. It is possible to choose ζ based on some applicable detection principle, e.g., the Neyman-Pearson rule for binary hypothesis tests.

The MMSE estimator yields a linear receiver with relatively low complexity. However, it does not perform well for under-determined systems. Clearly, we need to exploit the sparsity and/or finite-alphabet properties of $\mathbf{s}$ to improve the detection performance. The next two subsections review some existing methods that could be used in this respect.

3.2. l_1-norm Convex Optimization

To cope with under-determined systems, the sparsity property of $\mathbf{s}$ can be capitalized through an l_1-norm minimization formulation as follows [2]:

$$\tilde{\mathbf{s}}: \quad \min_{\tilde{\mathbf{s}} \in \mathcal{R}^N} ||\tilde{\mathbf{s}}||_1, \quad s.t. \;\; \mathbf{y} = \mathbf{H}\tilde{\mathbf{s}} \quad (4a)$$

$$\hat{\mathbf{s}} = (\tilde{\mathbf{s}} \geq \zeta); \qquad \zeta = 0.5. \quad (4b)$$

The linear constraint in (4a) yields simple convex linear programming which is also termed the Basis Pursuit (BP) method [3], but reduces the robustness of this method against additive noise. Noise resilience can be improved by adopting a quadratic constraint $||\mathbf{y} - \mathbf{H}\mathbf{s}||_2^2 \leq \delta$ in (4a) for a proper value of δ, yielding the BP de-noising (BPDN) method [3]. In both methods, the finite-alphabet nature of $\mathbf{s}$ is overlooked.

3.3. Generalized Sphere Decoding

Sphere decoding (SD) is a computationally-efficient search algorithm for solving an integer least squares problem. When the linear system is under-determined, several versions of generalized SD (GSD) have been developed [4]. For a non-sparse constant modulus signal $\mathbf{b} \in \{-1, 1\}^N$ passing through the same linear system $\mathbf{y} = \mathbf{H}\mathbf{b} + \mathbf{w}$ as in (2), one approach is to regularize the rank-deficient matrix $\mathbf{H}^T\mathbf{H}$ by adding a small diagonal loading term $\epsilon\mathbf{I}$, which amounts to imposing an l_2-norm constraint on $\mathbf{b}$, as follows [4]:

$$\min_{\mathbf{b} \in \{-1,1\}^N} ||\mathbf{y} - \mathbf{H}\mathbf{b}||_2^2 + \epsilon||\mathbf{b}||_2^2. \quad (5)$$

It is shown in [4] that (5) can be transformed into the following equivalent form:

$$\hat{\mathbf{b}} = \arg \min_{\mathbf{b} \in \{-1,1\}^N} ||\mathbf{R}(\boldsymbol{\rho} - \mathbf{b})||_2^2. \quad (6)$$

where $\boldsymbol{\rho} := \mathbf{G}^{-1}\mathbf{H}^T\mathbf{y}$ with $\mathbf{G} := \mathbf{H}^T\mathbf{H} + \epsilon\mathbf{I}$, and $\mathbf{R}$ is an upper triangular matrix satisfying $\mathbf{G} = \mathbf{R}^T\mathbf{R}$. Because $\mathbf{R}$ is now of full rank for any $\epsilon > 0$, the standard SD search steps applies directly on (6). This GSD algorithm utilizes the finite-alphabet constraint on $\mathbf{b}$ to search for a near-optimal solution at polynomial complexity. Nevertheless, its detection performance for under-determined systems exhibits an inevitable gap from full-rank systems.

4. SPARSE SIGNAL RECOVERY UNDER FINITE-ALPHABET CONSTRAINT

We aim to utilize both the sparse nature and the finite alphabet property of the input vector $\mathbf{s}$ to derive an accurate signal recovery algorithm with polynomial computational complexity. The sparsity property can be viewed as *a priori* knowledge and incorporated into the objective function under the Bayesian framework. Meanwhile, the finite-alphabet property constrains the search space for $\mathbf{s}$ on a lattice.

