ADAPTIVE SAMPLING FOR FAST SPARSITY PATTERN RECOVERY

Francisco Ramirez-Javega[1], *David Matas*[1], *Meritxell Lamarca*[1,2]

Signal Theory and Communications Department, Universitat Politecnica de Catalunya (UPC), Barcelona (Spain)[1]
Electrical and Computer Engineering Department, University of Delaware[2]
Email: javega@tsc.upc.edu, (david.matas, meritxell.lamarca)@upc.edu

ABSTRACT

In this paper we propose a low complexity adaptive algorithm for lossless compressive sampling and reconstruction of sparse signals. Consider a sparse non-negative real signal $\mathbf{x}$ containing only $k << n$ non-zero values. The sampling process obtains m measurements by a linear projection $\mathbf{y} = \mathbf{A}\mathbf{x}$ and, in order to minimize the complexity, we quantize them to binary values. We also define the measurement matrix $\mathbf{A}$ to be binary and sparse, enabling the use of a simple message passing algorithm over a graph. We show how to adaptively construct this matrix in a multi-stage process that sequentially reduces the search space until the sparsity pattern is perfectly recovered. As verified by simulation results, the process requires $O(n)$ operations and $O(k\log(n/k))$ samples.

1. INTRODUCTION

In multiple signal processing applications, the selection of an appropriate basis leads to a sparse representation of the data. The Compressive Sensing (CS) framework [1, 2] established that this sparsity can be exploited to recover the signal from a number of samples significantly lower than the one required by the Shannon-Nyquist theorem.

This problem considers the estimation of an unknown sparse signal vector $\mathbf{x} \in \mathbb{R}^n$ with minimum 0-norm $||\mathbf{x}||_0$ from a vector of linear observations $\mathbf{y} = \mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is a fixed random $m \times n$ real valued matrix, known as measurement (or sampling) matrix. Only a small number k (the sparsity index) of elements of $\mathbf{x}$ are non-zero. The set containing the positions of these elements is known as the *support set* or *the sparsity pattern*, defined as $\mathcal{S} \triangleq \{i \in 1, ..., n : x_i \neq 0\}$, with $|\mathcal{S}| = k$.

The recovery of this sparsity pattern is the challenging part of this estimation problem and, in many applications, of fundamental interest. Several publications focus on the sparsity pattern recovery problem, obtaining theoretical limits of the number of samples needed to recover the sparsity pattern perfectly or to achieve a certain error performance for a given (k, n) pair [3]. In the noiseless case, methods and theoretical limits do not differ from CS [4], which is intrinsically also a sparsity support recovery problem.

While the minimization of the 0-norm is a NP complex problem, the authors in [1, 2] advocated that the original optimization of the CS problem can be relaxed by using the 1-norm. This allowed the application of linear programming, with tractable complexity. The key point to guarantee that the solution is the sparsest under this relaxation is the proper definition of the measurement matrix $\mathbf{A}$. Restricted Isometry property (RIP) was found to be a sufficient condition for its design [5, 6]. These contributions preceded the proposal of many practical algorithms and measurement matrix constructions: greedy ones trying to solve the 0-norm minimization, convex optimization ones for the 1-norm relaxation and graph-based approaches, offering diverse trade-offs between the required number of samples and complexity (see [7, 8] and references therein).

In this paper, we follow a different approach: instead of fixing a measurement matrix with certain properties, we propose a low complexity adaptive sampling algorithm which constructs a matrix adapted to the sparse signal. To the authors knowledge, this approach was first applied in the CS framework in [9] and the basic concepts were introduced by Indyk [10] within a group testing framework. Applications of adaptive sampling methods can be found in the literature, usually under the name of *active sensing*, like topography reconstruction, medical image classification or spectrum sensing in communications [11,12]. This approach is related with the adaptive source coding scheme presented in [13].

