

CS754 Assignment 2 Report

Arpon Basu
Shashwat Garg

Spring 2022

Contents

1	Problem 1	2
1.1	Given $\delta_{2s} = 1$	2
1.2	Triangle Inequality	2
1.3	Relation between Norms	3
1.4	Summing up the norms	4
1.5	Extended Triangle Inequality	4
1.6	Reverse Triangle Inequality, for L1 norm	5
1.7	Rearranging	5
1.8	Using Cauchy-Schwartz	6
1.9	Using RIP	6
1.10	Using Lemma 2.1	7
1.11	AM-GM Inequality	7
1.12	Bounding norm of \mathbf{h}	7
1.13	Rearranging (again)	8
1.14	Condensing the inequalities	9
1.15	Bringing it all together	9
1.16	Wrapping up	10
2	Problem 2	11
3	Problem 3	12
3.1	The Submatrix Lemma	13
4	Problem 4	14
4.1	Imaging System	14
4.2	Designing the Coded Aperture	15
4.3	Optimization Technique	16
4.4	Results	16
5	Problem 5	18
6	Problem 6	18

Introduction

Welcome to our report on CS754 Assignment 2. We have tried to make this report comprehensive and self-contained. We hope reading this would give you a proper flowing description of our work, methods used and the results obtained.

Hope you enjoy reading the report. Here we go!

1 Problem 1

1.1 Given $\delta_{2s} = 1$

In this case, our RIP theorem inequality becomes-

$$0 \leq \|\Phi h\|_2^2 \leq 2 \|h\|_2^2$$

where h is a $2s$ -sparse vector.

We know that δ_{2s} is the smallest number such that the above holds for all s -sparse vectors for the matrix Φ .

This means that, unless the $\delta_{2s} = 1$ is present to take care of the upper bound on $\|\Phi h\|_2^2$, there will exist an s -sparse vector such that $\|\Phi h\|_2^2 = 0$, which implies that $\Phi h = \mathbf{0}$. Denoting the columns of Φ as $C_1, C_2 \dots C_n$ and the entries of h as $a_1, a_2, \dots a_n$, we can get that-

$$\sum_{i=1}^n C_i \cdot a_i = 0$$

If we focus only on the non-zero entries of h , we can get a linear combination of at-most $2s$ columns of Φ which sums up to 0. Thus, in this case, we can always say that $2s$ columns of Φ are linearly dependent.

Hence, when $\delta_{2s} = 1$, $2s$ columns of Φ may be linearly dependent.

1.2 Triangle Inequality

We are trying to solve the following-

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_1 \text{ where } \|y - \Phi \tilde{x}\|_2 \leq \epsilon$$

The solution we receive is called x^* and the actual, true, unknown solution is called x . Now, note that our reconstruction demands that $\|y - \Phi x_{estimated}\|_2 \leq \epsilon$

This implies that x^* will also obey the same property.

Secondly, in the premise of the Theorem, we assume that the noise magnitude is less than ϵ . This is one of the factors we take into account while choosing the ϵ value. Hence, by choice of ϵ in noise, we get that $\|y - \Phi x\|_2 \leq \epsilon$.

Finally, with these two tools, we can use the Triangle Inequality as follows-

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

Using the two arguments above, we can say that-

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

Thus, finally, we get that

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

Hence, proved.

1.3 Relation between Norms

We know that l_∞ norm gives out the element with the maximum magnitude in the vector. Let us assume entries of h_{T_j} to be a_1, a_2, \dots, a_n . WLOG, let us assume that, since h_{T_j} is s -sparse, a_1, a_2, \dots, a_s are non-zero and rest are 0. The re-ordering won't affect the solution since we would be taking the norm anyways. With this knowledge, we can see that-

$$\forall a_i : i \leq s, (a_i)^2 \leq \|h_{T_j}\|_\infty^2$$

Adding all these s inequalities up, since all a_i for $i > s$ are zero, we get-

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

From this, we directly get that-

$$\|h_{T_j}\|_2 \leq s^{1/2} \|h_{T_j}\|_\infty$$

Secondly, by construction, we know that every element of h_{T_j} is smaller than the every element of $h_{T_{j-1}}$. Thus, we can say the following-

$$\|h_{T_j}\|_\infty \leq |b_i|$$

where b_1, b_2, \dots, b_s are the s non-zero elements of $h_{T_{j-1}}$. Adding these up for the s inequalities, and using the fact that all b_i for $i > s$ are zero, we get-

