*International Journal of Engineering & Technology*, 7 (4.36) (2018) 569-573

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

*Research paper*

# Algorithms on Sparse Representation

**D. Khalandar Basha[1*], T. Venkateswarlu[2]**

[1]*Department of ECE, Institute of Aeronautical Engineering, Hyderabad.*
*Research Scholar, Sri Venkateswara University, Tirupathi.*
[2]*Professor, Department of ECE, Sri Venkateswara University, Tirupathi.*

**Abstract**

Representation of signals and images in sparse become more interesting for various applications like restoration, compression and recognition. Many researches carried out in the era of sparse representation. Sparse represents signal or image as a few elements from the dictionary atoms. There are various algorithms proposed by researchers for learning dictionary. This paper discuss some of the terms related to sparse like regularization term, $l_0$minimization, $l_1$minimization, $l_2$ minimizationfollowed by the pursuit algorithms for solving $p_0$ problem, greedy algorithms and relaxation algorithms. This paper gives algorithmic approaches for the algorithms.

***Keywords:*** *Sparse representation, image restoration, regularization, pursuit algorithm, greedy algorithm, relaxation algorithm*

## 1. Introduction

Sparse representation [1] is adopted in face recognition, image denoising and image compression. Consider an un-determined linear system Ax = b, where A ϵ $R^{nxm}$ with n < m, with more unknowns than equations. The system has infinitely many solutions with assumption that A is full rank and $AA^H$ is invertible. The $l_2$ and $l_1$norms are defined by (1), (2) respectively.

$$||x||_2^2 = \sum_i |x_i|^2 \quad (1)$$

$$||x||_1^1 = \sum_i |x_i| \quad (2)$$

The $l_2$norm is sum of squares which gives the energy of the signal X.

### 1.1. Least squares

The solution for solving Ax = b by using $l_2$norm method. The problem is defined as (3).

$$\arg arg_x^{min} \; ||x||_2^2 \; such \; that \; Ax = b \quad (3)$$

The solution for the above problem is as specified in (4).

$$x = A^H(AA^H)^{-1}b \quad (4)$$

Where $A^H$ is the complex conjugate of *A*.
If *b* is noisy image or signal then the solution cannot found exactly because of unknown noise quantity. The solution can be found by approximation method by minimization method.

$$\arg arg_x^{min} \; ||b - Ax||_2^2 + \lambda||x||_2^2 \quad (5)$$

The solution for the above problem is given by

$$x = (A^H A + \lambda I)^{-1}A^H b \quad (6)$$

### 1.2.Sparse solution

The other method for solving $Ax = b$ is optimizing the problem which computes the minimizing the sum of absolute values in x.

$$\arg arg_x^{min} \; ||x||_1^1 \; such \; that \; Ax = b \quad (7)$$

The problem in (7) is called basic pursuit (BP) method which gives solution by iterative functional algorithm. If the approximation solution found by using (8) then the minimizing the function is called basic pursuit denoising (BPD) algorithm also called Least Absolute Shrinkage and Selection Operator (LASSO) problem.

$$arg_x^{min} \; ||b - Ax||_2^2 + \lambda||x||_1^1 \quad (8)$$

When the solution of $x$ is expected as sparse then using $l_1$ norm gives advantage over $l_2$ norm.
Chapter brief about regularization terminology.
Chapter 3 discuss Pursuit algorithms, Chapter 4 describes Greedy algorithms and relaxation algorithms discussed in chapter 6.

## 2. Fundamentals

In image processing the unknown image b may be blurred or low resolution image. The matrix A is non invertible linear degraded operator.
The goal is to reconstruct the original image x from given observed image b. it is a typical linear inverse problem which may have invite solution but unique solution is desired. To get a unique solution for above linear system some additional criteria are needed. By using regularization additional criteria J(x) can be added.
The optimized problem represented in (9).

$$(P_J): \quad {}^{min}_{x} J(x) s.t. \quad Ax = b \quad (9)$$

Where $J(x)$ is regularization term, causes penalty to each possible solution. The best regularized tern is squared Euclidean norm i.e. $l_2$- norm. This norm gives a unique elucidation ẋ, called minimum norm solution.
The Lagrangian using Lagrange multiplier shown in (10).

$$\mathcal{L}(x) = ||x||_2^2 + \lambda^T(Ax - b) \quad (10)$$

For image processing$l_2$regularization gives poor performance. For better performance convex function is used for$J(x)$.
If $J(x)$ is strictly convex then unique solution can be obtained.