The aim of this work is to show how to construct an adaptive measurement matrix in a sequential multi-stage process. We propose an adaptive algorithm based on two steps that are repeated sequentially: first the construction of a random sparse measurement matrix and the sampling, and next a refinement process which recovers part of the signal, reducing the search space for next stage.

The measurement matrix is defined as binary and sparse, $\mathbf{A} \in \mathbb{F}_2^{m \times n}$, like in the application of expander graphs [14] and sudocodes [15] to CS. This enables the resolution of the problem with an affordable complexity by message passing propagation over a graph representation of this matrix. Simulation results show that this procedure ensures the recovery of $\mathcal{S}$ independently of the length of $\mathbf{x}$ with $m = O(k\log(n/k))$ samples in average while it requires $O(n)$ operations (binary comparisons).

In this paper, the proposed approach to sparsity pattern recovery is presented in section 2, we provide the insight that motivates the structure and design of the binary measurement matrix (section 3) and describe the proposed adaptive algorithm (section 4). Finally some performance results are shown and conclusions are drawn.

2. SPARSITY PATTERN RECOVERY PROBLEM

The objective is to reconstruct a signal, which is known to be sparse in a given basis, with as few number of measurements as possible compared to the dimension of the basis. These measurements are obtained by linear projections as introduced in the previous section. In CS, the measurement matrix is fixed, either randomly generated from a convenient ensemble [1] or with a deterministic construction [8] (e.g. chirps [16]) and then the reconstruction is approached as an optimization problem. Instead, we propose an adaptive sampling algorithm in which the measurement matrix is constructed in successive stages and its structure depends on the samples obtained previously.

As a result of this adaptability, the number of measure-

This work has been partially funded by the Catalan/Spanish Science and Technology Commissions and by FEDER funds from the European Commission: 2009SGR-1236, TEC2007-68094-C02-02, TEC2010-21245-C02-01, CSD2008-00010 and CEN-20101019

ments (i.e. rows in A) depends on the signal realization. For a perfect recovery of the sparsity pattern, samples are added until the pattern vector is completely recovered. The average number of samples ($\bar{m}$) is then the main metric of performance, but also the variance or the maximum number of samples required in the worst case (m^*). Alternatively, if a constraint on the maximum number of samples is imposed, then we are interested on the probability of perfect recovery of the support set $\mathcal{S}$.

We assume the signal to be real and non-negative, $\mathbf{x} \in \mathbb{R}^n_{\geq 0}$. We also require an *a priori* knowledge of the sparsity index k as an input parameter of the sampling process[1]. We will show that an approximated upper bound of k, $k \leq k^*$, suffices to ensure that the pattern vector is perfectly recovered (at the expense of an increased number of samples).

One of the main goals in order to make an adaptive sampling scheme attractive is to reduce the computational complexity as much as possible. With that in mind we design $\mathbf{A}$ to be sparse and binary and we also reduce the information provided by the samples to a binary basis. With that purpose, notice that the sparsity pattern recovery can be regarded as the reconstruction of a binary sequence $\mathbf{x}' \in \mathbb{F}^n$ of the same length as the sparse signal with ones at the non-zero positions. Let us then define the following nonlinear operator:

$$q(\alpha) = \begin{cases} 1 & \alpha > 0 \\ 0 & \alpha = 0 \end{cases} \tag{1}$$

which can be regarded as a 1-bit quantization or thresholding detection. Applying it to the measurements (element-wise), we observe that the signal before the linear projection is also reduced to a binary sequence, which corresponds to the sparsity pattern sequence $\mathbf{x}' = q(\mathbf{x})$:

$$\begin{gathered} \mathbf{y}' \triangleq q(\mathbf{A}\mathbf{x}) \\ \mathbf{y}' = q(\mathbf{A}q(\mathbf{x})) = q(\mathbf{A}\mathbf{x}') \end{gathered} \tag{2}$$

The measurement operation in this equation can be also expressed as the OR binary operation among sets of sparsity pattern bits $\mathbf{x}'$. This establishes a direct link between this problem and the non-linear binary source coding approach analyzed in [17][2]. From now on, we will work only with the 1-bit quantized version of $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}'$ and $\mathbf{y}'$.

