Kernel Reconstruction: an Exact Greedy Algorithm for Compressive Sensing

Belhassen Bayar and Nidhal Bouaynaya
Department of Electrical and Computer Engineering
Rowan University
Glassboro, NJ

Roman Shterenberg
Department of Mathematics
University of Alabama at Birmingham
Birmingham, AL

***Abstract*—Compressive sensing is the theory of sparse signal recovery from undersampled measurements or observations. Exact signal reconstruction is an NP hard problem. A convex approximation using the l_1-norm has received a great deal of theoretical attention. Exact recovery using the l_1 approximation is only possible under strict conditions on the measurement matrix, which are difficult to check. Many greedy algorithms have thus been proposed. However, none of them is guaranteed to lead to the optimal (sparsest) solution. In this paper, we present a new greedy algorithm that provides an exact sparse solution of the problem. Unlike other greedy approaches, which are only approximations of the exact sparse solution, the proposed greedy approach, called *Kernel Reconstruction*, leads to the exact optimal solution in less operations than the original combinatorial problem. An application to the recovery of sparse gene regulatory networks is presented.**

***Index Terms*—Compressive Sensing; Sparse Recovery; Greedy Algorithms; Gene Regulatory Networks.**

I. Introduction

We consider the problem of recovering a sparse vector $\boldsymbol{x} \in \mathbb{R}^m$ from under-determined measurements or observations $\boldsymbol{y} \in \mathbb{R}^{m \times n}$, $m < n$,

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

where $\boldsymbol{\Phi}$ is the $m \times n$ measurement matrix. We assume that $\boldsymbol{x}$ is k-sparse and the matrix $\boldsymbol{\Phi}$ is full rank. A vector $\boldsymbol{x}$ is called *k-sparse* if $\|\boldsymbol{x}\|_0 \leq k$, where the $\|\boldsymbol{x}\|_0 := |\text{supp}(\boldsymbol{x})| = \{j : x_j \neq 0\}$. It has become common to call $\|.\|$ the ℓ_0-norm. For $k \in \{1, 2, \cdots, n\}$,

$$\Sigma_k = \{\boldsymbol{x} \in \mathbb{C}^n : \|\boldsymbol{x}\|_0 \leq k\} \tag{2}$$

denotes the set of k-sparse vectors. Inferring the sparse vector $\boldsymbol{x}$ is equivalent to solving the ℓ_0-minimization problem [1]:

$$\min \|\boldsymbol{x}\|_0 \text{ subject to } \quad \boldsymbol{y} = \boldsymbol{\Phi x}, \tag{3}$$

where $\|\cdot\|_0$ is the number of nonzero elements in the vector $\boldsymbol{x}$. Compressive Sensing (CS) is the theory of reconstructing large dimensional signals from a small number of measurements by taking advantage of the signal sparsity. CS has been widely used and implemented in many applications including computized tomography [1], wireless communication [2], image processing [3] and camera design [4].

Unfortunately, the optimization problem in (3) is NP-hard. It has been shown, however, that if the measurement matrix $\boldsymbol{\Phi}$ satisfies the *null space property* (NSP), then sparse recovery is possible via ℓ_1 minimization [5]. The ℓ_1-minimization approach considers the solution of

$$\min \|\boldsymbol{z}\|_1 \quad \text{subject to } \boldsymbol{Az} = \boldsymbol{y}, \tag{4}$$

whcih is a convex optimization problem and can be seen as a convex relaxation of (3).

