# CS 754 - Advanced Image Processing
# Assignment 2 - Report

Shaan ul Haque - 180070053
Mantri Krishna Sri Ipsit - 180070032

February 18, 2021

## Question 1

### 1.1

If $\delta_{2s} = 1$, then for any $2s$ - sparse vector $h \in \mathbb{R}^n$, we have

$$0 \leq \|\Phi h\|_2^2 \leq 2 \|h\|_2^2 \tag{1}$$

If $2s$ columns of $\Phi$ are linearly dependent, then we have

$$\sum_i \alpha_i \phi_i = 0 \quad , \alpha_i \neq 0$$

Where $\{\phi_i\}$ are the $2s$ linearly dependent columns of $\Phi$ and $\{\alpha_i\}$ are some non-zero constants. This summation can be re-written as

$$\Phi v = 0, \quad \Phi = [\phi_1 | \ldots | \phi_{2s}], v^{\mathrm{T}} = [\alpha_1, \ldots, \alpha_{2s}]$$

Now, we can extend $v$ by adding rows of zeros so that the extended vector $v_1 \in \mathbb{R}^n$. Similarly, we can extend the columns of $\Phi$ by adding random columns (this works because we want the measurement matrix to be random in Compressed Sensing) so that $\Phi \in \mathbb{R}^{m \times n}$. Even after this extension, we still have

$$\Phi v_1 = 0$$

Clearly, we can see that $v_1$ is a $2s$ - sparse vector and also satisfies (1) trivially because $\|\Phi v_1\|_2 = 0$. Hence, the above justifies that when $\delta_{2s} = 1$, there exists a $2s$ - sparse vector $h$ such that $\Phi h = 0$ and $2s$ columns of $\Phi$ are linearly dependent.

### 1.2

We have $x^*$ and $x$ which satisfy the convex optimization problem

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_1 \quad \text{subject to } \|y - \Phi\tilde{x}\|_2 \leq \epsilon \tag{2}$$

Consider

$$\|\Phi(x^* - x)\|_2 = \|(\Phi x^* - y) + (y - \Phi x)\|_2$$

1

Then, using the triangle inequality, we have

$$\|\Phi(x^* - x)\|_2 = \|(\Phi x^* - y) + (y - \Phi x)\|_2 \leq \|\Phi x^* - y\|_2 + \|y - \Phi x\|_2$$

But using (2) we can say that

$$\|\Phi x^* - y\|_2 + \|y - \Phi x\|_2 \leq \epsilon + \epsilon = 2\epsilon$$

Hence, we have

$$\|\Phi(x^* - x)\|_2 \leq \|\Phi x^* - y\|_2 + \|y - \Phi x\|_2 \leq 2\epsilon$$

## 1.3

We know that each $h_{T_j}$ is $s$ sparse. We also know that the $\ell_\infty$ norm of a vector denotes the maximum element of that vector. As $h_{T_j}$ is $s$ sparse, we have, for each non-zero entry $\alpha_i$ of $h_{T_j}$

$$\alpha_i^2 \leq \left\|h_{T_j}\right\|_\infty^2 \quad i = 1, 2, 3, \ldots, s$$

Adding all the above $s$ inequalities, we have

$$\left\|h_{T_j}\right\|_2^2 \leq s \left\|h_{T_j}\right\|_\infty^2$$

$$\implies \left\|h_{T_j}\right\|_2 \leq s^{1/2} \left\|h_{T_j}\right\|_\infty \tag{3}$$

We know that, by the construction of the vectors $\{h_{T_j}\}$, the largest element of $h_{T_j}$ is smaller than the smallest element of $h_{T_{j-1}}$. In fact, the $\left\|h_{T_j}\right\|_\infty$ is smaller than every (non-zero) element of $h_{T_{j-1}}$. Re-writing this mathematically, we have

$$\left\|h_{T_j}\right\|_\infty \leq |\beta_i| \quad i = 1, 2, 3, \ldots, s$$

Adding all the above $s$ inequalities, we get

$$s \left\|h_{T_j}\right\|_\infty \leq \left\|h_{T_{j-1}}\right\|_1$$

$$\implies s^{1/2} \left\|h_{T_j}\right\|_\infty \leq s^{-1/2} \left\|h_{T_{j-1}}\right\|_1 \tag{4}$$

Combining the inequalities (3) and (4), we have

$$\left\|h_{T_j}\right\|_2 \leq s^{1/2} \left\|h_{T_j}\right\|_\infty \leq s^{-1/2} \left\|h_{T_{j-1}}\right\|_1$$

## 1.4

Using the relation between the 1st and the 3rd expression in the above inequality, we have

$$\left\|h_{T_j}\right\|_2 \leq s^{-1/2} \left\|h_{T_{j-1}}\right\|_1$$

Summing up all such above inequalities for $j \geq 2$, we have

$$\sum_{j \geq 2} \left\|h_{T_j}\right\|_2 \leq s^{-1/2} \left( \left\|h_{T_1}\right\|_1 + \left\|h_{T_2}\right\|_1 + \ldots \right) \tag{5}$$