### 2.1. Preferable regularization term

The convexity of the squared Euclidean norm is trivial. The other choice of $J(x)$ that are convex or strictly are $l_p$ – norms for $p \geq 1$ given by (11).

$$||x||_p^p = \sum_i |x_i|^p \quad (11)$$

The $l_\infty$ - norm, $l_1$ – norm are popular. The $l_\infty$ - norm gives maximal entry of the vector and $l_1$ – norm sums the absolute entries. The $l_1$provides sparse solution.

### 2. 1. 1 $l_1$ Minimization

The regularization term $J(x) = ||x|_1$ is convex but not strictly. The problem P1 is given by

$$(P_1): \quad \min_x ||x||_1 s.t. \; b = Ax \quad (12)$$

The problem in equ (12) have more than one solution. The optimal solution is one that have no more than n non-zeros. Consider the problem (13)

$$\min_x ||x||_1 s.t. \; ||x||_2 = 1 \quad (13)$$

Consider the vectors of x on the unit $l_2$ sphere and search for one vector which is shortest in$l_1$ in this set. Among all $l_2$ – normalized vectors, the shortest in $l_1$ are sparest possible. The algebraic solution is geometric view in Fig 1.

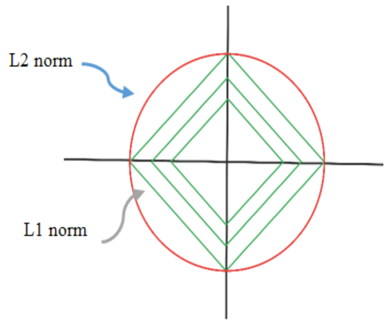



**Fig. 1:** Geometricview of$l_1$-norm and $l_2$-norm

The sparest solution is on the corner of the $l_1$sphere [2 – 5]. The shortest vectors in $l_1$ are extremely sparse having one non-zero entry.$l_1$-norm regularization function provide sparse solution. Fig 2 shows the optimized solutions of$l_1 - norm and l_2 - norm$. It is observed that $l_2 - norm$ provides dense solution [6 -8] where as $l_1 - norm$provides sparse solution.

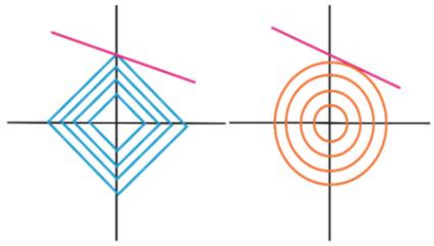


**Fig. 2:** Geometricsolution for$l_1$-norm and $l_2$-norm

In Fig 2 Pink colour line shows feasible solutions. Blue colour line indicates $l_1 - norm$ ball and orange colour portion indicates $l_2 - norm$ ball after blowing to obtain optimal solution.

### 2. 1. 2 $l_0$ Minimization

If the norms with norm value less than 1 will provide sparse solution but they are not convex. The behaviour of $|x|^p$ for various values of $p$ shown in Fig 3.

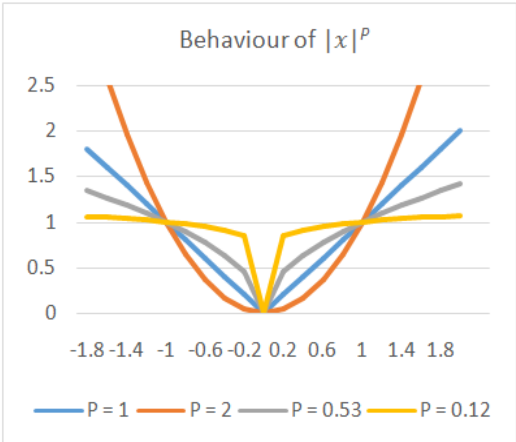


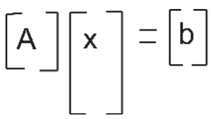

**Fig. 3:** Behavior of $|x|^p$

The $l_0 -$ norm simply counts the number of non – zeros in$x$.

