Gini Index based Approach for Sparse Signal Recovery

Shree Ranga Raju N M
Department of Systems Design Engineering,
University of Waterloo, Waterloo - N2L 3G1
Email : srrajuna@uwaterloo.ca

February 3, 2015

1 The gini index based sparse recovery algorithm

Algorithm 1 GI based sparse recovery algorithm (GIMP)

Inputs: $\mathbf{A}_{M\times N}$, $\mathbf{b}_{M\times 1}$, maxiter = Maximum number of iterations, and ε = Convergence threshold to stop the iterations.

Initialize: $\mathbf{x}^0 = 0$, $\mathbf{r}^0 = \mathbf{b}$,
$T = \emptyset$ (Final Set. Empty at the beginning)
$C^0 = \emptyset$ (Candidate Set)
$D^0 = \emptyset$ (Delete Set)
$n = 1$;
repeat
  C^n = find the indices which satisfies,
    $| < \mathbf{r}^{n-1}, \Phi_{C^n} > | \geq \mathrm{GI}(\mathbf{r}^{n-1})\ .\ \max_{j\in\Psi}| < \mathbf{r}^{n-1}, \Phi_j > |$,
    Ψ = [1,2,3,...,N].
  $\mathbf{x}^n_{T\cup C^n} = \mathbf{A}^{\dagger}_{T\cup C^n}\mathbf{b}$;
  D^n = choose indices such that it satisfies,
    $|\mathbf{x}^n_{T\cup C^n}| < \mathrm{GI}(\mathbf{x}^n_{T\cup C^n})\ .\ \max|\mathbf{x}^n_{C^n}|$
  $T = (T \cap C^n) \setminus D^n$;
  $\mathbf{x}^n_T = \mathbf{A}^{\dagger}_T\mathbf{b}$;
  $\mathbf{r}^n = \mathbf{b} - \mathbf{A}_T\mathbf{x}^n_T$;
until ($\|\mathbf{r}^n\|_2 < \varepsilon$) or (n == maxiter)
$n = n + 1$;

Output: $T, \mathbf{x}_T = \mathbf{A}^{\dagger}_T\mathbf{b}$, and $\mathbf{x}_{T^c} = 0$

This algorithm (GIMP) falls under the category of adaptive greedy pursuit algorithms. From Algorithm 1, T is the estimated support set until the current iteration, and matrix $\mathbf{A}_T$ consists of the columns of $\mathbf{A}$ with coordinates listed in T. Like most greedy pursuit algorithms, the first step

of the proposed method is to select the candidate set C^n whose correlation between the columns of $\mathbf{A}$ and the residual $\mathbf{r}^{n-1}$ is not less than GI($\mathbf{r}^{n-1}$) . $\max_{j\in\Psi}| < \mathbf{r}^{n-1}, \Phi_j >|$. The GI($\mathbf{r}^{n-1}$), i.e., Gini Index (GI) of ($\mathbf{r}^{n-1}$) is calculated using the equation (1). Here Ψ = [1,2,3,...,N] is the whole coordinate set. GI($\mathbf{r}^{n-1}$) is used to decide the number of atoms (columns of $\mathbf{A}$) to be chosen at each time of the iteration.