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

This is equivalent to-

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

Hence, finally, we have obtained the required relations-

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

1.4 Summing up the norms

By just summing up the above inequality, we can say that-

$$\begin{aligned} \|h_{T_j}\|_2 &\leq s^{-1/2} \|h_{T_{j-1}}\|_1 \\ \implies \sum_{j \geq 2} \|h_{T_j}\|_2 &\leq s^{-1/2} (\|h_{T_1}\|_1 + \|h_{T_2}\|_1 + \dots) \end{aligned}$$

Also, we know that-

$$\|h_{T_0}\|_1 + \|h_{T_1}\|_1 + \|h_{T_2}\|_1 + \dots = \|h\|_1$$

Thus, we can say that

$$\|h_{T_0^c}\|_1 = \|h\|_1 - \|h_{T_0}\|_1 = \|h_{T_1}\|_1 + \|h_{T_2}\|_1 + \dots$$

Hence, finally we have-

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

1.5 Extended Triangle Inequality

By definition of h_{T_i} , we can say that, $h_{(T_0 \cup T_1)^c} = \sum_{j \geq 2} h_{T_j}$ Thus, obviously

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

Also, from previous parts, we know that

$$\sum_{j \geq 2} \|h_{T_j}\|_2 \leq s^{-1/2} \|h_{T_0^c}\|_1$$

Now, by Triangle inequality, we know that-

$$\|h_{T_2} + h_{T_3} + \dots\|_2 \leq \|h_{T_2}\|_2 + \|h_{T_3}\|_2 + \dots$$

This is equivalent to-

$$\left\| \sum_{j \geq 2} h_{T_j} \right\|_2 \leq \sum_{j \geq 2} \|h_{T_j}\|_2$$

Combining everything, we get-

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

1.6 Reverse Triangle Inequality, for L1 norm

Note that x^* is the minimizer of the l1 norm. Thus, if we have chosen a correct error bound, we will get that-

$$\|x\|_1 \geq \|x^*\|_1 \implies \|x\|_1 \geq \|x + h\|_1$$

Now from reverse triangle inequality, I can say that, $\|a + b\|_1 \geq \|a\|_1 - \|b\|_1$. Also I can split the L1-norm of $(x+h)$ into terms in T_0 and those in T_0^c .

Thus finally, can say the following, by repeatedly applying triangle Inequality. Note that this is not the L2 norm inequality, but it works similarly, due to behaviour of the $\|$ function. We are considering $(x+h)$ to be $(x-(-h))$, while applying the inequality.

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

$$\sum_{i \in T_0^c} |x_i + h_i| \geq \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1^1$$

Adding these two, along with our previous results, we get-

$$\|x\|_1 \geq \|x + h\|_1 = \sum_{i \in T_0} |x_i + h_i| + \sum_{i \in T_0^c} |x_i + h_i| \geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1$$

1.7 Rearranging

Now, we have the following-

$$\begin{aligned} \|x\|_1 &\geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1 \\ \implies \|h_{T_0^c}\|_1 &\leq \|h_{T_0}\|_1 + \|x\|_1 - \|x_{T_0}\|_1 + \|x_{T_0^c}\|_1 \end{aligned}$$

Now, note that $x_{T_0^c} = x - x_{T_0}$. Thus, using triangle inequality, we can say that-

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

Combining the above two, we get-

$$\|h_{T_0^c}\|_1 \leq \|h_{T_0}\|_1 + \|x\|_1 - \|x_{T_0}\|_1 + \|x_{T_0^c}\|_1 \leq \|h_{T_0}\|_1 + \|x_{T_0^c}\|_1 + \|x_{T_0^c}\|_1 = \|h_{T_0}\|_1 + 2\|x_{T_0^c}\|_1$$

Hence, proved.