## 3. Pursuit algorithms

The algorithms used for solving $p_0$ problem are called pursuit algorithms [9]. The objective of the pursuit algorithms is given in equation (14)

$$(p_0): \quad \min_x ||x||_0 s.t. \quad Ax = b \quad (14)$$




Where A ϵ Rnxm, x ϵ Rmx1 and b ϵ Rnx1.
The $(p_0)$ problem is discrete in nature. The unknown in $(p_0)$ is the support of the solution. This problem is highly non smooth because of$l_0 - norm$.

### 3.1. Exhaustive search algorithms

For solving $(p_0)$ problem search over all possible supports. Let k is the number of zeros in the solution. In this algorithm checks all possible nodes for the number of non-zeros in the optimal solution for k = 1, 2, 3, - - - till obtaining sparest solution.
The algorithm 1 for exhaustive search over all possible solution.

**Algorithm 1: Exhaustivesearchalgorithmforsparsesolution**

1. Setk = 1
2. Findallsupports $\{S_i\}$ ofcardinalityk
3. ForeachsupportSolvetheleastsquareproblem $\min_x ||x||_2^2$ s.t. $\sup(x) = S_i$
4. Evaluateerror
5. iferror $\leq e^2$ Stopelsek $\leftarrow k+1$ andgotostep 2

This algorithm provides accurate optimal solution for $P_0$ problem. This is combinatorial problem recognized to be NP-Hard [10, 11].

## 3.2. Approximation algorithms

The alternative for exhaust analysis is approximation algorithms. These algorithms sacrifices accuracy and not attain the strictly optimal solution. These algorithms are called Greedy Algorithms. Greedy algorithms search the tree of possibilities by removing many unlikely states, which leads various strategies to exist.

To approximate the solution of

$$(p_0) \quad \min_x ||x||_0 s.t. \qquad Ax = b,$$

Two strategies exists namely Greedy algorithms, relaxation algorithms.

### 3.2.1. Greedy algorithms

Greedy methods highlight the discrete nature of the problem and build the support

## 3.3. Relaxation algorithms

Relaxation methods smooth$p_0$ to become pleasant to handle continuous optimization task.

# 4. Greedy algorithms

Greedy algorithms searches set of support possibilities each node while trimming the unlikely states, which leads various strategies [12]. Because of this various strategies belongs to wide family of greedy methods. The various greedy algorithms are

- The Thresholding Algorithm
- Weak Matching Pursuit (WMP)
- Matching Pursuit (MP)
- Orthogonal Matching Pursuit (OMP)
- Least- Squares Orthogonal Matching Pursuit (LS-OMP)

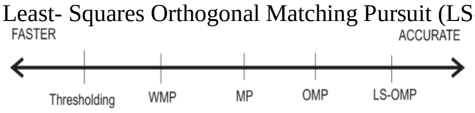



**Fig. 4:** Comparison of greedy algorithms

## 4.1. The thresholding algorithm

Thresholding algorithm is simplest and basic algorithm. It calculates$|A^T b|$and absolute value of this vector conveys association of all atoms w.r.t. b. The bigger the inner product it is marked as atom is in describing b.

**Algorithm 2: Weak Matching Pursuit (WMP) algorithm for sparse solution**

1. Compute $\beta = |A^T b|$ and sort this vector by absolute desending order
2. $|\beta_{i_1}| \geq |\beta_{i_2}| \geq |\beta_{i_3}| \geq \ldots\ldots \geq |\beta_{i_k}|$
3. Set $k = 0, x_0 = 0, S_0 = \{\}$
4. $k \leftarrow k+1$
5. Update $S_k = S_{k-1} \cup \{i_k\}$
6. Update $x_k$: $x_k = A_{S_k}^T b$
7. Update the residual : $r_k = b - Ax_k$
8. if $||r_k||_2 \leq e$ Stop else goto step 3

## 4.2. Weak Matching Pursuit (WMP)

In WMP $|a_i^T r_{k-1}|$ computed and stop the algorithm if its value is big enough i.e if its value is above t (<1) times the upper boundary. WMP if faster and it is less accurate.