The output of GI is always between 0 and 1 for any given vector. GI is directly proportional to number of zeroes in a vector i.e., higher the value of GI (tending more towards 1), more number of zeroes in a given vector. In the above algorithm this property is used as an advantage for sparse signal recovery. If GI is exactly equal to 1, it just chooses the maximum correlation atom each time, similar to that of Orthogonal Matching Pursuit method [3]. In proposed algorithm GI of residual is used to form the search space, because residual decreases sequentially as the algorithm converges i.e., the number of zeroes in the residue vector increases, and hence it is more likely that the number of atoms chosen for the next iteration is less. Therefore, GI of residual will be larger (almost equal to 1) making number of atoms chosen for the next iteration less. This avoids choosing a specific number of atoms, unlike Subspace Pursuit [5] or Comoressive Sampling Matching Pursuit [6], where exactly K (number of non-zero or sparsity level) or $2K$ number of atoms in each iteration chosen. In the second step of the algorithm, we remove some of the atoms whose approximate coefficients $\mathbf{x}^n_{T\cup C^n}$ are smaller than GI($\mathbf{x}^n_{T\cup C^n}$) times the maximum amplitude of $\mathbf{x}^n_{C^n}$, where $\mathbf{x}^n_{C^n}$ are the corresponding entries of $\mathbf{x}^n_{T\cup C^n}$ in the current selected set C^n and GI($\mathbf{x}^n_{T\cup C^n}$) is calculated using equation (1). Here, $\mathbf{x}_T = \mathbf{A}^\dagger_T$ is the recovered sparse signal.

However, there is no mathematical proof for this experiment.

A brief overview of GI [1] is given as follows:

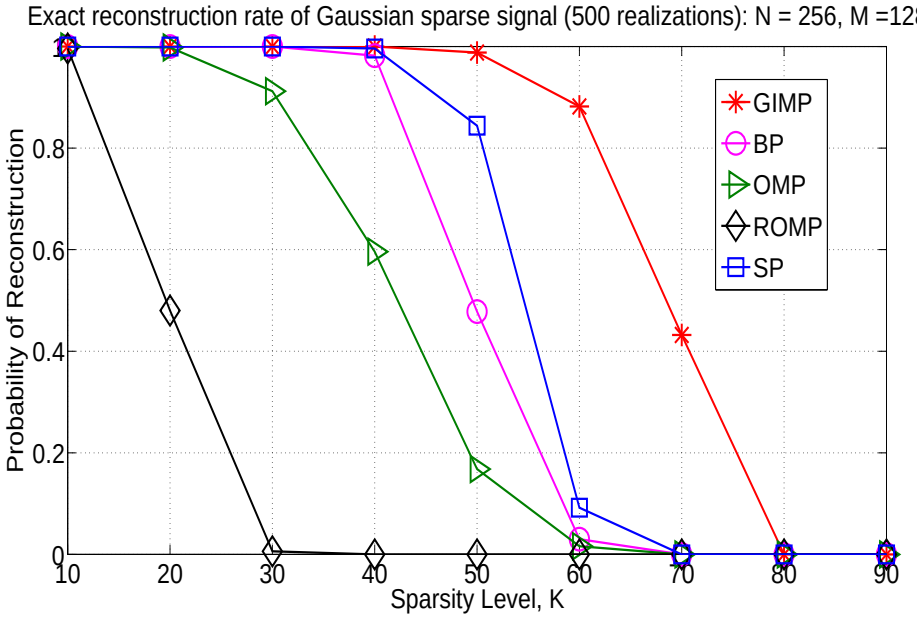
$$GI(\mathbf{g}) = 1 - 2 \sum_{k=1}^{N} \frac{\mathbf{g_{[k]}}}{\|\mathbf{g}\|_1} \Big(\frac{N-k+0.5}{N}\Big) \tag{1}$$

where $\|\mathbf{g}\|_1$ is the l_1 norm of vector $\mathbf{g}$.