The NSP is difficult to show directly. The *restricted isometry property* (RIP) may be easier to handle. A matrix $\boldsymbol{A}$ is said to satisfy the RIP of order k if the smallest number δ_k such that

$$(1 - \delta_k)\|\|_2^2 \leq \|\boldsymbol{Az}\|_2^2 \leq (1 + \delta_k)\|\boldsymbol{z}\|_2^2, \tag{5}$$

for all $\boldsymbol{z} \in \Sigma_k$, satisfies $\delta_k \in (0, 1)$. The RIP implies the NSP as shown in [5]. If $\boldsymbol{\Phi}$ satisfies the RIP of order $3k$ with $\delta_{3k} < 1/3$, then recovery of all k-sparse vectors by ℓ_1-minimization is guaranteed. The RIP is also a difficult condition to check. It can be shown that δ_k can be equivalently defined as

$$\delta_k = \max_{T \subset \{1, \cdots, n\}, \#T \leq k} \|\boldsymbol{A}_T^* \boldsymbol{A}_T - Id\|_2, \tag{6}$$

where $\boldsymbol{A}_T$ denotes the column submatrix of $\boldsymbol{A}$ corresponding to the columns indexed by T.

Since both the NSP and RIP conditions are difficult to check, many greedy algorithms have been suggested to infer the sparse signal [6] [7] [8]. In particular, a family of thresholding algorithms has been suggested. The *Hard Thresholding* (HT) algorithm [6] makes an initial guess at the support of $\boldsymbol{x}$ and then projects the measurements $\boldsymbol{y}$ onto this support. An iterative version called Iterative Hard Thresholding (IHT) updates the residual and estimates a new $\boldsymbol{x}$ at every iteration until the stopping criterion is satisfied. Another version of lower complexity per iteration has been suggested called *Matching Pursuit* (MP). In particular, the *Orthogonal Matching Pursuit* (OMP) [9] is an iterative greedy algorithm that selects at each step the column which is most correlated with the current residuals. This algorithm estimates the nonzero entries in the vector $\boldsymbol{x}$ along with their indices. The combination of the orthogonalization of OMP [9] with the selection of a full candidate support set in HT leads to two-stage greedy algorithms, such as *Compressive Sampling Matching Pursuit* (CoSaMP) [8] and *Subspace Pursuit* (SP)[13]. These two-stage greedy algorithms incorporate a least squares problem in each iteration [10].

The greedy algorithms and the ℓ_p-minimization approaches, for $0 < p \leq 1$, have shown good performance in recovering sparse signals. However, none of them guarantees an exact recovery, and error bounds on their approximate solutions are unknown. Moreover, greedy algorithms require a priori knowledge of the exact number k of nonzero elements in the vector $\boldsymbol{x}$. In this paper, we present a new greedy algorithm that guarantees exact recovery of any sparse signal with unknown sparsity level k.

II. The Kernel Reconstruction Algorithm

We consider the linear operator $\boldsymbol{\Phi} : \mathbb{C}^n \longrightarrow \text{Range}(\boldsymbol{\Phi})$. We have $\mathbb{C}^n = \text{Range}(\boldsymbol{\Phi}^T) \oplus \text{Ker}(\boldsymbol{\Phi})$, where $\dim(\text{Ker}(\boldsymbol{\Phi})) = s$. Let $\boldsymbol{x}_0 \in \text{Range}(\boldsymbol{\Phi}^T)$ be a particular solution. We have

$$\begin{aligned} & \boldsymbol{y_0} = P_{\text{Range}(\Phi)} \boldsymbol{\Phi} P_{\text{Range}(\Phi^T)} \boldsymbol{x}_0 \\ \Longrightarrow \quad & \boldsymbol{x_0} = \left(P_{Range(\Phi)} \boldsymbol{\Phi} P_{Range(\Phi^T)}\right)^{-1} \boldsymbol{y_0}. \end{aligned} \tag{7}$$