**Algorithm 3: Weak Matching Pursuit (WMP) algorithm for sparse solution**

1. Set $k = 0, x_0 = 0, r_0 = b - Ax_0 = b, S_0 = \{\}$
2. $k \leftarrow k+1$
3. Compute $p(i) = |a_i^T r_{k-1}|$ for $1 \leq i \leq m$
4. Choose $i_0$, as soon as $p(i) \geq t * |r_{k-1}|_2$
5. Update $S_k = S_{k-1} \cup \{i_0\}$
6. Update $x_k := x_k = x_{k-1} \& x_k(i_0) = x_k(i_0) + a_{i_0}^T r_{k-1}$
7. Update the residual : $r_k = b - Ax_k$
8. if $||r_k||_2 \leq e$ Stop else goto step 2

## 4.3. Matching Pursuit (MP)

The term matching refers to the correlation between the residual and the atoms in A to find the next atom. MP retains $x_{k-1}$ and update it by adding new atom with its coefficient [13].MP may choose the same atom twice. MP is faster since it avoids LS computations.

**Algorithm 4: Matching Pursuit (MP) algorithm for sparse solution**

1. Set $k = 0, x_0 = 0, r_0 = b - Ax_0 = b, S_0 = \{\}$
2. $k \leftarrow k+1$
3. Compute $p(i) = |a_i^T r_{k-1}|$ for $1 \leq i \leq m$
4. Choose $i_0$, s.t. $\forall 1 \leq i \leq m$, $p(i_0) \geq p(i)$
5. Update $S_k = S_{k-1} \cup \{i_0\}$
6. Update $x_k := x_k = x_{k-1} \& x_k(i_0) = x_k(i_0) + a_i^T r_{k-1}$
7. Update the residual : $r_k = b - Ax_k$
8. if $||r_k||_2 \leq e$ Stop else goto step 2

The solution in each step is chosen such that

## 4.4. Orthogonal Matching Pursuit (OMP)

The OMP generates a series of solutions with gradually growing support by adding one non-zero at a time $x_0, x_1, x_2, \ldots\ldots, x_k, \ldots\ldots$ The proposed solutions deviates the equality $Ax = b$ and calculate the deviation error called residual vector at each step. The residual at $k^{th}$ step is given by

$$r_k = b - Ax_k \qquad (15)$$

OMP uses the residual in each step to choose the next atom [14]. With initial support$x_0 = 0$, so the residual is b. Next $x_0$is added with one non-zero element ($x_1$), then residual will become smaller, next add one more non zero then evaluate the residual [15]. Repeat the process till residual reaches closer to zero. The OMP algorithm is as follows

**Algorithm 5: Orthogonal Pursuit algorithm for sparse solution**

1. Set $k = 0, x_0 = 0, r_0 = b - Ax_0 = b, S_0 = \{\}$
2. $k \leftarrow k+1$
3. Compute energy $E(i) = \min_x ||z * a_i - r_{k-1}||_2^2$ for $1 \leq i \leq m$
4. Choose next atom $x_0$, s.t. $\forall 1 \leq i \leq m$, $E(i_0) \leq E(i)$
5. Update $S_k = S_{k-1} \cup \{i_0\}$
6. Caluate LS: $x_k = \min_x ||Ax - b||_2^2$ s.t. $Sup\{x\} = S_k$
7. Update the residual : $r_k = b - Ax_k$
8. if $r_k \leq e^2$ Stop else goto step 2

The solution in each step is chosen such that the new residual $r_k$ is orthogonal to all selected atoms in A. OMP will never choose the same atom again [16]. If any atom is selected second time its inner product with the same residual is zero. The main parts of the OMP algorithm are sweep stage and least square stage. In sweep stage the next atom is selected, which requires maximal value of $A^T r_{k-1}$ and updating of solution $x_k$will take place.

### 4.5. Least- Squares Orthogonal Matching Pursuit (LS-OMP)

LS-OMP computes the actual error directly. It performs LS multiple times to select the next atom whereas OMP relies on residual. LS-OMP is more efficient. LS-OMP complexity is more than OMP complexity.