In the absence of the finite-alphabet constraint, a sparsity-constrained optimization formulation for recovering $\mathbf{s}$ can be expressed as[1]

$$\min_{\mathbf{s} \in \mathcal{R}^N} ||\mathbf{y} - \mathbf{H}\mathbf{s}||_2^2 + \lambda||\mathbf{s}||_l^l. \quad (7)$$

Sparsity is induced by the l-norm penalty term $\lambda||\mathbf{s}||_l^l$ for $l \in [0, 2)$, while exact sparsity corresponds to $l = 0$ [2, 3]. Setting $l = 1$ and viewing λ as a Lagrange multiplier, (7) subsumes several noted l_1-regularization algorithms, including LASSO [6] and BPDN [3]. When $l = 2$, (7) no longer results in a sparse representation of $\mathbf{s}$, but the expression resembles to that of GSD in (5) except for the different search spaces.

For $\mathbf{s} \in \{0, 1\}^N$, it holds that

$$||\mathbf{s}||_0 = ||\mathbf{s}||_1 = ||\mathbf{s}||_2^2; \quad \text{and} \quad ||\mathbf{s}||_1 = \mathbf{s}^T\mathbf{1} = \mathbf{1}^T\mathbf{s}. \quad (8)$$

Note that the nonlinear, non-differentiable l_1-norm function reduces to a linear form, which is amenable to gradient-based algorithms. Setting $l = 1$ and choosing $0 < \epsilon \leq \lambda$, we rewrite (7) as

$$\min_{\mathbf{s} \in \{0,1\}^N} ||\mathbf{y} - \mathbf{H}\mathbf{s}||_2^2 + \frac{\lambda - \epsilon}{2}\left(\mathbf{s}^T\mathbf{1} + \mathbf{1}^T\mathbf{s}\right) + \epsilon||\mathbf{s}||_2^2. \quad (9)$$

In (9), the second term is a linear representation of the sparsity-inducing l_1-norm on $\mathbf{s}$, and the third term of the l_2-norm is useful in regularizing the rank-deficiency issue of the measurement matrix $\mathbf{H}$. Our goal next is to develop computationally-efficient algorithms that solve (9).

[1] With $l = 1$, this is an MAP formulation assuming Gaussian noise and sparsity-inducing Laplace prior on $\mathbf{s}$ [5].

4.1. Algorithm 1: Sphere Decoding

Our first approach is to re-formulate (9) such that it is conducive to the use of SD, even when $\mathbf{H}$ is rank deficient. Adopting the definitions $\mathbf{G} := \mathbf{H}^T\mathbf{H} + \epsilon\mathbf{I}$ and $\mathbf{G} = \mathbf{R}^T\mathbf{R}$ as in (6), we re-write the objective function in (9) as follows:

$$\begin{aligned} J(\mathbf{s}) &:= ||\mathbf{y} - \mathbf{H}\mathbf{s}||_2^2 + \frac{\lambda - \epsilon}{2}\left(\mathbf{s}^T\mathbf{1} + \mathbf{1}^T\mathbf{s}\right) + \epsilon||\mathbf{s}||_2^2 \\ &= \mathbf{s}^T(\mathbf{H}^T\mathbf{H} + \epsilon\mathbf{I})\mathbf{s} - (\mathbf{y}^T\mathbf{H} - \frac{\lambda-\epsilon}{2}\mathbf{1}^T)\mathbf{s} \\ &\qquad -\mathbf{s}^T(\mathbf{H}^T\mathbf{y} - \frac{\lambda-\epsilon}{2}\mathbf{1}) + \mathbf{y}^T\mathbf{y} \\ &= \mathbf{s}^T\mathbf{R}^T\mathbf{R}\mathbf{s} - (\mathbf{y}^T\mathbf{H} - \frac{\lambda-\epsilon}{2}\mathbf{1}^T)\mathbf{G}^{-1}\mathbf{R}^T\mathbf{R}\mathbf{s} \\ &\qquad -\mathbf{s}^T\mathbf{R}^T\mathbf{R}\mathbf{G}^{-1}(\mathbf{H}^T\mathbf{y} - \frac{\lambda-\epsilon}{2}\mathbf{1}) + \mathbf{y}^T\mathbf{y} \\ &= ||\mathbf{R}(\boldsymbol{\rho}_\lambda - \mathbf{s})||_2 + C, \qquad \mathbf{s} \in \{0,1\}^N \qquad (10) \end{aligned}$$

where $\boldsymbol{\rho}_\lambda := \mathbf{G}^{-1}(\mathbf{H}^T\mathbf{y} - \frac{\lambda-\epsilon}{2}\mathbf{1})$ and $C := \mathbf{y}^T\mathbf{y} - \boldsymbol{\rho}_\lambda^T\mathbf{G}\boldsymbol{\rho}_\lambda$.