By construction of $\{h_{T_j}\}$, we can say that

$$\left\|h_{T_1}\right\|_1 + \left\|h_{T_2}\right\|_1 + \ldots = \left\|h_{T_0^c}\right\|_1$$

Hence, (5) becomes

$$\sum_{j \geq 2} \left\|h_{T_j}\right\|_2 \leq s^{-1/2} \left( \left\|h_{T_1}\right\|_1 + \left\|h_{T_2}\right\|_1 + \ldots \right) \leq s^{-1/2} \left\|h_{T_0^c}\right\|_1 \tag{6}$$

## 1.5

We know that

$$h_{(T_0 \cup T_1)^c} = \sum_{j \geq 2} h_{T_j}$$

Hence, we have

$$\left\| h_{(T_0 \cup T_1)^c} \right\|_2 = \left\| \sum_{j \geq 2} h_{T_j} \right\|_2$$

Using the extended version of triangle inequality for the above equation and using the inequality (6) above, we get

$$\left\| h_{(T_0 \cup T_1)^c} \right\|_2 = \left\| \sum_{j \geq 2} h_{T_j} \right\|_2 \leq \sum_{j \geq 2} \left\| h_{T_j} \right\|_2 \leq s^{-1/2} \left\| h_{T_0^c} \right\|_1 \tag{7}$$

## 1.6

We have the *reverse triangle inequality* for any two real numbers as follows:

$$|x - y| \geq \left| |x| - |y| \right|$$

Using this inequality for $|x_i + h_i|, i \in T_0$, we have

$$|x_i + h_i| \geq \left| |x_i| - |h_i| \right| \geq |x_i| - |h_i|$$

Summing over all $i \in T_0$, we have

$$\sum_{i \in T_0} |x_i + h_i| \geq \left\| x_{T_0} \right\|_1 - \left\| h_{T_0} \right\|_1 \tag{8}$$

Again using the reverse triangle inequality for $|h_i + x_i|, i \in T_0^c$, we have

$$|h_i + x_i| \geq \left| |h_i| - |x_i| \right| \geq |h_i| - |x_i|$$

Summing over all $i \in T_1$, we have

$$\sum_{i \in T_0^c} |x_i + h_i| \geq \left\| h_{T_0^c} \right\|_1 - \left\| x_{T_0^c} \right\|_1 \tag{9}$$

By adding the inequalities (8) and (9) we get

$$\sum_{i \in T_0} |x_i + h_i| + \sum_{i \in T_0^c} |x_i + h_i| \geq \left\| x_{T_0} \right\|_1 - \left\| h_{T_0} \right\|_1 + \left\| h_{T_0^c} \right\|_1 - \left\| x_{T_0^c} \right\|_1 \tag{10}$$

## 1.7

As per the convex optimization problem that we are solving, we have

$$\|x\|_1 \geq \|x + h\|_1$$

Using this in the inequality (10) we have

$$\|x\|_1 \geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1 \tag{11}$$

By definition we have

$$\|x_{T_0^c}\|_1 = \|x - x_s\|_1$$

By using the reverse triangle inequality, we have

$$\|x_{T_0^c}\|_1 = \|x - x_s\|_1 \geq \|x\|_1 - \|x_s\|_1$$

Noting the fact that $\|x_s\| = \|x_{T_0}\|$ and using inequality (11) we have

$$\|x_{T_0^c}\|_1 \geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1 - \|x_{T_0}\|_1$$

which is same as

$$\|h_{T_0^c}\|_1 \leq \|h_{T_0}\|_1 + 2\|x_{T_0^c}\|_1 \tag{12}$$

## 1.8

Using inequalities (7) and (12) we have

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq s^{-1/2}\left(\|h_{T_0}\|_1 + 2\|x_{T_0^c}\|_1\right)$$

Using the known inequality $\|v\|_1 \leq \sqrt{k}\|v\|_2$ for a $k$ sparse vector v, we have

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq s^{-1/2}\left(\|h_{T_0}\|_1 + 2\|x_{T_0^c}\|_1\right) \leq s^{-1/2}\left(\sqrt{s}\|h_{T_0}\|_2 + 2\|x_{T_0^c}\|_1\right)$$

Hence we have

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq \|h_{T_0}\|_2 + 2e_0, \quad e_0 \equiv s^{-1/2}\|x - x_s\|_1 \tag{13}$$

## 1.9

We know that by Cauchy-Schwartz inequality, we have

$$|\langle \Phi h_{T_0 \cup T_1}, \Phi h\rangle| \leq \|\Phi h_{T_0 \cup T_1}\|_2 \|\Phi h\|_2 \tag{14}$$