For illustrative purposes, a simple example gives an idea of the reconstruction capability loss due to this quantization. Consider a measurement that is built adding two nonzero elements $y' = a+b$. In the linear case, if other measurements provide the knowledge of a then b can be easily obtained as $y'-a$ (this is how sudocodes or LDPC for CS work [15], [18]). On the contrary, in the nonlinear binary case, knowing $a = 1$ provides no information about b when $y' = 1$. In spite of this, simulation results show that the proposed multi-stage process approaches the theoretical limits, while working with a binary basis makes the algorithm complexity extremely low.

3. RANDOM SEQUENTIAL MEASUREMENT MATRIX CONSTRUCTION

There is a straightforward set of binary measurement matrices that recover $\mathbf{x}$ with the minimum possible number of samples $(k+1)$ [19]. These matrices are constructed as follows: if we look only at the columns indexed by the support set $\mathcal{S}$, k rows form a permutation of the identity matrix and the remaining row is zero in those positions and is one in all the rest, so it measures all the zero components of $\mathbf{x}$.

Unfortunately, this naive matrix cannot be directly constructed unless we already know $\mathcal{S}$. However, it tells us that obtaining measurements that are equal to zero is the key point in the CS approach for sparse signal recovery of non-negative signals, specially for graph-based methods [15, 18]. The idea is that if the measurement is zero then all the signal components that contributed to that measurement are also zero. Actually, at the expense of adding more samples than $k+1$, the recovery of all the zeros can be achieved with more than one row, each one measuring smaller subsets of zeros (not necessarily disjoint).

The objective of the proposed method is to build a matrix with that desired structure by means of an adaptive process which concatenates random matrices sequentially. Let us see first how a completely random binary matrix can approximate this perfect reconstruction matrix. For simplicity we assume a matrix $\mathbf{A} = [\phi_1 ... \phi_j ... \phi_m]^T$ with an uniform number of ones per row d_o. We choose $d_o << n$ so we force the matrix to be sparse. A given row ϕ_j^T is chosen uniformly from the set of sequences with d_o ones. Its non-null components are indexed by the set $\mathcal{I}_j \triangleq \{v \in 1, ..., n : \phi_{j,v} \neq 0\}$, $|\mathcal{I}_j| = d_o$.

Let us call p_0 the probability of a signal component being 0. Then, when all the ones of ϕ_j overlap with zero components of $\mathbf{x}'$, $|\mathcal{I}_j \cap \mathcal{S}| = 0$, the measurement $y'_j = \phi_j^T \mathbf{x}'$ is equal to zero. This happens with probability:

$$p(|\mathcal{I}_j \cap \mathcal{S}| = 0) = p_0^{d_0}. \tag{3}$$

and tells us the probability of recovering d_o zeros of the sparsity pattern from a single measurement.

Obtaining one of the k desired row vectors which overlap only with a one of $\mathbf{x}'$, i.e. $|\mathcal{I}_j \cap \mathcal{S}| = 1$, happens with probability

$$p(|\mathcal{I}_j \cap \mathcal{S}| = 1) = \binom{d_o}{1} p_0^{d_0 - 1}(1 - p_0). \tag{4}$$

However, we also generate ambiguous vectors that will not be useful to recover $\mathbf{x}'$, the ones such that $|\mathcal{I}_j \cap \mathcal{S}| > 1$, with a probability

$$p(|\mathcal{I}_j \cap \mathcal{S}| > 1) = 1 - p(|\mathcal{I}_j \cap \mathcal{S}| = 1) - p(|\mathcal{I}_j \cap \mathcal{S}| = 0). \tag{5}$$

As commented in section 2, this is the price to pay for working exclusively in $\mathbb{F}_2$.