¹(PS-Please note how the inequality was applied in the opposite sense)

1.8 Using Cauchy-Schwartz

In section 1.5, we have already seen the following result-

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

In the previous section, we saw the following-

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

Also, from the Cauchy-Schwartz inequality, we can say that the dot product of two vectors is less than the product of their L2 norms. Thus, we can say that-

$$\|h_{T_0}\|_1 \leq s^{1/2} \|h_{T_0}\|_2$$

This is obtained by taking a signum function on h_{T_0} , which returns a vector with +1 in places where h_{T_0} is positive and -1 in places where h_{T_0} is negative. A dot product of the two, gives us the L1 norm on the LHS, and the other 2 L2 norms on the RHS. The inequality is the classic $\cos x \leq 1$

Finally, we can say that-

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

This is equivalent to the desired result, when $e_0 = s^{-1/2} \|x - x_s\|_1$

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

1.9 Using RIP

We now use RIP for the first time in the proof. First, from definition of dot product or Cauchy-Schwartz inequality, we can say that-

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

We will now put individual bounds on the two expressions. Firstly, from RIP of order $2s$, we can say that-

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

Next, we can say the following, using results derived above-

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

Combining the two, we get the following-

$$\langle \Phi h_{T_0 \cup T_1} | \Phi h \rangle \leq \|\Phi h_{T_0 \cup T_1}\|_2 \|\Phi h\|_2 \leq 2\epsilon(1 + \delta_{2s})^{1/2} \|h_{T_0 \cup T_1}\|_2$$

1.10 Using Lemma 2.1

We are not giving the proof of Lemma 2.1 again. We will invoke it on h_{T_0} and h_{T_j} . Both are s -sparse vectors by definition. Thus, we can apply the lemma directly to get the result. The lemma states-

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

for all x and x' which are s -sparse. We just plug in x as h_{T_0} and x' as h_{T_j} , to get-

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

Hence, proved.

1.11 AM-GM Inequality

Note that-

$$\|h_{T_0 \cup T_1}\|_2^2 = \sum_{i \in T_0 \cup T_1} h_i^2 = \sum_{i \in T_0} h_i^2 + \sum_{j \in T_1} h_j^2 = \|h_{T_1}\|_2^2 + \|h_{T_0}\|_2^2$$

Also, note that, from AM-GM inequality, we have that-

$$\frac{A+B}{2} \geq \sqrt{A \cdot B}$$

Substituting $A = \|h_{T_0}\|_2^2$ and $B = \|h_{T_1}\|_2^2$, we get that-

$$\frac{\|h_{T_0}\|_2^2 + \|h_{T_1}\|_2^2}{2} \geq \sqrt{\|h_{T_0}\|_2^2 \cdot \|h_{T_1}\|_2^2}$$

Finally, we can say that-

$$\|h_{T_0 \cup T_1}\|_2^2 = \frac{\|h_{T_0}\|_2^2}{2} + \frac{\|h_{T_1}\|_2^2}{2} + \frac{\|h_{T_0}\|_2^2}{2} + \frac{\|h_{T_1}\|_2^2}{2} \geq \frac{\|h_{T_1}\|_2^2}{2} + \frac{\|h_{T_0}\|_2^2}{2} + \sqrt{\|h_{T_0}\|_2^2 \cdot \|h_{T_1}\|_2^2}$$

We can write the same as-

$$2 \cdot \|h_{T_0 \cup T_1}\|_2^2 \geq (\|h_{T_0}\|_2 + \|h_{T_1}\|_2)^2$$

Taking the square-root, we have-

$$\sqrt{2} \cdot \|h_{T_0 \cup T_1}\|_2 \geq \|h_{T_0}\|_2 + \|h_{T_1}\|_2$$

Hence, shown.