As $h = x^* - x$, we have by inequality justified in section 1.3,

$$\|\Phi h\|_2 = \|\Phi(x^* - x)\|_2 \leq 2\epsilon$$

also, as $h_{T_0 \cup T_1}$ is a $2s$ sparse vector and $\Phi$ follows RIP of order $2s$, we have

$$\|\Phi h_{T_0 \cup T_1}\|_2 \leq \sqrt{1 + \delta_{2s}}\|h_{T_0 \cup T_1}\|_2$$

Using the above two inequalities in (14), we have

$$|\langle \Phi h_{T_0 \cup T_1}, \Phi h\rangle| \leq \|\Phi h_{T_0 \cup T_1}\|_2 \|\Phi h\|_2 \leq 2\epsilon\sqrt{1 + \delta_{2s}}\|h_{T_0 \cup T_1}\|_2 \tag{15}$$

4

## 1.10

Lemma 2.1 from the paper states

$$|\langle \Phi x, \Phi x' \rangle| \leq \delta_{s+s'} \, \|x\|_2 \, \|x'\|_2$$

for all $x, x'$ supported on disjoint subsets $T, T' \subset \{1, 2, \ldots, n\}$ with $|T| \leq s, |T'| \leq s'$ Using this lemma for $x = h_{T_0}$ and $x' = h_{T_j}$, we have

$$|\langle \Phi h_{T_0}, \Phi h_{T_j} \rangle| \leq \delta_{2s} \, \|h_{T_0}\|_2 \, \|h_{T_j}\|_2 \tag{16}$$

## 1.11

We know that for two positive real numbers $a$ and $b$, the QM-AM inequality is as follows:

$$\frac{a+b}{2} \leq \sqrt{\frac{a^2+b^2}{2}}$$

Taking $a = \|h_{T_0}\|_2$ and $b = \|h_{T_1}\|_2$ and substituting in the above inequality, we get

$$\|h_{T_0}\|_2 + \|h_{T_1}\|_2 \leq \sqrt{2} \, \sqrt{\|h_{T_0}\|_2^2 + \|h_{T_1}\|_2^2} = \sqrt{2} \, \|h_{T_0 \cup T_1}\|_2 \tag{17}$$

as $T_0$ and $T_1$ are disjoint.

## 1.12

By replacing $T_0$ with $T_1$ in inequality (16) and adding both, we get

$$|\langle \Phi h_{T_0}, \Phi h_{T_j} \rangle| + |\langle \Phi h_{T_1}, \Phi h_{T_j} \rangle| \leq \delta_{2s}(\|h_{T_0}\|_2 + \|h_{T_1}\|_2) \, \|h_{T_j}\|$$

Making use of triangle inequality for the LHS of above inequality, we get

$$|\langle \Phi h_{T_0} + \Phi h_{T_1}, \Phi h_{T_j} \rangle| \leq \delta_{2s}(\|h_{T_0}\|_2 + \|h_{T_1}\|_2) \, \|h_{T_j}\|$$

Replacing $\Phi h_{T_0} + \Phi h_{T_1} = \Phi h_{T_0 \cup T_1}$ and above inequalities for $j \geq 2$, we get

$$\sum_{j \geq 2} |\langle \Phi h_{T_0 \cup T_1}, \Phi h_{T_j} \rangle| \leq \delta_{2s}(\|h_{T_0}\|_2 + \|h_{T_1}\|_2) \left( \sum_{j \geq 2} \|h_{T_j}\|_2 \right)$$

Again making use of triangle inequality for the LHS of the above inequality, we get

$$\left| \left\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right| \leq \delta_{2s}(\|h_{T_0}\|_2 + \|h_{T_1}\|_2) \left( \sum_{j \geq 2} \|h_{T_j}\|_2 \right)$$

Finally, using (17), we finally have

$$\left| \left\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right| \leq \delta_{2s} \sqrt{2} \, \|h_{T_0 \cup T_1}\| \left( \sum_{j \geq 2} \|h_{T_j}\|_2 \right) \tag{18}$$

Now, using the fact that $\Phi h = \Phi h_{T_0 \cup T_1} + \sum_{j \geq 2} \Phi h_{T_j}$ in the inequality (15) above, we get

$$\left| \left\langle \Phi h_{T_0 \cup T_1}, \Phi h_{T_0 \cup T_1} + \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right| \leq 2\epsilon\sqrt{1 + \delta_{2s}} \, \|h_{T_0 \cup T_1}\|_2$$

Using the linearity property of inner product, we have

$$\left| \|\Phi h_{T_0 \cup T_1}\|_2^2 + \left\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right| \leq 2\epsilon\sqrt{1 + \delta_{2s}} \, \|h_{T_0 \cup T_1}\|_2 \tag{19}$$

Now, adding the inequalities (18) and (19), we get

$$\left| \|\Phi h_{T_0 \cup T_1}\|_2^2 + \left\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right| + \left| \left\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right| \leq 2\epsilon\sqrt{1 + \delta_{2s}} \, \|h_{T_0 \cup T_1}\|_2 +$$

$$\delta_{2s} \sqrt{2} \, \|h_{T_0 \cup T_1}\| \left( \sum_{j \geq 2} \|h_{T_j}\|_2 \right)$$