The interesting point is that we can extract some information of the sparsity pattern after sampling with a sparse random matrix constructed as stated before. This brings the idea of sequentially repeating this process. With a given probability, every stage will reduce the search space of the next one when some components of $\mathbf{x}'$ are recovered.

4. SEQUENTIAL MULTI-STAGE SAMPLING

In this section we present the proposed sampling algorithm, which constructs the measurement matrix $\mathbf{A}$ as the sparsity pattern vector is recovered. It is based on a message passing process defined over a bipartite graph representation of the measurement matrix. The complexity of the algorithm is very low due to the sparsity of $\mathbf{A}$ and that the update rules of the nodes can be implemented only with comparisons. It is known that the complexity of these systems is $O(n)$ [20].

[1]It must be noted that k can be inferred from $\mathbf{y}$ with good reliability assuming that n and m are large enough. We can count the measurements equal to zero and apply the relationship in equation (3).

[2]The AND operator was applied to locate the few zero components in a sequence with a larger number of ones. It was shown that the performance of the system can be predicted with a set of recursive expressions.

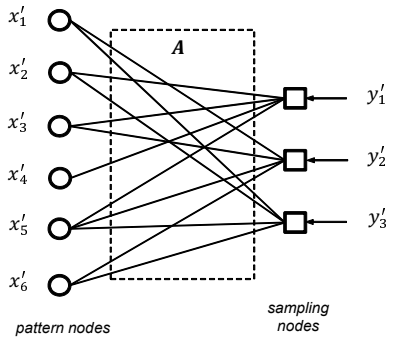

Figure 1: Graph Representation of the sampling

4.1 Graph Representation and Message Passing

The sampling process $\mathbf{y}' = q(\mathbf{A}\mathbf{x}')$ can be represented as a bipartite or Tanner graph, as the example one in figure 1. Left nodes represent the sparsity pattern components $\{x'_i\}$ while right nodes represent the quantized sampling operation, with the corresponding values $\{y'_j\}$. The connections are determined by the non-zero entries of the measurement matrix $\mathbf{A}$.

The value of the sparsity pattern components can be reconstructed over this graph employing a message passing algorithm like the ones proposed in [15]. We denote this process as refinement, and in our case reduces to a very simple algorithm (see "Algorithm 1"), as we work with binary quantized measurements. Notice that it does not require iterations nor arithmetic operations. In a general case in which we did not apply a binary quantization, another message passing algorithm could be employed as refinement process, while the proposed adaptive sampling and matrix generation procedure could still be applied.

As a result of the refinement algorithm, we obtain an estimation of $\mathbf{x}'$ with some perfectly recovered zeros and ones and a set of unknown components that cannot be verified (that we denote as '?'). The proposed multi-stage process is aimed at reducing this set to an empty one.

Considering the i-th stage, let us call $\mathbf{A}_i$ the corresponding measurement matrix and $\mathbf{y}'_i$ the obtained samples through it. Let us define the concatenation of all the matrices generated up to this stage as the i-th *aggregated matrix* $\mathbf{A}_i^a = [\mathbf{A}_1^T, ..., \mathbf{A}_i^T]^T$ and the sequence of samples as $\mathbf{y}_i^{'a} = [\mathbf{y}_1^{'T}, ..., \mathbf{y}_i^{'T}]^T$. These samples will be employed in a refinement process to obtain $\hat{\mathbf{x}}'_i$, an estimate of $\mathbf{x}'$ at the i-th stage, and the set of *unrecovered components* as $\mathcal{Q}_i \triangleq \{v \in 1, ..., n : \hat{x}'_{i,v} = ?\}$. This set defines the search space for the next stage, i.e. it determines the elements to be further processed, enabling the adaptive sampling procedure to concentrate only on the unknown part of the signal.

Now let us focus on the process to construct the matrix at the i-th stage, $\mathbf{A}_i$.