Let $\boldsymbol{B} = \text{Null}(\boldsymbol{\Phi})$ be the $n \times s$ matrix whose columns are the s vectors that span the subspace $\text{Ker}(\Phi)$. A solution of the under-determined problem in (1) can be written as

$$\boldsymbol{x} = \boldsymbol{x}_0 + \sum_{j=1}^{s} a_j \boldsymbol{b}_j, \tag{8}$$

where $\boldsymbol{b}_j$'s are the kernel vectors or the columns of the matrix B, and the a_j's are the coefficients of the linear combination. It is easy to check from Eq. (8) that $\boldsymbol{\Phi x} = \boldsymbol{\Phi x}_0 = \boldsymbol{y}$. Writing Eq. (8) in matrix form, we obtain

$$\boldsymbol{x} = \boldsymbol{x}_0 + B\boldsymbol{a}, \tag{9}$$

where $\boldsymbol{a} = (a_1, \cdots, a_s)^T$. Thus, to find $\boldsymbol{x}$ we need to compute the vector $\boldsymbol{a}$. To do this, we assume that the sparsest solution $\boldsymbol{x}^*$ contains at least $s = \dim(\text{Ker}(\Phi))$ zeros, or equivalently that the vector $\boldsymbol{x} \in \Sigma_{n-s}$. Since $\text{rank}(\boldsymbol{B}) = s$, there exists s linearly independent rows of $\boldsymbol{B}$ that span a space that we call L. Let $\boldsymbol{P}_s$ be the projection matrix that projects $\boldsymbol{x}$ onto the space L. Thus, $\boldsymbol{P}_s \boldsymbol{x} = 0$, and $\boldsymbol{a}$ can be computed as follow

$$\begin{aligned} \boldsymbol{P}_s \boldsymbol{x} \quad &= \quad 0 \\ &\Longleftrightarrow \quad \boldsymbol{P}_s \boldsymbol{x}_0 + \boldsymbol{P}_s \boldsymbol{B} \boldsymbol{a} = 0 \\ &\Longrightarrow \quad \boldsymbol{a} = -\left(\boldsymbol{P}_s \boldsymbol{B}\right)^{-1} \boldsymbol{P}_s \boldsymbol{x}_0. \end{aligned} \tag{10}$$

Generically any s rows of $\boldsymbol{B}$ will be linearly independent. However, we do not know exactly which entries of $\boldsymbol{x}$ are equal to zero and thus we do not know which form of $\boldsymbol{P}_s$ to choose. Therefore, a combinatorial search should be performed in order to find all possible linearly independent rows of $\boldsymbol{B}$ and find out which basis leads to the sparsest solution.

The computational complexity of the proposed Kernel Reconstruction algorithm is $\mathcal{O}(n^s)$ since it requires $\binom{n}{s}$ operations in order to find the s linearly independent rows of B corresponding to the exact sparest solution. As a comparison, the computational complexity of the original ℓ_0-based problem is $\mathcal{O}(n^k)$ since it requires $\sum_{\ell=1}^{k} \binom{n}{\ell}$ operations to find the exact solution. Recall that in order to have an exact solution, we should have $m \geq k+1$. For a full rank matrix Φ, we have $s = n - m$. Therefore, it is easy to show that $s \leq n - k - 1$. Thus, for $k \geq \lceil \frac{n-1}{2} \rceil$ we have $\mathcal{O}(n^s) << \mathcal{O}(n^k)$.

KR algorithm: The Kernel Reconstruction algorithm is summarized below.

Input: The vector $\boldsymbol{y_0}$ and the full rank matrix $\Phi \in \mathbb{R}^{m \times n}$ $(m < n)$ satisfying the under-determined model $\boldsymbol{y_0} = \boldsymbol{\Phi x}$.

1 Compute the particular solution $\boldsymbol{x_0}$ in Eq. (7).
2 Compute $\boldsymbol{B} = \text{Null}(\Phi)$.
3 For each s linearly independent rows of $\boldsymbol{B}$, compute $\boldsymbol{a}$ in Eq. (10), and then the corresponding solution $\boldsymbol{x}$ in Eq. (8).
4 Choose the sparsest solution $\boldsymbol{x}$.