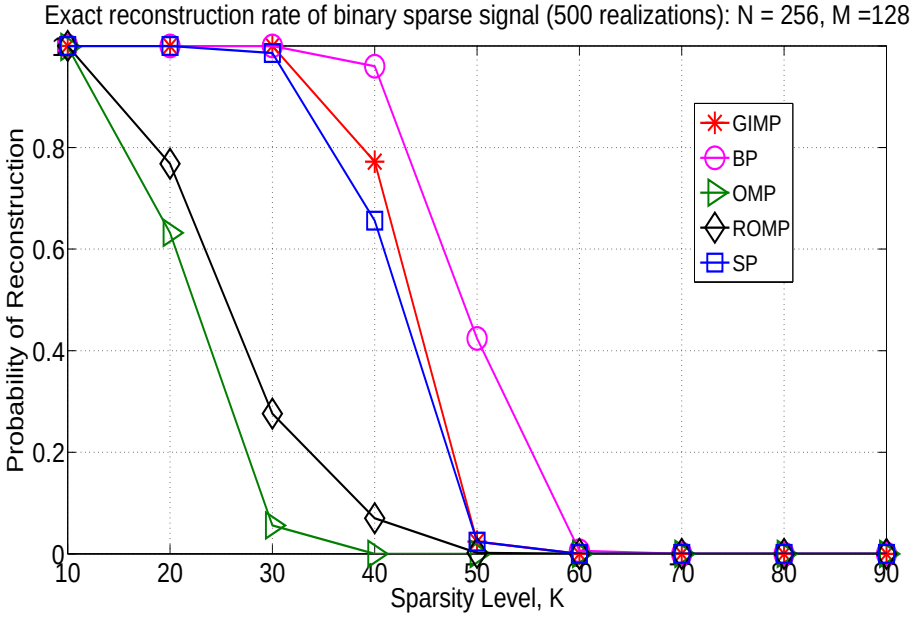
2 Experimental Results

In this section, we perform MATLAB simulations to show the benefits of the algorithm when compared with other greedy algorithms like OMP [3], ROMP [4], SP [5] methods and the l_1 norm minimization algorithm (BP). [2]

In our experiments we considered Gaussian sparse signals and binary sparse signals of length $N = 256$ to be the original signals. In each case, support was selected uniformly at random and K (sparsity level) number of non-zero enteries were generated from either Gaussian distribution or binary

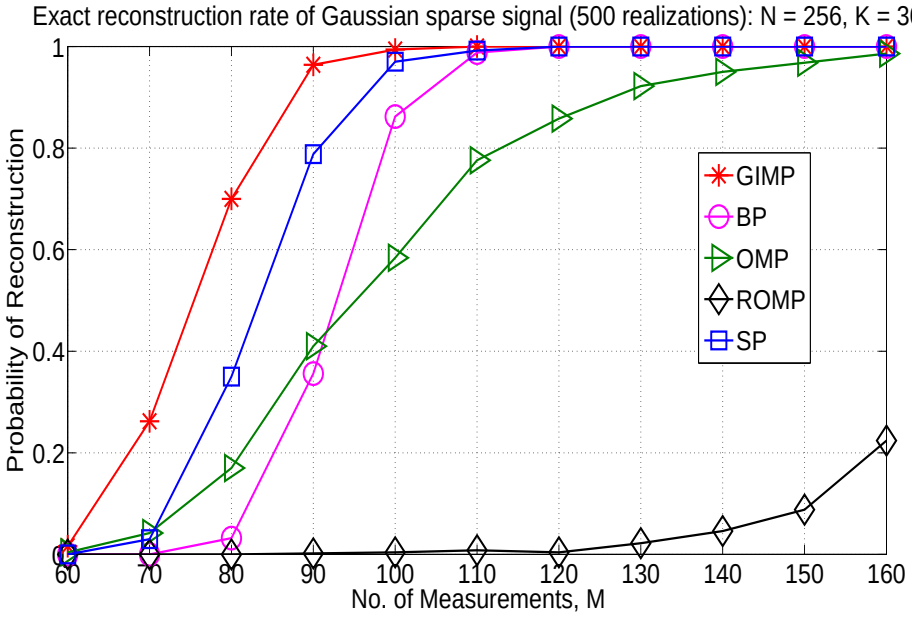


(a)