The standard SD algorithm can now be employed to search for $\mathbf{s}$ that minimizes (10). In view of the sparsity-related terms in the objective function, we term this algorithm SD-CS (SD with Compressive Sampling).

4.2. Algorithm 2: Semi-Definite Relaxation

Alternatively, we express (9) in a quadratic form as follows:

$$\begin{aligned} J(\mathbf{s}) &= ||\mathbf{y} - \mathbf{H}\mathbf{s}||_2^2 + \frac{\lambda - \epsilon}{2}\left(\mathbf{s}^T\mathbf{1} + \mathbf{1}^T\mathbf{s}\right) + \epsilon||\mathbf{s}||_2^2 \\ &= \begin{bmatrix}\mathbf{s}\\ 1\end{bmatrix}^T \underbrace{\begin{bmatrix}\mathbf{H}^T\mathbf{H} + \epsilon\mathbf{I} & -\mathbf{H}^T\mathbf{y} + \frac{\lambda-\epsilon}{2}\mathbf{1} \\ -\mathbf{y}^T\mathbf{H} + \frac{\lambda-\epsilon}{2}\mathbf{1}^T & \mathbf{y}^T\mathbf{y}\end{bmatrix}}_{\tilde{\mathbf{Q}}_s} \underbrace{\begin{bmatrix}\mathbf{s}\\ 1\end{bmatrix}}_{\tilde{\mathbf{s}}} \\ &= \tilde{\mathbf{s}}^T\tilde{\mathbf{Q}}_s\tilde{\mathbf{s}}, \qquad \tilde{\mathbf{s}} \in \{0,1\}^{(N+1)}, \ \tilde{\mathbf{s}}_{N+1} = 1 \qquad (11) \end{aligned}$$

where $\tilde{\mathbf{Q}}_s$ is a positive-semidefinite matrix for any $\epsilon \geq 0$.

The form in (11) is a Boolean quadratic programming problem which permits several efficient algorithms, such as the semi-definite relaxation (SDR) method [7]. To employ SDR, we express $J(\mathbf{s})$ as a function of $\mathbf{b} := 2\mathbf{s} - \mathbf{1} \in \{-1,1\}^N$, yielding

$$\begin{aligned} J(\mathbf{b}) &= \tilde{\mathbf{b}}^T\tilde{\mathbf{Q}}_b\tilde{\mathbf{b}} = \text{trace}\{\tilde{\mathbf{Q}}_b\tilde{\mathbf{B}}\}, \qquad (12) \\ \text{where} \quad & \tilde{\mathbf{b}} := [\mathbf{b}^T\ 1]^T \in \{-1,1\}^{(N+1)}; \\ & \tilde{\mathbf{B}} := \tilde{\mathbf{b}}\tilde{\mathbf{b}}^T, \quad \tilde{B}_{nn} = 1, \forall n. \end{aligned}$$

The SDR-based quasi-maximum-likelihood algorithm in [7] can be applied to solving (12). It involves two steps: first, relax the rank-1 constraint on $\tilde{\mathbf{B}}$ and solve (12) with respect to $\tilde{\mathbf{B}}$ using semi-definite programming; second, find an approximate Boolean solution to $\tilde{\mathbf{b}}$ (hence $\mathbf{s}$) via randomization.

The SDR algorithm is based on solving a convex optimization problem; hence, it does not suffer from local minima. Besides, it allows for $\epsilon = 0$ and naturally takes care of the rank-deficiency issue of $\mathbf{H}^T\mathbf{H}$.

5. EXEMPLARY APPLICATIONS

The problem of detecting sparse finite-alphabet signals appears in several signal processing applications. This section gives two examples.

5.1. Detection of Multiple Sources

Suppose that there are M signal sources scattered in a sensing field. Setting a grid of sampling points of adequate resolution, the locations of these sources can be represented by the grid. Indexing all the sampling points into a length-N sequence, we define an $N \times 1$ state vector $\mathbf{s} \in \{0,1\}^N$ to represent the locations of signal sources. The index of a nonzero element in $\mathbf{s}$ corresponds to a signal being 'on' at that sampling point; hence, the known location of that sampling point reveals the source location. Obviously, the number of 1's in $\mathbf{s}$ is M.