We know from the triangle inequality that

$$\|\Phi h_{T_0 \cup T_1}\|_2^2 \leq \left| \|\Phi h_{T_0 \cup T_1}\|_2^2 + \left\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right| + \left| -\left\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \right\rangle \right|$$

Using this, we finally have,

$$\|\Phi h_{T_0 \cup T_1}\|_2^2 \leq \|h_{T_0 \cup T_1}\|_2 \left( 2\epsilon\sqrt{1 + \delta_{2s}} + \sqrt{2} \, \delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_2 \right) \tag{20}$$

From the restricted isometry property, we have

$$(1 - \delta_{2s}) \, \|h_{T_0 \cup T_1}\|_2^2 \leq \|\Phi h_{T_0 \cup T_1}\|_2^2$$

Hence, we have

$$(1 - \delta_{2s}) \, \|h_{T_0 \cup T_1}\|_2^2 \leq \|\Phi h_{T_0 \cup T_1}\|_2^2 \leq \|h_{T_0 \cup T_1}\|_2 \left( 2\epsilon\sqrt{1 + \delta_{2s}} + \sqrt{2} \, \delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_2 \right) \tag{21}$$

## 1.13

By combining the inequality (6) with RHs of (21), we have

$$\|h_{T_0 \cup T_1}\|_2 \left( 2\epsilon\sqrt{1 + \delta_{2s}} + \sqrt{2} \, \delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_2 \right) \leq \|h_{T_0 \cup T_1}\|_2 \left( 2\epsilon\sqrt{1 + \delta_{2s}} + \sqrt{2}\delta_{2s} \, s^{-1/2} \, \|h_{T_0^c}\|_1 \right)$$

Using the above inequality with (21), we have

$$(1 - \delta_{2s}) \, \|h_{T_0 \cup T_1}\|_2^2 \leq \|h_{T_0 \cup T_1}\|_2 \left( 2\epsilon\sqrt{1 + \delta_{2s}} + \sqrt{2}\delta_{2s} \, s^{-1/2} \, \|h_{T_0^c}\|_1 \right)$$

On simplifying the above inequality, we get

$$\|h_{T_0 \cup T_1}\|_2 \leq \alpha\epsilon + \rho s^{-1/2} \, \|h_{T_0^c}\|_1, \quad \alpha \equiv \frac{2\sqrt{1 + \delta_{2s}}}{1 - \delta_{2s}}, \rho \equiv \frac{\sqrt{2}\delta_{2s}}{1 - \delta_{2s}} \tag{22}$$

## 1.14

We now conclude from (22) and (12) that

$$\|h_{T_0 \cup T_1}\|_2 \leq \alpha\epsilon + \rho s^{-1/2} \|h_{T_0}\|_1 + 2\rho e_0 \leq \alpha\epsilon + \rho \|h_{T_0}\|_2 + 2\rho e_0$$

As $T_0$ and $T_1$ are disjoint and $h_{T_0}$ and $h_{T_1}$ are both $s$ sparse, we can say that

$$\|h_{T_0}\|_2 \leq \|h_{T_0} + h_{T_1}\|_2 = \|h_{T_0 \cup T_1}\|_2$$

Using this in the above inequality, we get

$$\|h_{T_0 \cup T_1}\|_2 \leq \alpha\epsilon + \rho \|h_{T_0 \cup T_1}\|_2 + 2\rho e_0$$

On simplifying, we get

$$\|h_{T_0 \cup T_1}\|_2 \leq (1 - \rho)^{-1} (\alpha\epsilon + 2\rho e_0) \tag{23}$$

## 1.15

We can write $h$ as

$$h = h_{T_0 \cup T_1} + h_{(T_0 \cup T_1)^c}$$

Hence, by triangle inequality, we have

$$\|h\|_2 \leq \|h_{T_0 \cup T_1}\|_2 + \|h_{(T_0 \cup T_1)^c}\|_2$$

But from (13), we know that

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq \|h_{T_0}\|_2 + 2e_0 \leq \|h_{T_0 \cup T_1}\| + 2e_0$$

Combining the above two inequalities, we get

$$\|h\|_2 \leq 2 \|h_{T_0}\|_2 + 2e_0 \tag{24}$$

Now combining (24) with (23) we get

$$\|h\|_2 \leq 2 \|h_{T_0}\|_2 + 2e_0 \leq 2 (1 - \rho)^{-1} (\alpha\epsilon + (1 + \rho)e_0) \tag{25}$$

## 1.16

We know that

$$\|h_{T_0}\|_1 \leq \rho \|h_{T_0^c}\|_1$$

Adding $\|h_{T_0^c}\|_1$ on both sides, we get

$$\|h\|_1 \leq (1 + \rho) \|h_{T_0^c}\|_1 \tag{26}$$

Using the result

$$\|h_{T_0^c}\|_1 \leq 2(1 - \rho)^{-1} \|x_{T_0^c}\|_1$$

in (26), we get

$$\|h\|_1 \leq 2(1 + \rho)(1 - \rho)^{-1} \|x_{T_0^c}\|_1 \tag{27}$$

# Question 2

Actual Barbara image is shown below.