***Algorithm 6: Least Squares Orthogonal Pursuit algorithm for sparse solution***

1. $Set\ k = 0, x_0 = 0, r_0 = b - Ax_0 = b,\ S_0 = \{\}$
2. $k \leftarrow k + 1$
3. $Compute\ energy\ E(i) = \min_x ||Ax - b||_2^2\ \ s.t. sup\{x\} = S_{k-1} \cup \{i\}\quad for\ 1 \leq i \leq m$
4. $Choose\ i_0\quad s.t.\quad \forall 1 \leq i \leq m,\ E(i_0) \leq E(i)$
5. $Update\ S_k = S_{k-1} \cup \{i_0\}$
6. $Caluate\ LS{:}\ x_k = \min_x ||Ax - b||_2^2\quad s.t. Sup\{x\} = S_k$
7. $Update\ the\ residual:\ r_k = b - Ax_k$
8. $if\ r_k \leq e^2\ \ Stop\ else\ \ goto\ step\ 2$

In LS-OMP the residual is calculated only for stopping criterion. In step6the calculation already done on step3. So step6 not needed to be again. The LS-OMP requires more time for computation to find the solution.

### 4.6. Designing greedy algorithms

For approximating the solution of $\min_x ||x||_0 s.t.\qquad Ax = b$

- Step1: Select A of size nxm then draw random sparse vector$x_0$.Where $||x||_0 \epsilon R^m$and$||x||_0 = S \ll n$.
- Step2: Create a vector by multiplying $Ax_0$
- Setp3: For the pair A and b, it may have S - sparse solutions.
- Step4: For this given pair feed it to tested algorithms and compare the result $\hat{x}_0$to the original$x_0$.

The stopping criteria is based on error. The error is calculated by equation (16)

$$e = \frac{||\hat{x} - x_0||_2^2}{||x_0||_2^2} \quad (16)$$

## 5. Relaxation algorithms

The problem $p_0$is highly non smooth due to $l_0$and the problem has many local minima points most of them are not solutions. To overcome these difficulties just smooth the $l_0$norm term by the methods called relaxation methods.
$l_0 - norm$is given by

$$||x||_0 = \sum_{k=1}^{m} p(x_k) \quad where\ \ p^*(x) = \begin{cases} 0 & x = 0 \\ 1 & x \neq 0 \end{cases} \quad (17)$$

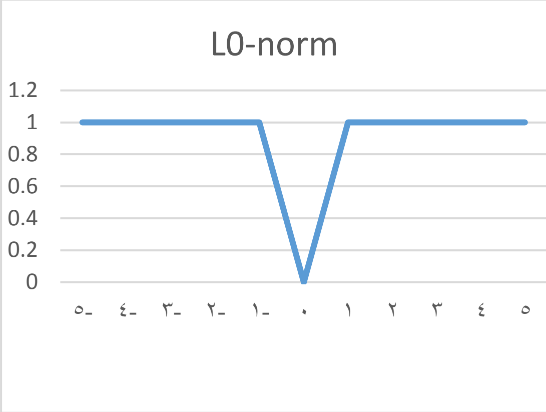



**Fig. 5:** $l_0 -$ norm

Relaxation methods will smooth this curve. There are many ways to smooth this. Commonly used for relaxing the $l_0$ expression are

$$p_\alpha(x) = 1 - e^{\left\{-\frac{x^2}{\alpha}\right\}} \quad (18)$$

$$p_\alpha(x) = \frac{x^2}{\alpha + x^2} \quad (19)$$

$$p_\alpha(x) = |x|^\alpha \quad (20)$$

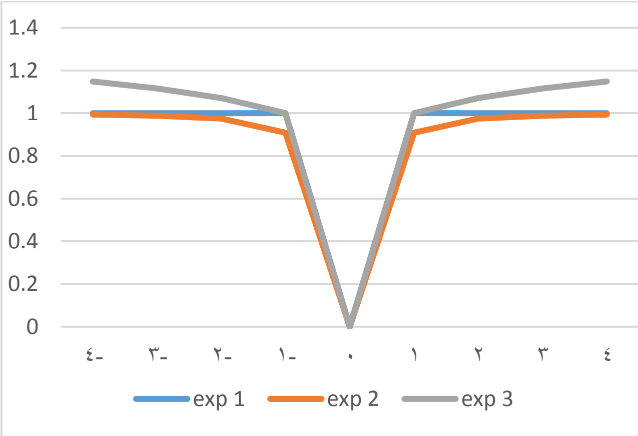



**Fig. 6:** Different smoothing functions of $l_0 -$ norm with = 0.1

As α decreases these functions closer to$l_0$.

### 5.1. Graduation optimization

This technique starts with a wide smoothing and solves the problem and narrow the option to recomputethe solution.