III. Application to Gene Regulatory Networks

A. Simulation results

We compare the proposed KR algorithm to the ℓ_1 and ℓ_2 approaches as well as the Compressive Sampling Matching Pursuit (CoSaMP) greedy algorithm [8]. To this aim, we generate a synthetic model that satisfies Eq. (1) where the entries of $\boldsymbol{x}$, $\boldsymbol{y_0}$ and the matrix Φ are random and normally distributed. In Fig. 1, we plot the percentage error $\frac{\|x - x^*\|_2}{\|x\|_2}$ versus the number of measurements m for $(n, k) = (20, 7)$. Observe from Fig. 1 that our approach can exactly recover the k-sparse signal for $m \geq 7$. This is there are at least $s = 13$ zeros in $\boldsymbol{x}$ for $m \geq 7$. For $m < 7$, we do not have exact reconstruction because we assume that we have at least $s = 14$ zeros in $\boldsymbol{x}$. One can notice that even for the approximate solution ,i.e., $m < 7$, the KR algorithm outperforms the other approaches.

B. Application to Gene Regulatory Networks

In this paper, we model the dynamics of genetic profiles as suggested in [11]. We consider the following linear model for the gene regulatory network,

$$\boldsymbol{Y} = \boldsymbol{AX}, \tag{11}$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m] \in \mathbb{R}^{n \times m}$, $\boldsymbol{Y} = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m] \in \mathbb{R}^{n \times m}$ and $\boldsymbol{A} = \{a_{ij}\}_{1 \leq i,j \leq n}$, n is the number of genes, m is the number of measurements or experiments and $m < n$. $\boldsymbol{x}_i \in \mathbb{R}^n$ is the vector of gene expression measurements (e.g., microarray data) for all n genes at experiment i. The entry a_{ij} of the connectivity matrix $\boldsymbol{A}$ represents the influence of gene j on gene i. The columns $\boldsymbol{y}_i$ represent the rate of change of the expression of genes and the external perturbations at time i. Our goal is to estimate the connectivity matrix $\boldsymbol{A}$. Genetic regulatory networks are known to be sparse, and thus the connectivity matrix $\boldsymbol{A}$ is assumed to be sparse.

We use the KR algorithm to infer GRNs. We generate synthetic genomic networks that satisfies the model in (11) with $n = 10$ genes and varying number of measurements m. We use a sparse model of the connectivity matrix, $\|\boldsymbol{A}\|_0 = 0.2n^2$. The entries of the matrix $\boldsymbol{A}$ are drawn from a standard normal distribution with zero-mean and unit variance, i.e.,

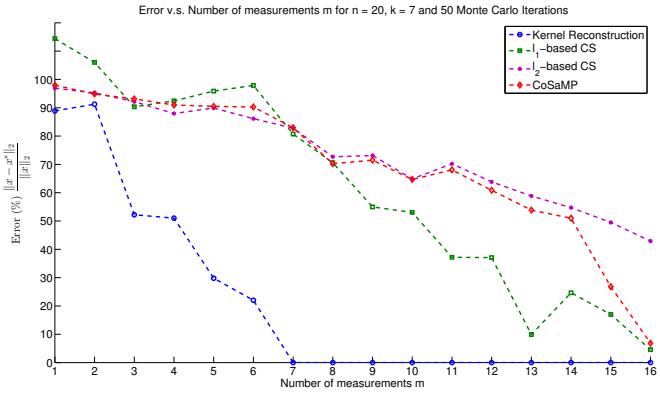

Fig. 1. Performance comparison of Kernel Reconstruction with ℓ_1,ℓ_2-based CS and CoSaMP for $n = 10$ and $k = 3$

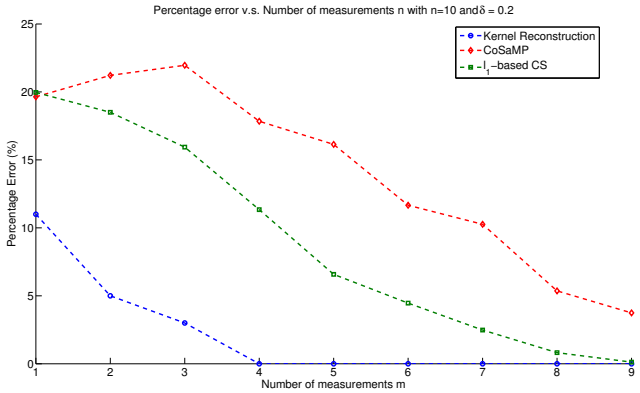

Fig. 2. Percentage error v.s. Number of measurements n with $n = 10$ and δ = 0.2

$a_{i,j} \in \mathcal{N}(0,1)$. We estimate the entries of the matrix $\boldsymbol{A}$ row by row,