4.2 Design and Construction of the Matrix $\mathbf{A}_i$

The measurement matrix at the i-th stage is constructed randomly according to the observations made in section 3. Instead of operating over the whole $\mathbf{x}'$, it samples only the unrecovered pattern nodes indexed by $\mathcal{Q}_{i-1}$. Moreover, we force every connected pattern node to have degree one. Therefore, the columns of $\mathbf{A}_i$ will have only a one[3].

The probabilities (3) and (4), tied to the probability of recovering zeros and ones of $\mathbf{x}'$, depend critically on d_o and p_0. The former is the degree of the sampling nodes, while the latter is the probability of having a zero in the connected pattern nodes.

[3]This forces $\mathbf{A}_i$ to be full row rank.

Instead of fixing a value of d_o for all rows of $\mathbf{A}_i$, we choose to design a probability distribution $\rho_i(d_o)$ based on the pattern estimated by the previous stage, $\hat{\mathbf{x}}'_{i-1}$. We design the matrix to balance the number of ones and zeros of the sequence of measurements $\mathbf{y}'_i$, obtaining a binary entropy close to one. Therefore, the probability of obtaining a measurement equal to zero expressed in (3) has to be close 1/2. With that purpose, we find an average $\hat{d}_o$ from

$$p_0^{\hat{d}_0} = 1/2. \tag{6}$$

and then choose a distribution with the two closest integers around it and the appropriate weights such that $\hat{d}_o = \sum_{l=1,2} d_l \rho_i(d_l)$.

We use an estimation of p_0 given by the a priori sparsity index k^* of the signal and the number of components that are already recovered:

$$p_0 = 1 - \frac{k^* - k_{i-1}}{|\mathcal{Q}_{i-1}|} \tag{7}$$

where k_{i-1} is the number of ones in $\hat{\mathbf{x}}'_{i-1}$. At the first stage, $p_0 = 1 - k^*/n$.

Notice that with this construction, the number of samples generated at this stage is $m_i = |\mathcal{Q}_{i-1}| \sum_{\forall d_o} \frac{\rho_i(d_o)}{d_o}$.

Algorithm 2 describes the matrix construction process according to this design. After constructing $\mathbf{A}_i$, the corresponding new samples $\mathbf{y}'_i$ are obtained.

Algorithm 1 Refinement

Find zeros: If a measurement is zero, all connected pattern nodes are verified as zero.
Pruning: Remove all pattern nodes verified as zero.
Find ones: If the measurement is one and only one pattern node is connected, verify it as one.

Algorithm 2 Generation of $\mathbf{A}_i$

Initialize: $\mathcal{U} = \mathcal{Q}_{i-1}$
while $|\mathcal{U}| > 0$ **do**
  $d \leftarrow$ select according to ρ_i
  $\mathcal{V} \leftarrow$ choose randomly $min(d, |\mathcal{U}|)$ elements from $\mathcal{U}$
  add row to $\mathbf{A}_i$ with ones at the positions in $\mathcal{V}$
  $\mathcal{U} \leftarrow$ remove $\mathcal{V}$ from $\mathcal{U}$
end while

4.3 Sequential multi-stage process

In the two previous subsections we have explained how to construct the measurement matrix at every stage and how to estimate the sparsity pattern. To define the proposed adaptive sampling procedure, we just have to concatenate them as an stage and sequentially repeat it. By operating with this muti-stage procedure, the dimensionality of the problem is reduced at each stage and a more efficient use of

Algorithm 3 Sequential multi-stage sampling

Initialize: $i = 1$, $\mathcal{Q}_0 = \{1, ..., n\}$,
while $\mathcal{Q}_{i-1} \neq \emptyset$ or stop criterion **do**
  Construct $\mathbf{A}_i$
  $\mathbf{y}'_i = q(\mathbf{A}_i \mathbf{x}')$
  $(\mathcal{Q}_i, \hat{\mathbf{x}}'_i) \leftarrow$ Refinement with $\mathbf{A}_i^a$ and $\mathbf{y}_i^{'a}$.
  $i = i + 1$
end while