Figure 1: Barbara image

$$\text{RMSE} = ||X(:) - \hat{X}(:)||_2 / ||X(:)||_2$$

## 2.a

**Root Mean square error due to noise: 0.0146**
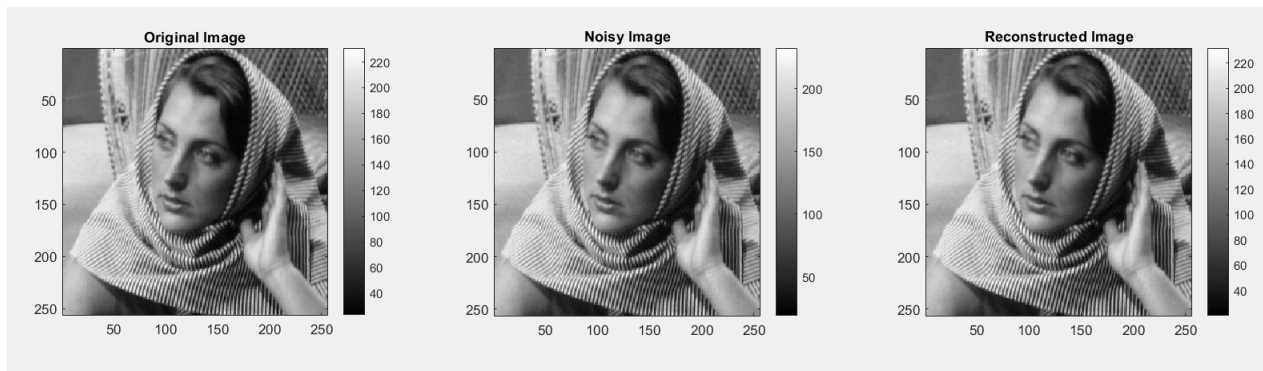**Root Mean square error after reconstruction: 0.0135**



Figure 2: All figures at one place for easier observation

## 2.b

**Root Mean square error after reconstruction: 0.406**



Figure 3: Reconstructed Barbara Image using DCT as the sparsing basis

## 2.c

**Root Mean square error after reconstruction: 0.5678**
This code took almost 1.5hrs to compile even when the number of iterations for ista was 20.



Figure 4: Reconstructed Barbara Image using DWT as the sparsing basis

## 2.d

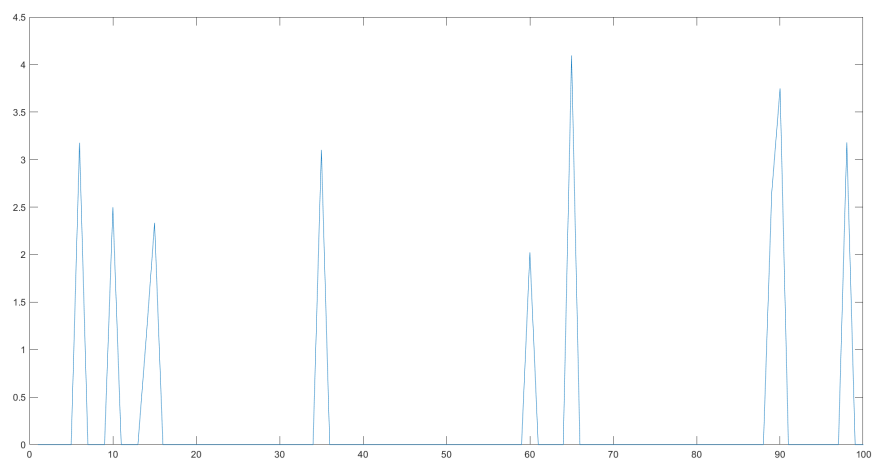**Root Mean square error after reconstruction: 0.853**
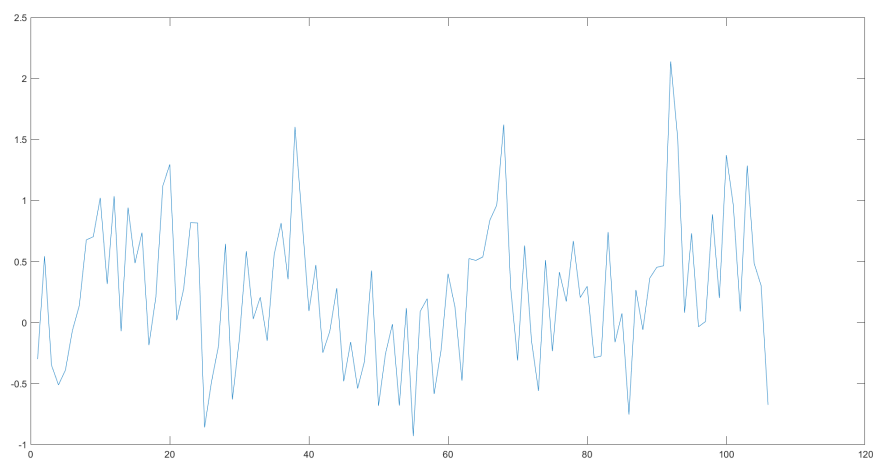
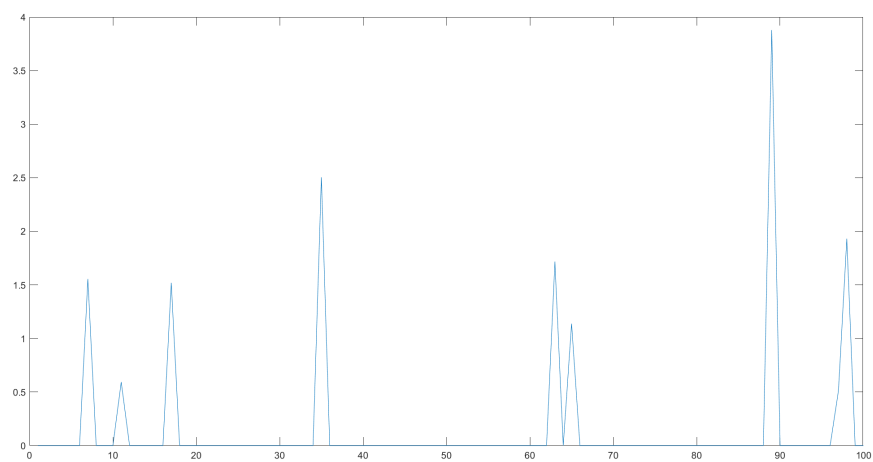Figure 5: Original sparse signal