$$\boldsymbol{Y}(i,:) = \boldsymbol{a}_i^T \boldsymbol{X}, \ i = 1, \cdots, n, \tag{12}$$

where $\boldsymbol{Y}(i,:)$ is the i^{th} row of $\boldsymbol{Y}$. The performance of the algorithm is assessed through the following measure suggested in [12]

$$E = \sum_{i=1}^{n}\sum_{i=j}^{n} e_{i,j} \quad \text{with} \quad e_{i,j} = \begin{cases} 1, & \text{if } |a_{ij} - \hat{a}_{ij}| > \delta \\ 0, & \text{otherwise,} \end{cases} \tag{13}$$

where a_{ij} and $\hat{a}_{ij}$ denote, respectively, the true and estimated connectivity entries. δ is a fixed threshold set to $\delta = \frac{1}{2}\min_{i,j}|a_{i,j}| \neq 0$. The percentage error is equal to E/n^2.

Figure 2 shows the percentage error versus the number of measurements m for $n = 10$-gene networks and $\|A\|_0 = 0.2n^2$. The proposed KR algorithm exactly recovers the network for $m > 3$. In particular, the KR algorithm outperforms the CoSamp greedy algorithm and the ℓ_1-minimization approach for all m.

In Fig. 3 we visualize the estimated directed networks using the different algorithms. Table I shows the detection of the known interactions in the original network. One can notice that our approach was able to detect all the 9 interactions in the original network, whereas the ℓ_1-minimization and the CoSaMP approaches have respectively detected 6 and 4 interactions.

IV. Conclusion

We presented a new algorithm that exactly recovers k-sparse signals from linear under-determined systems. We assume in our approach that the sparse solution has at least s zero entries, where s is the dimension of the kernel of the measurement matrix $\boldsymbol{\Phi}$. Although the greedy algorithm that we suggest requires $\binom{n}{s}$ combinatorial operations to find the optimal solution, it is still computationally less expensive than the ℓ_0-based approach. This combinatorial search is the price to pay for an exact solution (zero error) to be found. Finally, we compared our algorithm to the two ℓ_p approaches and the CoSaMP greedy algorithm [8]. In our future work, we will explore efficient implementations of the proposed kernel reconstruction algorithm, where we can find the optimal kernel basis without enumerating all $\binom{n}{s}$ basis.

Acknowledgment

This project is supported by Award Number R01GM096191 from the National Institute Of General Medical Sciences (NIH/NIGMS). The content is solely the responsibility of the