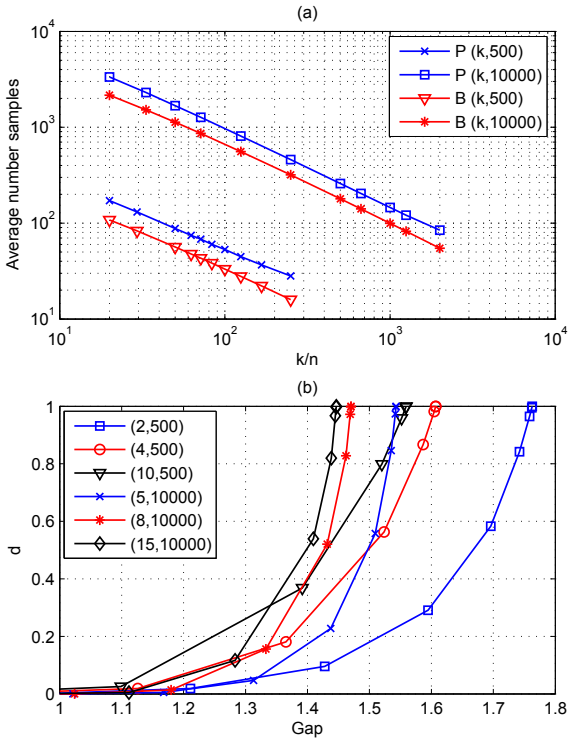

Figure 2: (a) Average number of samples for perfect recovery and a given sparsity ratio k/n. (b) Error and gap evolution curves through the stages of the adaptive sampling. In both plots $2 \cdot 10^5$ sparsity patterns were simulated.

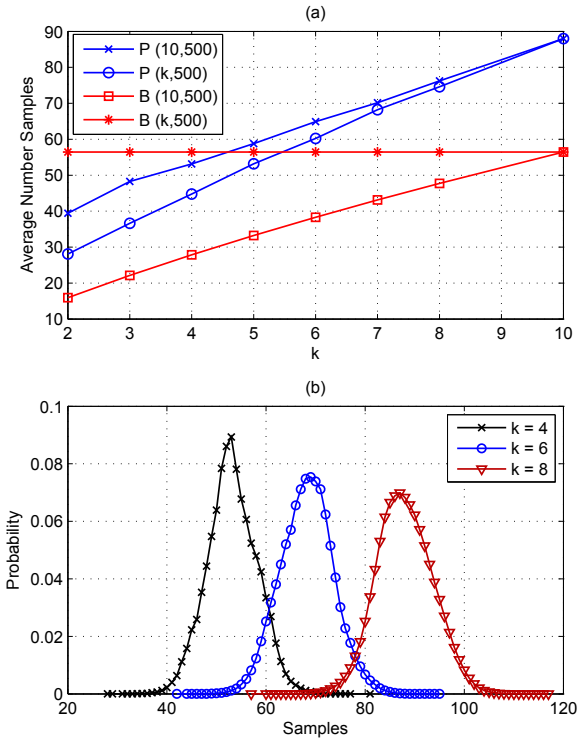

Figure 3: (a) Average number of samples for perfect recovery for $n = 500$ comparing results when k is known ($P(k, 500)$) with the ones when and upper value $k^* = 10$ is known ($P(10, 500)$). The bounds for nonadaptive sampling schemes are included for comparison. (b) Probability Mass function of the number of samples for $n = 500$ and several k.

the samples is made. This procedure can be continued until the sparsity pattern is perfectly recovered (as it was proved in [17]) or it can be stopped when the number of samples reaches a maximum value($m = m^*$) or the cardinality of $\hat{\mathcal{S}}$ reach a given value. Algorithm 3 describes this process.

The estimated support set of the sparse signal is the support set of $\hat{\mathbf{x}}'_i$ at the last stage, $\hat{\mathcal{S}} = \{i \in 1, ..., n : \hat{x}'_i \neq 0\}$.