1.12 Bounding norm of h

Firstly, from the RIP property, we have- From the restricted isometry property, we have

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

For the second part, we need some more results. Note that we can use the Triangle Inequalities, 1.10 and 1.11 to show the following-

$$\langle \Phi h_{T_0} + \Phi h_{T_1} | \Phi h_{T_j} \rangle \leq \langle \Phi h_{T_0} | \Phi h_{T_j} \rangle + \langle \Phi h_{T_1} | \Phi h_{T_j} \rangle$$

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

Combining the two inequalities above and adding the over j-index.

$$\begin{aligned} \langle \Phi h_{T_0} + \Phi h_{T_1} | \sum_{j \geq 2} \Phi h_{T_j} \rangle &\leq \delta_{2s} \sqrt{2} \|h_{T_0 \cup T_1}\| \left(\sum_{j \geq 2} \|h_{T_j}\|_2 \right) \\ \implies \langle \Phi h_{T_0 \cup T_1} | \sum_{j \geq 2} \Phi h_{T_j} \rangle &\leq \delta_{2s} \sqrt{2} \|h_{T_0 \cup T_1}\| \left(\sum_{j \geq 2} \|h_{T_j}\|_2 \right) \end{aligned}$$

Using 1.9, we can invoke the following-

$$\langle \Phi h_{T_0 \cup T_1} | \Phi h \rangle \leq (1 + \delta_{2s})^{1/2} \|h_{T_0 \cup T_1}\|_2$$

which is the same as-

$$\langle \Phi h_{T_0 \cup T_1} | \Phi h_{T_0 \cup T_1} + \sum_{j \geq 2} \Phi h_{T_j} \rangle \leq 2\epsilon (1 + \delta_{2s})^{1/2} \|h_{T_0 \cup T_1}\|_2$$

Invoking the triangle inequality, I can say that-

$$\langle \Phi h_{T_0 \cup T_1} | \Phi h_{T_0 \cup T_1} \rangle \leq \langle \Phi h_{T_0 \cup T_1} | \sum_{j \geq 2} \Phi h_{T_j} \rangle + \langle \Phi h_{T_0 \cup T_1} | \Phi h_{T_0 \cup T_1} + \sum_{j \geq 2} \Phi h_{T_j} \rangle$$

Thus, using the two inequalities derived above, we get that-

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

Thus, we finally have the following-

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

Hence, justified.

1.13 Rearranging (again)

Let us remind ourselves of the following from 1.5-

$$\sum_{j \geq 2} \|h_{T_j}\|_2 \leq s^{-1/2} \|h_{T_0^c}\|_1$$

The previous inequality said that-

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

Combining the two above, we get that-

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

Removing a factor of $\|h_{T_0 \cup T_1}\|_2^2$, we get-

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

which is the same as-

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

where, $\alpha = \frac{2\sqrt{1+\delta_{2s}}}{1-\delta_{2s}}$ and $p = \frac{\sqrt{2}\delta_{2s}}{1-\delta_{2s}}$. Hence shown.

1.14 Condensing the inequalities

From 1.7, we know that-

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

Using this in the above inequality, we can say-

$$\|h_{T_0 \cup T_1}\|_2 \leq \alpha\epsilon + ps^{-1/2} \|h_{T_0^c}\|_1 \leq \alpha\epsilon + ps^{-1/2} \|h_{T_0}\|_1 + 2ps^{-1/2} \|x_{T_0^c}\|_1$$

Now, using triangle inequality and Cauchy-Schwartz, like a lot of times before-

$$s^{-1/2} \|h_{T_0}\|_1 \leq \|h_{T_0}\|_2 \leq \|h_{T_0 \cup T_1}\|_2$$

This gives us-

$$\|h_{T_0 \cup T_1}\|_2 \leq \alpha\epsilon + p \|h_{T_0 \cup T_1}\|_2 + 2pe_0$$