Figure 6: Noisy Convolved Signal



Figure 7: Reconstructed Signal

# Question 3

## 3.a

The oracular solution will be nothing but the solution of the linear equation

$$y = \mathbf{\Phi}_S x_S$$

which is nothing but

$$\tilde{x} = \mathbf{\Phi}_S^\dagger y$$

## 3.b

$$
\begin{aligned}
\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_2 &= \left\|\mathbf{\Phi}_S^\dagger \boldsymbol{y} - \boldsymbol{x}\right\|_2 \\
&= \left\|\mathbf{\Phi}_S^\dagger(\mathbf{\Phi}\boldsymbol{x} - \boldsymbol{\eta}) - \boldsymbol{x}\right\|_2 \\
&= \left\|\mathbf{\Phi}_S^\dagger \boldsymbol{\eta}\right\|_2 \leq \left\|\mathbf{\Phi}_S^\dagger\right\|_2 \|\boldsymbol{\eta}\|_2
\end{aligned}
$$

## 3.c

We know that if $\lambda_i$ are the non-zero singular values of $\mathbf{\Phi}$, then the singular values of $\mathbf{\Phi}_S^\dagger$ will be $\lambda_i^{-1}$. Also, if the eigen values of $\mathbf{\Phi}$ are denoted as $e_i$, then we know $\lambda_i = \sqrt{e_i}$.

As $\mathbf{\Phi}$ satisfies restricted isometry property of order $2k$, we have

$$(1 - \delta_{2k}) \|\boldsymbol{x}\|_2^2 \leq \|\mathbf{\Phi}\boldsymbol{x}\|_2^2 \leq (1 + \delta_{2k}) \|\boldsymbol{x}\|_2^2$$

Re-writing $\|A\|_2^2 = A^T A$, we can say that

$$1 - \delta_{2k} \leq e_i \leq 1 + \delta_{2k}$$

which is equivalent to

$$\frac{1}{\sqrt{1 + \delta_{2k}}} \leq \frac{1}{\lambda_i} \leq \frac{1}{\sqrt{1 - \delta_{2k}}}$$

Hence we can say that the largest singular value of $\mathbf{\Phi}_S^\dagger$ lies between $\frac{1}{\sqrt{1 + \delta_{2k}}}$ and $\frac{1}{\sqrt{1 - \delta_{2k}}}$

## 3.d

From Noisy recovery theorem, we have

$$\|x^* - x\|_2 \leq C_0 k^{-1/2} \|x - \tilde{x}\|_1 + C_1 \epsilon$$

We know that for a $2k$ sparse vector $x - \tilde{x}$, we have

$$\|x - \tilde{x}\|_1 \leq \sqrt{2k} \|x - \tilde{x}\|_2$$

Hence, we have

$$\|x^* - x\|_2 \leq \sqrt{2} C_0 \|x - \tilde{x}\|_2 + C_1 \epsilon$$

11

But we have

$$\|x - \tilde{x}\|_2 \leq \frac{\epsilon}{\sqrt{1 - \delta_{2k}}}$$

Hence, we will have

$$\|x^* - x\|_2 \leq C_2 \frac{\epsilon}{\sqrt{1 - \delta_{2k}}}$$

where

$$C_2 = C_0 \sqrt{2} + C_1 \sqrt{1 - \delta_{2k}}$$

which is a constant factor.