5. RESULTS

In this section simulation results illustrate the performance of the proposed algorithm when samples are generated until perfect recovery is achieved and also when a maximum number of samples is imposed. In the second case, the performance is measured with the following recovery metric:

$$d = d(\mathcal{S}, \hat{\mathcal{S}}) = \frac{|\mathcal{S} \cap \hat{\mathcal{S}}|}{\max(|\mathcal{S}|, |\hat{\mathcal{S}}|)} \tag{8}$$

where $\hat{\mathcal{S}}$ and $\mathcal{S}$ are the estimated and the real support set, as defined before. Notice that $\hat{\mathcal{S}}$ includes recovered indexes from the support set ($\hat{x}'_i = 1$) and the unrecovered positions ($\hat{x}'_i =?$) and that $d = 1$ when the support set is perfectly recovered.

Figure (2.a) shows the average number of samples needed for perfect recovery of the proposed scheme (P) versus the sparsity rate (k/n) assuming that k is exactly known, for different values of the pair (k, n). It is compared with the bound $k \log(n/k)$ (B). We can observe that there exists an approximately constant multiplicative factor with respect to this theoretical bound $g = \dfrac{\bar{m}}{k \log(n/k)}$, which is larger than 1 ($g \approx 1.5$ for $n = 500$ and $g \approx 1.4$ for $n = 10000$). We will refer to this factor as gap hereafter.

Figure (2.b) shows the evolution of the recovery metric versus the gap through the stages of the adaptive sampling algorithm, i.e. for increasing number of measurements, for different values of the pair (k, n), and k also assumed exactly known. Each point of a given curve represents the resulting gap and recovery metric at the end of a stage. We have chosen to represent the gap instead of the number of samples because the normalization with the bound allows us to observe the dependence of the performance on (k, n): In a closer look, the gap decreases when the sparsity index k increases for a fixed length n, and also decreases with n.

The plotted curves helps us to explain the behavior of the sequential process: While the number of samples is below the bound (i.e. $g < 1$), only zeros of the sparsity pattern are found, so the recovery metric stays at 0. However, the knowledge of the zeros produces an avalanche effect when the number of samples exceeds a threshold. Next stages rapidly determine the ones of the sparsity pattern until convergence when recovery is perfect. As it can be seen, the gap is kept well below 2.

Figure (3.a) compares the performance of the proposed algorithm when k is exactly known with the case in which an upper bound is known instead ($k \leq k^*$). The average number of samples for perfect recovery versus k for $n = 500$ is shown and compared with the theoretical bound. As it can be seen, when assuming $k^* = 10$ the algorithm still works, at the expense of adding a limited number of additional samples

k	m_{OMP}	m^*	$\bar{m}$
4	56	50	39.3
8	96	76	64.0
12	136	98	85.2
16	184	117	104.4
20	228	135	121.4

Table 1: Comparison with OMP [21] in terms of required number of samples for $n = 256$ and several k.

(more the further is k from k^*).

Figure (3.b) shows the probability mass function (p.m.f.) of the number of samples for a given pair (k, n). Two millions of samples from each distribution were simulated for $k = \{4, 6, 8\}$ and $n = 500$. It can be seen that the variance increases with the sparsity index and also the p.m.f. becomes more Gaussian-shaped. Notice that the probability mass is concentrated around the mean; this means that the event of requiring an much larger number of measurements to recover $\mathcal{S}$ is very unlikely. In the simulations, none of the realizations required more than 81, 95 and 117 samples for $k = 4, 6, 8$, respectively.

Results have been compared with the ones obtained employing Orthogonal Maching Pursuit (OMP) [21], which employs a Gaussian measurement matrix. Table 1 sumarizes the maximum number of samples required for OMP (m_{OMP}) for $n = 256$ and several k values as found in [21] and compares it with the average and maximum number ($\bar{m}$ and m^*) of samples obtained with the proposed method. A significant improvement can be observed.