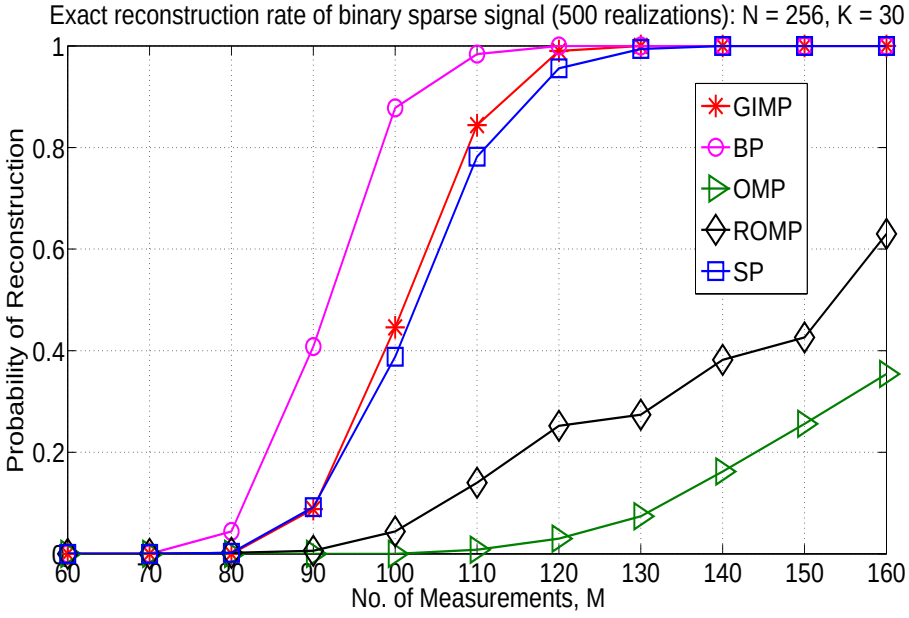


(b)

Figure 1: Performance in terms of sparsity level vs probability of exact reconstruction. (a) Gaussian Sparse Signals and (b) Binary Sparse Signals. ($N = 256$, $M = 128$)



(a)



(b)

Figure 2: Performance in terms of number of measurements vs probability of exact reconstruction. (a) Gaussian Sparse Signals and (b) Binary Sparse Signals. ($N = 256$, $K = 30$)

distribution. A Gaussian random matrix, $\mathbf{A}$ of size $M \times N$ was generated as the measurement matrix. Each experiment is repeated over 500 randomly generated samples of the signal. The reconstruction was considered to be exact when the maximum magnitude difference between the original signal $\mathbf{x}$ and the reconstucted signal $\hat{\mathbf{x}}$ is less than 10^{-3}, i.e., $|\mathbf{x} - \hat{\mathbf{x}}|_{max} < 10^{-3}$. In our experiments, OMP uses K iterations, SP and ROMP method uses parameters given in [5] and [4] respectivel, and BP code (MATLAB) is taken from [7]. The proposed method uses g which is Gini Index of the original signal, $\varepsilon = 10^{-6}$, maxiter = no. of measurements (M) as the input parameters.

References

[1] Corrado Gini, "Variability and Mutability", 1912.

[2] Scott S. chen and D.L. Donoho, Michael, and A. Saunders, "Atomic decomposition by basis pursuit", *SIAM journal on scientific computing*, vol.20, pp.33-61, 1998.

[3] J.A. tropp and A.C. Gilbert, "Signal Recovery from random measurements via orthogonal matching pursuit", *IEEE Trans. on Information Theory* vol. 53, no. 12, pp. 4655 - 4666, dec. 2007.

[4] D. Needell and R. Vershynin, "Uniform Uncertainity priniciple and signal recovery via regularized orthogonal matching pursuit", *Found. Computational Mathematics*, vol.9, pp. 317-334, 2009.

[5] Wei dai and O. Milenkovic, "Subspace Pursuit for compressive sensing signal reconstruction", *IEEE Trans. on Information theory* vol. 55, no.5, pp 2230 -2249, may 2009.

[6] 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.

[7] Justin Romberg, "l1-magic", http://users.ece.gatech.edu/ justin/l1magic.