There are K sensors employed to collect linear measurements from these M sources, yielding a $K \times 1$ data vector $\mathbf{y}$ that obeys the linear model in (2). In many cases, the measurement matrix $\mathbf{H}$ is known based on the relative locations between the sampling points and the sensors. Here, M can be small, whereas N is typically set to a large value to attain high resolution for a large sensing field. To save sensing resources, the number of active sensors K can be smaller than N. As a result, a compressed sensing problem arises.

Such a binary detection problem can be found in several applications. For example: a) localization of multiple targets, where the sensing field is a geographical area; b) spectrum sensing of active transmitters emitting on multiple channels, where the sensing field is a wide spectrum band [8]; c) object detection in imaging, where the sensing field is the field of view of an optical device.

We test the proposed SD-CS algorithm for a toy problem, using parameters $N = 10$, $M = 3$, $K = 5$ and $L = 500$ simulation trials. In each trial, the measurement matrix $\mathbf{H}$ is randomly generated with *i.i.d.* Gaussian-distributed entries of zero mean and equal variance, where the variance is set according to the signal-to-noise-ratio (SNR) value. The noise $\mathbf{w}$ is Gaussian distributed with variance $\sigma_w^2 = 1$. In the SD-CS algorithm in (10), the coefficients λ and ϵ reflect the weights on sparsity-inducing l_1-norm and diagonal-loading l_2-norm. We test two selections: i) *SD-CS.i*: $\lambda = \epsilon = \sigma_w^2$, which uses diagonal loading only; ii) *SD-CS.ii*: $\epsilon = \sigma_w^2$ and $\lambda = \max\{\sqrt{2\log N}\sigma_w/||\mathbf{H}||_2, \sigma_w^2\}$, where λ is an empirical value suggested in [3] for the l_1-norm penalty term on real-valued $\mathbf{s}$ in (7). The MMSE, BP and BPDN are also tested for comparison. It is shown from Figure 1 that SD-CS.i yields better probability of detection (P_d) than SD-CS.ii, but the corresponding probability of false alarms (P_{fa}) is worse. When comparing the probability of correct estimation (P_c), both algorithms perform similarly, outperforming MMSE, BP and BPDN (BP is not shown; it performs silimarly to BPDN). The latter three algorithms ignore the finite-alphabet constraint, which explains the performance loss.

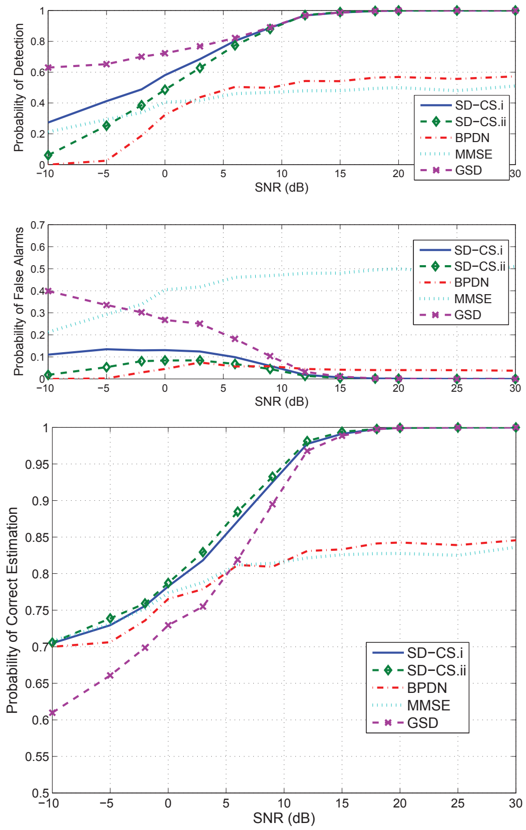

Fig. 1. Detection of multiple sources: P_d, P_{fa} and P_c.

The GSD algorithm in (5) is also applied to this detection problem after substituting the on-off state vector $\mathbf{s}$ by a non-sparse polar vector $\mathbf{b} = 2\mathbf{s} - 1 \in \{-1, 1\}^N$. Interestingly, both the GSD and SD-CS.i have a similar form of objective functions and perform similar lattice search steps, but the GSD has worse P_c performance in low SNR region. To explain this gap, we note that the l_2-norm term in the respective objective functions has different implications. For GSD, $\epsilon||\mathbf{b}||_2^2$ in (5) is a dummy term used to numerically regularize the rank-deficient channel covariance $\mathbf{H}^T\mathbf{H}$; indeed, $\mathbf{b} \in \{-1, 1\}^N$ herein and hence $\epsilon||\mathbf{b}||_2^2 = \epsilon N$ is constant. For SD-CS, $\epsilon||\mathbf{s}||_2^2$ in (10) plays a dual role: to regularize $\mathbf{H}^T\mathbf{H}$ and to induce sparsity on $\mathbf{s} \in \{0, 1\}^N$ since $\epsilon||\mathbf{s}||_2^2 = \epsilon||\mathbf{s}||_0$.