where, $e_0 = s^{-1/2} \|x_{T_0^c}\|_1$

This can be easily rearranged to give-

$$\|h_{T_0 \cup T_1}\|_2 \leq (1 - p)^{-1}(\alpha\epsilon + 2pe_0)$$

1.15 Bringing it all together

By Triangle Inequality, we have that-

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

From 1.8, we know that-

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

This gives-

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

Using the previous inequality, we have-

$$\|h_{T_0 \cup T_1}\|_2 \leq (1 - p)^{-1}(\alpha\epsilon + 2pe_0)$$

Substituting $\|h_{T_0 \cup T_1}\|_2$ from above, we have-

$$\|h\|_2 \leq \|h_{T_0 \cup T_1}\|_2 + \|h_{(T_0 \cup T_1)^c}\|_2 \leq 2\|h_{T_0 \cup T_1}\|_2 + 2e_0 \leq (1-p)^{-1}(\alpha\epsilon + 2pe_0) + 2\epsilon$$

And now, finally after simplifying, we get-

$$\|h\|_2 \leq \|h_{T_0 \cup T_1}\|_2 + \|h_{(T_0 \cup T_1)^c}\|_2 \leq 2\|h_{T_0 \cup T_1}\|_2 + 2e_0 \leq 2(1-p)^{-1}(\alpha\epsilon + (1+p)e_0)$$

which is what we needed to show.

1.16 Wrapping up

By Definition, we can say that-

$$\|h\|_1 = \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1$$

Now, from our long journey of proving inequations, we can state the following-

$$\|h_{T_0}\|_1 \leq s^{1/2} \|h_{T_0}\|_2 \leq s^{1/2} \|h_{T_0 \cup T_1}\|_2$$

Since we have been given that h is in the nullspace of Φ , this means that error $\epsilon = 0$, as the case is noiseless.

Hence, from 1.13, we can say that, with $\epsilon = 0$ -

$$\begin{aligned} \|h_{T_0 \cup T_1}\|_2 &\leq ps^{-1/2} \|h_{T_0^c}\|_1 \\ \implies s^{1/2} \|h_{T_0 \cup T_1}\|_2 &\leq p \|h_{T_0^c}\|_1 \end{aligned}$$

This gives-

$$\|h_{T_0}\|_1 \leq s^{1/2} \|h_{T_0 \cup T_1}\|_2 \leq p \|h_{T_0^c}\|_1$$

Finally-

$$\|h\|_1 = \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 \leq \|h_{T_0^c}\|_1 + p \|h_{T_0^c}\|_1 = (1+p) \|h_{T_0^c}\|_1$$

From 1.7, we can get that-

$$\begin{aligned} \|h_{T_0^c}\|_1 &\leq \|h_{T_0}\|_1 + 2 \|x_{T_0^c}\|_1 \leq p \|h_{T_0^c}\|_1 + 2 \|x_{T_0^c}\|_1 \\ \implies \|h_{T_0^c}\|_1 &\leq 2(1-p)^{-1} \|x_{T_0^c}\|_1 \end{aligned}$$

Combining all, we get-

$$\|h\|_1 = \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 \leq 2(1+p)(1-p)^{-1} \|x_{T_0^c}\|_1$$

And for the last time,

Hence, proved.

2 Problem 2

(a) First of all, note that $\Phi \mathbf{x} = \Phi_S \tilde{\mathbf{x}}$, since all other entries of \mathbf{x} are zero, and hence all the corresponding columns of Φ get nullified.

Now, we introduce the notion of **pseudo-inverse** as mentioned in the question: For any matrix A (which can be rectangular too), there exists a unique **Moore-Penrose pseudo-inverse** A^\dagger which mimics many of the properties that the actual inverse satisfies. In particular, if it so happens that the columns of A are linearly independent, then $A^\dagger A = I$.

Also note that since we are assuming Φ satisfies RIC of order S , it automatically implies that any S columns of Φ are linearly independent. Thus hereon we can assume WLOG that $\Phi_S^\dagger \Phi_S = I$.