# Question 4

As derived in lecture, we have the following relationship:

$$\delta_s = max\{1 - \lambda_{min}, \lambda_{max} - 1\}$$

where, $\lambda_{max}$ and $\lambda_{min}$ the maximum and minimum eigenvalues of the matrix $(\boldsymbol{A_\Gamma})^T(\boldsymbol{A_\Gamma})$ over all sets $\Gamma$. Here, $\Gamma$ is an index set of columns of $\boldsymbol{A}$ whose cardinality is less than or equal to s(sparsity of the vector $\theta$) and $\boldsymbol{A_\Gamma}$ denotes a sub matrix obtained by retaining those columns of $\boldsymbol{A}$ whose index belongs to $\Gamma$.

Now let us take $t \geq s$. In this case all the possible sets whose cardinality is less than $t$ will be covered by $\Gamma$. Thus, $\Gamma$ will cover sets with cardinality less than or equal to $s$ also. When we take maximum and minimum Eigen-values over all such sets then we have the following two relationships:

$$\lambda_{s,max} \leq \lambda_{t,max}; \quad \lambda_{t,min} \leq \lambda_{s,min}$$

which is obvious as searching for maximum and minimum values over a larger set will give us more rigorous values. Thus we have:

$$\lambda_{s,max} - 1 \leq \lambda_{t,max} - 1; \quad 1 - \lambda_{t,min} \geq 1 - \lambda_{s,min}$$

which finally gives:

$$max\{1 - \lambda_{t,min}, \lambda_{t,max} - 1\} \geq max\{1 - \lambda_{s,min}, \lambda_{s,max} - 1\}$$

Thus, we have $\delta_t \geq \delta_s$ for $t \geq s$.

# Question 5

- Title: Boolean Compressed Sensing and Noisy Group Testing

- Link: https://arxiv.org/pdf/0907.1061.pdf

- Objective Function:
  Here, we try to maximize the likelihood of decoder which in essense minimizes the error probability. For $N$ items and a defective set of size $K$, the decoder goes through all $\binom{N}{K}$

possible sets of size $K$ and chooses the set that is most likely. If we denote the tests' outcomes as $Y^T$, we choose $\omega^*$ for which

$$p(Y^T | \boldsymbol{X}_{S_{\omega^*}}) > p(Y^T | \boldsymbol{X}_{S_\omega}); \quad \forall \omega \neq \omega^*$$

Here

- $N$ is the total number of items, $k$ is the known number of defectives (or positive items), $p$ denotes the probability that an item is part of a given test, and $T$ is the total number of tests

- Codewords: For the $j$-th term, $X_j^T$ is a binary vector $\in \{0,1\}^T$, with the $t$-th entry $X_j(t) = 1$ if the $j$-th item is pooled in test $t$, and 0 otherwise. Following an information theoretic convention, we call it the $j$-th codeword. The observation vector $Y^T$ is a binary vector of length $T$, with entries equal to 1 for the tests with positive outcome. Similarly $Y(t)$ denotes the $t$-th component of the vector $Y^T$.

- $\boldsymbol{X} \in \{0,1\}^{N \times T}$ is the measurement matrix, or the codebook, which is a collection of $N$ codewords defining the pool design, i.e., the assignment of items to tests.

$$\boldsymbol{X} = [X_1^T; X_2^T; \ldots; X_N^T]$$

- Given a subset $S \subset \{1, 2, \ldots N\}$ with cardinality $|S|$, the matrix $\boldsymbol{X}_S$ is an $|S| \times T$ matrix formed from the rows indexed by $S$. In other words, $\boldsymbol{X}_S$ denotes the codewords (each of length $T$) corresponding to the items in $S$. Similarly, $X_S$ denotes a vector, whose components are restricted to the set of components indexed by $S$. Thus, $X_S$ is a column of the matrix $\boldsymbol{X}_S$. When indexing by test is needed, $X_S(t)$ is used to specifically denote the $t$-th column of the matrix $\boldsymbol{X}_S$, and $X_j(t)$ is the $t$-th component of the vector $X_j^T$.

- Index the different sets of items of size $K$ as $S_\omega$ with index $\omega$. Since there are $N$ items in total, there are $\binom{N}{K}$ such sets, hence

$$\omega \in \mathcal{I} = \left\{ 1, 2, \ldots, \binom{N}{K} \right\}$$

- Differences between Tapestry and the proposed method:

  1. Tapestry models the task from a linear algebra perspective whereas the proposed method models the task from an Information Theory perspective making use of channel coding concepts and probability and random processes.

  2. Tapestry solves a constrained convex optimization problem whereas the proposed method solves an un-constrained optimization problem.