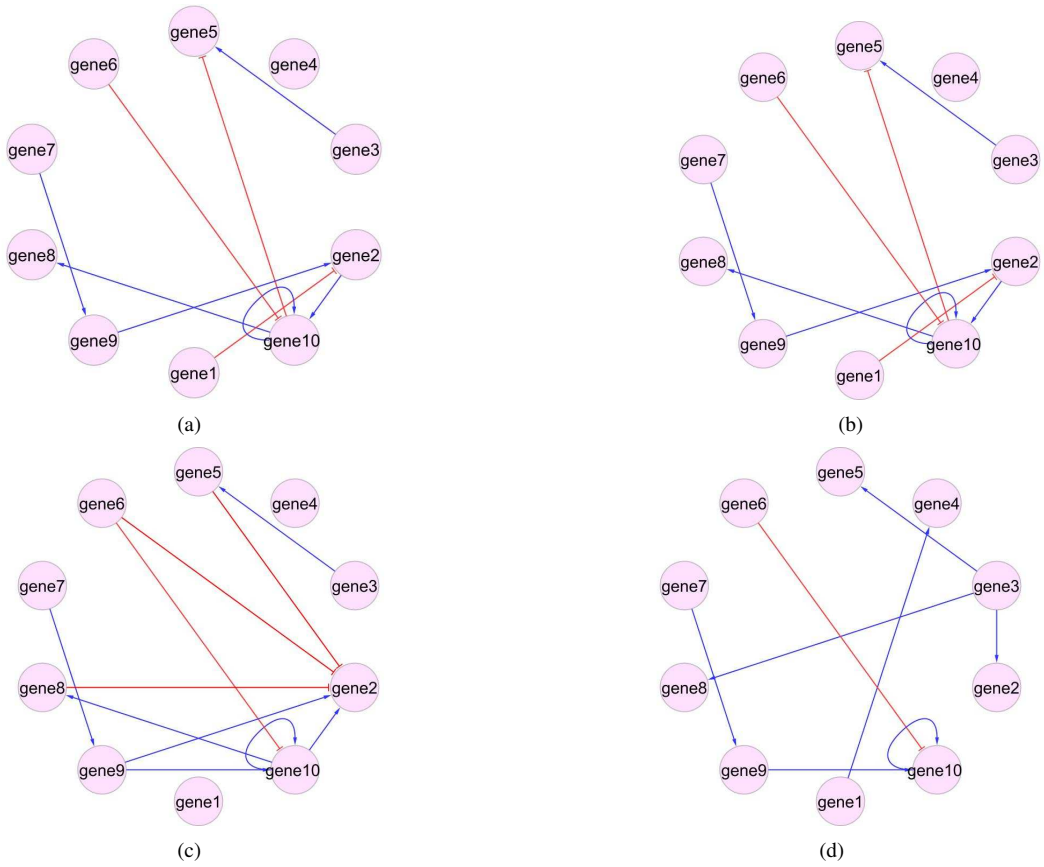

Fig. 3. Inference of a gene regulatory network with $n = 10$, $m = 5$ and $\|A\|_0 = 0.1n^2$. Blue and red edges denote, respectively, positive and negative interactions. (a) Ground truth interactions; (b) Kernel Reconstruction algorithm; (c) ℓ_1-minimization; (d) CoSaMP [8].

TABLE I
DETECTION PERFORMANCE

	(*g1,g2*)	(*g2,g10*)	(*g3,g5*)	(*g6,g10*)	(*g7,g9*)	(*g9,g2*)	(*g10,g5*)	(*g10,g8*)	(*g10,g10*)
Kernel Reconstruction	✓	✓	✓	✓	✓	✓	✓	✓	✓
ℓ_1-minimization	×	×	✓	✓	✓	✓	×	✓	✓
CoSaMP	×	×	✓	✓	✓	×	×	×	✓

authors and does not necessarily represent the official views of the National Institute Of General Medical Sciences or the National Institutes of Health.

REFERENCES

[1] 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, 2006.

[2] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 255–271, 2010.

[3] J. Bobin, J.-L. Starck, and R. Ottensamer, "Compressed sensing in astronomy," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 718–726, 2008.

[4] R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, 2008.

[5] M. Fornasier and H. Rauhut, "Compressive sensing," in *Handbook of Mathematical Methods in Imaging*. Springer, 2011, pp. 187–228.

[6] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[7] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

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

[9] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[10] J. D. Blanchard and J. Tanner, "Gpu accelerated greedy algorithms for compressed sensing," *Mathematical Programming Computation*, vol. 5, no. 3, pp. 267–304, 2013.

[11] M. Bansal, V. Belcastro, A. A. Impiombato, and D. di Bernardo, "How to infer gene networks from expression proles," *BioMed Central Systems Biology*, 2007.

[12] M. K. S. Yeung, J. Tegner, and J. J. Collins, "Reverse engineering gene networks using singular value decomposition and robust regression," *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 6163–6168, April 2002.