6. CONCLUSIONS

In this paper we introduced and analyzed the design of binary and sparse adaptive measurement matrices in a multistage process. We showed that the application of this method allows to perfectly recover, with high probability, the sparsity pattern of the signal with a limited number of measurements $O(k \log(n/k))$. We also showed that an upper bound of the sparsity index suffices to recover $\mathcal{S}$. At last we saw that the number of samples required to recover a given sequence is highly concentrated around the mean of the mass function, hence we can stop the adaptive process and ensure that the algorithm recovers the support set with a given (high) probability. It remains for future work to test the performance of the scheme considering real measurements. Then an enhanced refinement process (as the one presented in [15]) can be employed, possibly allowing to reduce the required number of samples.

REFERENCES

[1] E.J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[2] D.L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] G. Reeves and M. Gastpar, "Fundamental tradeoffs for sparsity pattern recovery," *CoRR*, abs/1006.3128, 2010.

[4] D. L. Donoho and J. Tanner, "Counting faces of randomly-projected polytopes when the projection radically lowers dimension," *J. of the American Mathematical Soc.*, pp.1–53, 2009.

[5] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, 28:253–263, 2008. 10.1007/s00365-007-9003-x.

[6] E.J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[7] T. Blumensath and M.E. Davies, "Stagewise weak gradient pursuits," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4333–4346, Nov. 2009.

[8] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *IEEE J. of Sel. Topics in Signal Process.*, vol. 4, no. 2, pp. 358–374, Apr. 2010.

[9] J. Haupt, R. Castro, and R. Nowak, "Adaptive discovery of sparse signals in noise," In *42nd Asilomar Conference on Signals, Systems and Computers*, Oct. 2008.

[10] P. Indyk, "Explicit constructions for compressed sensing of sparse signals," In *ACM-SIAM Symposium on Discrete Algorithms*, Jan. 2008.

[11] A. Tajer, R. Castro, and Xiaodong Wang, "Adaptive spectrum sensing for agile cognitive radios," In *IEEE Intl. Conf. on Acoustics Speech and Signal Processing (ICASSP), 2010* , March 2010.

[12] Wenyi Zhang, A.K. Sadek, Cong Shen, and S.J. Shellhammer, "Adaptive spectrum sensing," In *2010 Inf. Theory and Applications Workshop*, Feb. 2010.

[13] F. Ramirez Javega, M. Lamarca, and J. Garcia-Frias, "Progressive encoding with non-linear source codes for compression of low-entropy sources," In *6th Intl. Symposium on Turbo Codes and Iterative Information Processing, 2010*, Sept. 2010.

[14] S. Jafarpour, Weiyu Xu, B. Hassibi, and R. Calderbank, "Efficient and robust compressed sensing using optimized expander graphs," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4299–4308, Sep. 2009.

[15] S. Sarvotham, D. Baron, and R.G. Baraniuk, "Sudocodes - fast measurement and reconstruction of sparse signals," In *2006 IEEE Intl. Symposium on Information Theory*, Jul. 2006.

[16] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Applied and Computational Harmonic Analysis*, vol. 26, no. 2, pp. 283–290, 2009.

[17] F. Ramirez Javega and M Lamarca, "Variable-length source compression using successive refinement and nonlinear graph-based codes," In *Data Compression Conference, 2011*, Mar. 2011.

[18] Fan Zhang and H. D. Pfister, "On the iterative decoding of high-rate LDPC codes with applications in compressed sensing," *CoRR*, abs/0903.2232, 2009.

[19] Ping Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," In *IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, May 1996.

[20] A. Shokrollahi, "LDPC codes: An introduction," In *Coding, Cryptography and Combinatorics*, vol. 23 of *Progress in Computer Science and Applied Logic Ser.*, pp. 85–110. Springer-Verlag New York, LLC, Jun. 2004.

[21] J. A. Tropp, Anna, and C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, pp.4655–4666, 2007.