5.2. Digital Communication with Binary Modulation

Consider a multiuser communication system using binary polar modulation. The baseband input-output relationship can be described by a linear model $\mathbf{y} = \mathbf{H}\mathbf{b} + \mathbf{w}$, where $\mathbf{H}$ is a $K \times N$ channel response matrix and $\mathbf{b} \in \{-1, 1\}^N$ is the transmitted symbol vector. This model also describes a general filterbank transceiver [9].

When the system is overloaded, $\mathbf{H}$ is a fat matrix with $K < N$. The GSD algorithm in (5) can be used to solve for $\mathbf{b}$. However, the bit-error-rate (BER) suffers due to system overloading. Alternatively, we substitute $\mathbf{b}$ by $\mathbf{s} := \frac{\mathbf{b}+\mathbf{1}}{2} \in \{0, 1\}^N$, which results in an equivalent linear model $\mathbf{y}_1 = \mathbf{H}_1\mathbf{s} + \mathbf{w}$ where $\mathbf{y}_1 = \mathbf{y} + \mathbf{H}\mathbf{1}$ and $\mathbf{H}_1 = 2\mathbf{H}$. Probabilistically, half of the elements in $\mathbf{s}$ are zeros, which makes $\mathbf{s}$ much sparser than $\mathbf{b}$. Capitalizing on the prior knowledge of sparsity, we resort to the SD-SC algorithm in (10) to solve for $\mathbf{s}$.

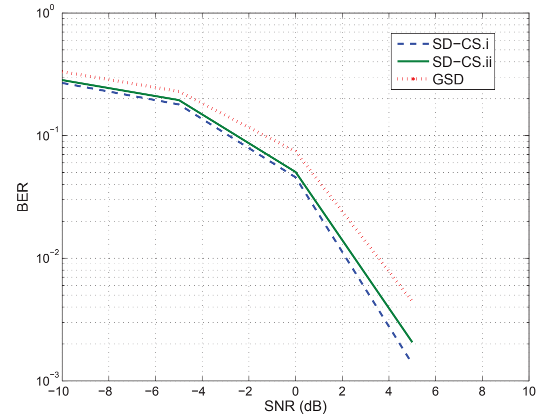

Fig. 2. Demodulation in binary communication: BER.

Figure 2 plots the BER of the GSD and SD-CS algorithms, for $N = 16$ and $K = 10$. The SD-CS performs slightly better, again due to the sparsity knowledge it exploits.

6. REFERENCES

[1] M. F. Duarte, M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, "Sparse Signal Detection from Incoherent Projections," *Proc. of ICASSP*, vol. 3, pp. III.305-308, May 2006.

[2] E. J. Candes, J. Romberg and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207 - 1223, May 2006.

[3] S.S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129 - 159, 2001.

[4] T. Cui and C. Tellambura, "An Efficient Generalized Sphere Decoder for Rank-Deficient MIMO Systems," *IEEE Communications Letters*, vol. 9, no. 5, pp. 423-425, May 2005.

[5] D. P. Wipf, and B. D. Rao, "Sparse Bayesian Learning for Basis Selection," *IEEE Trans. on Signal Processing*, vol. 50, no.4, pp. 912-922, April 2002.

[6] R. Tibshirani, "Regression shrinkage and selection via lasso," *J. Royal Statistics Soc., Ser. B*, vol. 58, pp. 267-288, 1996.

[7] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching, "Quasi-ML multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Trans. on Signal Proc.*, vol. 50, no.4, pp. 912-922, April 2002.

[8] J.-A. Bazerque and G. B. Giannakis, "Distributed Spectrum Sensing for Cognitive Radios by Exploiting Sparsity," *Proc. of Asilomar Conf.*, Pacific Grove, CA, Oct. 2008.

[9] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers: Part I" *IEEE Trans. on Signal Processing*, vol. 47, pp. 1988-2006, July 1999.