Armed with this knowledge, it's very easy to determine the oracular solution: Note that while deriving any **deterministic** theoretical solution, we must ignore the noise $\boldsymbol{\eta}$ as we have no precise way of quantifying it apart from an upper bound on it's norm. Thus, $\mathbf{y}_{\text{theoretical}} = \Phi \mathbf{x} = \Phi_S \tilde{\mathbf{x}}$, and thus

$$\tilde{\mathbf{x}} = \Phi_S^\dagger \mathbf{y}_{\text{theoretical}}$$

Note that we don't actually know the value of $\mathbf{y}_{\text{theoretical}}$. This is a shortcoming that is fixed below.

(b) Clearly

$$\begin{aligned} \mathbf{y} &= \Phi \mathbf{x} + \boldsymbol{\eta} = \Phi_S \tilde{\mathbf{x}} + \boldsymbol{\eta} \\ \Rightarrow \Phi_S^\dagger \mathbf{y} &= \Phi_S^\dagger \Phi_S \tilde{\mathbf{x}} + \Phi_S^\dagger \boldsymbol{\eta} = \tilde{\mathbf{x}} + \Phi_S^\dagger \boldsymbol{\eta} \end{aligned}$$

Treating $\Phi_S^\dagger \mathbf{y}$ as the “actual” \mathbf{x} , we get that

$$\begin{aligned} \mathbf{x} &= \tilde{\mathbf{x}} + \Phi_S^\dagger \boldsymbol{\eta} \\ \Rightarrow \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 &= \|\Phi_S^\dagger \boldsymbol{\eta}\|_2 \end{aligned}$$

Now, from the theory of **Singular Value Decomposition** we recall that the largest singular value of any matrix A could be framed as the maximum of the following optimization problem

$$\sigma_{\max} = \|A\|_2 = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2$$

Then it's a direct consequence of the above maximization that $\|\Phi_S^\dagger \boldsymbol{\eta}\|_2 \leq \|\Phi_S^\dagger\|_2 \|\boldsymbol{\eta}\|_2$, as desired.

(c) We first prove a short lemma:

Lemma: Let A be any matrix. Then the non zero singular values of A^\dagger are the reciprocals of the non zero singular values of A .

Proof: Let $A = U\Sigma V^*$ be the SVD of A . Then

$$A^\dagger = (A^* A)^{-1} A^* = ((V\Sigma^* U^*) U \Sigma V^*)^{-1} (V\Sigma^* U^*) = V(\Sigma^* \Sigma)^{-1} \Sigma^* U^* = V\Sigma^\dagger U^*$$

since U and V being unitary matrices satisfy the property that $U^* = U^{-1}$ and $V^* = V^{-1}$.

Now, from the properties of SVD we know that Σ is a rectangular matrix whose diagonal entries contain the singular values of A .

Also, by the properties of the pseudo-inverse, one can obtain that the pseudo-inverse of Σ is obtained in a similar way in which we invert a diagonal matrix, ie: Take all the non zero entries of Σ (which occur in the diagonal), take their reciprocals, and finally take the transpose of the matrix that is obtained (the transpose is a consequence of the pseudo-inverse: The pseudo-inverse of a rectangular matrix has the dimensions of its transpose).

But then by the above derivation we also saw that the singular values of A^\dagger were basically the diagonal entries of Σ^\dagger .

Thus, we obtain that the (non zero) singular values of Φ_S^\dagger are the reciprocals of the singular values of Φ_S (just plugin $A = \Phi_S$ in the previous paragraph).

Also note that for any matrix A , it's singular values are the square roots of the eigenvalues of AA^* , and thus the singular values of A^\dagger are the square roots of the reciprocals of the eigenvalues of AA^* ².

Now, since Φ is RIC with order $|S| = 2k$, from the slides (CS.Theory, Pg.101/109) we have that $\delta_{2k} = \max(\lambda_{\max} - 1, 1 - \lambda_{\min})$ where λ_{\max} is the maximum eigenvalue of $\Phi_\Gamma \Phi_\Gamma^*$ and λ_{\min} is the minimum eigenvalue of $\Phi_\Gamma \Phi_\Gamma^*$ over all possible sets $\Gamma \subseteq [n]$, $|\Gamma| \leq |S|$ (note that λ_{\max} and λ_{\min} both don't necessarily come from some same set Γ).