  3. In Tapestry pooling, we model noise as a Poisson random variable whereas in the proposed method they model it as a Bernoulli random variable.

# Question 6

Let $\boldsymbol{x}_{min}$ be the vector which minimizes the lasso cost function. Thus we have:

$$J'(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_{min}} = 2||\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}||_2|_{\boldsymbol{x}=\boldsymbol{x}_{min}} \frac{d||\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}||_2}{d\boldsymbol{x}}|_{\boldsymbol{x}=\boldsymbol{x}_{min}} + \lambda \frac{d||\boldsymbol{x}||_1}{d\boldsymbol{x}}|_{\boldsymbol{x}=\boldsymbol{x}_{min}} = 0$$

As hinted by the question let $||\boldsymbol{y} - \boldsymbol{\Phi x}||_2|_{\boldsymbol{x}=\boldsymbol{x}_{min}} = \epsilon'$. Thus, the condition finally becomes:

$$-\frac{2}{\lambda}\epsilon'\frac{d||\boldsymbol{y} - \boldsymbol{\Phi x}||_2}{d\boldsymbol{x}}|_{\boldsymbol{x}=\boldsymbol{x}_{min}} = \frac{d||\boldsymbol{x}||_1}{d\boldsymbol{x}}|_{\boldsymbol{x}=\boldsymbol{x}_{min}} \tag{28}$$

Now we have three cases:

**Case1:** A global minima was found for $||\boldsymbol{x}||_1$ and $\epsilon' = 0$. In this case exact reconstruction was possible as $||\boldsymbol{y} - \boldsymbol{\Phi x}_{min}||_2 = 0$ and thus this solution would also satisfy the P1 problem and we can have $\epsilon$ as any positive value.

**Case2:** A global minima was found for $||\boldsymbol{x}||_1$ and $\frac{d||\boldsymbol{y}-\boldsymbol{\Phi x}||_2}{d\boldsymbol{x}}|_{\boldsymbol{x}=\boldsymbol{x}_{min}} = 0$. In this case exact reconstruction was not possible but we achieved the best possible reconstruction as global minima for $||\boldsymbol{y} - \boldsymbol{\Phi x}||_2$ was also achieved. For this case we can set $\epsilon$ as any value greater than $\epsilon'$. Thus, when solving P1 problem with such constraint, we will be guaranteed that we are searching and finally finding the best possible solution and reconstruction. Therefore, solution to the lasso problem would also solve the P1 problem given that $\epsilon \geq \epsilon'$.

**Case3:** When both sides of the Eq(28) were non-zero. This means that the cost function is minimized at $\boldsymbol{x} = \boldsymbol{x}_{min}$ which is not the global minimum for our function in the P1 problem. The P1 problem is:

$$min_{\boldsymbol{x}}||\boldsymbol{x}||_1 \ s.t \ ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2 \leq \epsilon$$

We know that the solution to such constrained optimization problem lies at the boundary which is explained clearly in the fig. given below.
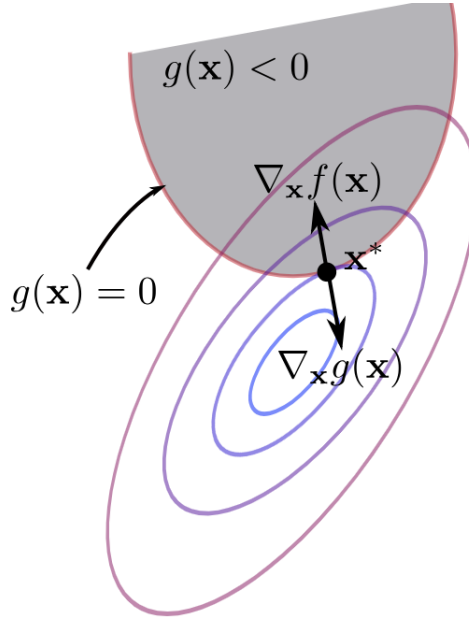


Figure 8: Figure taken from here

Thus, we can apply Lagrangian Multiplier method to solve the optimization, which gives:

$$c\frac{d(||\boldsymbol{y} - \boldsymbol{\Phi x}||_2 - \epsilon)}{d\boldsymbol{x}} + \frac{d||\boldsymbol{x}||_1}{d\boldsymbol{x}} = 0$$

where c is Lagrangian Multiplier constant and $\epsilon$ is some constant whose derivative is zero. Looking at this equation and the Eq(28) we see that if we put $\epsilon = \epsilon'$ then c will become $-\frac{2}{\lambda}\epsilon'$

and we will be solving the same set of equations to get the elements of vector $\boldsymbol{x}$. Hence, we have proved the existence of the constraint parameter $\epsilon$ such that the solution corresponding to lasso will be solution for P1 problem also.

Therefore, we see for any case there exist some $\epsilon$ such that the solution to lasso problem will also be a solution to P1 problem.