***Algorithm 7: Graduted Optimization***

1. $Set\ j = 1, x_0 = 0,\ \alpha\ with\ large\ value$
2. $For\ every\ x_{j-1}\qquad solve\ \alpha\ smoothed\ P_0\ \ problem\ \ ie\ P_0\{\alpha\}$
3. $j \leftarrow j + 1$
4. $goto\ step\ 2\ till\ the\ problem\ for\ narrow\ option$

### 5.2. Numerical solution of the relaxed $P_0$

There are various ways to solve the problem. One of the methods is Iterative Reweighted Least-Squares (IRLS). This algorithm is also known as Focal Underdetermined System Solver (FOCUSS). The smoothed $l_0$- norm expression evaluate weighted$l_2$. The weight is given by

$$w_k = \frac{P_\alpha(x_k)}{x_k^2} \quad (21)$$

IRLS iterates between a solutions of $l_2$ problem and update the weights.

## 6. Conclusion

Many researches working on sparse approximation for image restoration, image compression and pattern recognition. This paper discussed few elements of sparse representation, pursuit algorithms, greedy algorithms and relaxation algorithms. Pursuit algorithms solve $p_0$ problem. The problems are non-smooth and it's have more local minima. Relaxation algorithms smooth $l_0$ norm.

## References


[1] Elad M, *Sparse and redundant representations: from theory to applications in signal and image processing*, Springer Science & Business Media, (2010).

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

[3] Efron B, Hastie T, Johnstone I & Tibshirani R, "Least angle regression", *The Annals of statistics*, Vol.32, No.2, (2004), pp.407-499.

[4] Yang J & Zhang Y, "Alternating direction algorithms for \ell_1-problems in compressive sensing", *SIAM journal on scientific computing*, Vol.33, No.1, (2011), pp.250-278.

[5] Schmidt M, Fung G & Rosales R, "Fast optimization methods for l1 regularization: A comparative study and two new approaches", *European Conference on Machine Learning*, (2007), pp.286-297.

[6] Naseem I, Togneri R & Bennamoun M, "Linear regression for face recognition", *IEEE transactions on pattern analysis and machine intelligence*, Vol.32, No.11, (2010), pp.2106-2112.

[7] Zhang L, Yang M & Feng X, Sparse representation or collaborative representation: Which helps face recognition?. *IEEE international conference on Computer vision (ICCV)*, (2011), pp.471-478.

[8] Xu Y, Zhang D, Yang J & Yang JY, "A two-phase test sample sparse representation method for use with face recognition", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.21, No.9, (2011), pp.1255-1262.

[9] Chen SS, Donoho DL & Saunders MA, "Atomic decomposition by basis pursuit", *SIAM Journal on Scientific Computing*, Vol.20, No.1, (1998), pp.33-61.

[10] Wright J, Yang AY, Ganesh A, Sastry SS & Ma Y, "Robust face recognition via sparse representation", *IEEE transactions on pattern analysis and machine intelligence*, Vol.31, No.2, (2009), pp.210-227.

[11] Donoho DL, "For most large underdetermined systems of linear equations the minimal $\ell$1-norm solution is also the sparsest solution", *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, Vol.59, No.6, (2006), pp.797-829.

[12] Tropp JA, Gilbert AC & Strauss MJ, "Algorithms for simultaneous sparse approximation Part I: Greedy pursuit", *Signal Processing*, Vol.86, No.3, (2006), pp.572-588.

[13] Mallat SG & Zhang Z, "Matching pursuits with time-frequency dictionaries", *IEEE Transactions on Signal Processing*, Vol.41, (1993), pp.3397–3415.

[14] Pati YC, Rezaiifar R & Krishnaprasad PS, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition", *IEEE Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, (1993), pp. 40-44.

[15] Davis G, Mallat S & Avellaneda M, "Adaptive greedy approximations", *Constructive approximation*, Vol.13, No.1, (1997), pp.57-98.

[16] Tropp JA & Gilbert AC, "Signal recovery from random measurements via orthogonal matching pursuit", *IEEE Transactions on information theory*, Vol.53, No.12, (2007), pp.4655-4666.

[17] Needell D & Vershynin R, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit", *Foundations of computational mathematics*, Vol.9, No.3, (2009), pp.317-334.