Then, considering the fact that $\Phi_\Gamma \Phi_\Gamma^*$ and $\Phi_\Gamma^* \Phi_\Gamma$ have the same non-zero eigenvalues³, we see that every eigenvalue of $\Phi_S^* \Phi_S$ lies in $[1 - \delta_{2k}, 1 + \delta_{2k}]$ (because note that S is just an instance of Γ and $\delta_{2k} = \max(\lambda_{\max} - 1, 1 - \lambda_{\min})$ implies that all eigenvalues, are first bounded by λ_{\max} and λ_{\min} which are in turn bounded by $1 - \delta_{2k}$ and $1 + \delta_{2k}$.

Refer to the solution of Q3 for an elaboration of the same), and consequently **EVERY** singular value of Φ_S^\dagger , which as we showed above were the square roots of the reciprocals of the eigenvalues of $\Phi_S^* \Phi_S$, must lie within $[\frac{1}{\sqrt{1+\delta_{2k}}}, \frac{1}{\sqrt{1-\delta_{2k}}}]$, as desired.

(d) Note that Thm. 3 said that $\|x^* - x\|_2 \leq \frac{C_0}{\sqrt{S}} \|x - \tilde{x}\|_2 + C_1 \epsilon$. But we already know that $\|x - \tilde{x}\|_2 \leq \frac{\epsilon}{\sqrt{1-\delta_{2k}}} \leq \frac{\epsilon}{\sqrt{2-\sqrt{2}}}$, since $\delta_{2k} \leq \sqrt{2} - 1$ for Thm. 3 to hold.

Thus, $\|x^* - x\|_2 \leq (\frac{C_0}{\sqrt{(2-\sqrt{2})S}} + C_1) \epsilon$, and since C_0 and C_1 are known bounded constants we obtain that (the errors in) both the solutions are some constant factors of ϵ , and consequently some constant factors of each other.

3 Problem 3

We establish a certain lemma first before proving the asked for result.

²note that the eigenvalues of AA^* are non-negative by the Spectral theorem since AA^* is Hermitian

³this is again a linear algebra property: For any matrix A , AA^* and A^*A have the same non zero eigenvalues

3.1 The Submatrix Lemma

Claim: Let $\mathbf{A} \in \mathbb{K}^{m \times n}$ be a matrix and let $\mathbf{A}_s \in \mathbb{K}^{m \times s}$ be a submatrix of \mathbf{A} . Then all the eigenvalues of $\mathbf{A}_s^* \mathbf{A}_s$ are also eigenvalues of $\mathbf{A}^* \mathbf{A}$.

Proof: Notice that the sampling of \mathbf{A}_s from \mathbf{A} can be done via a **selection matrix** $\mathbf{B} \in \mathbb{K}^{n \times s}$ such that $\mathbf{A}_s = \mathbf{A}\mathbf{B}$. Notice that the columns of \mathbf{B} are just “one-hot” columns corresponding to the columns of \mathbf{A} which were sampled to make \mathbf{A}_s . One of the consequences of this is that $\mathbf{B}\mathbf{B}^* = \mathbf{I}_n \in \mathbb{K}^{n \times n}$. We shall use this below.

Also, WLOG assume that \mathbf{A}_s are the first s columns of \mathbf{A} (if not, then we can always swap columns without affecting “global” matrix properties such as eigenvalues).

Then, note that

$$\begin{aligned} \mathbf{A}_s^* \mathbf{A}_s \mathbf{x} &= \lambda \mathbf{x}, \text{ where } \mathbf{x} \neq 0 \text{ is an eigenvector of } \mathbf{A}_s^* \mathbf{A}_s \\ &\Rightarrow (\mathbf{A}\mathbf{B})^* (\mathbf{A}\mathbf{B}) \mathbf{x} = \lambda \mathbf{x} \\ &\Rightarrow \mathbf{B}^* \mathbf{A}^* \mathbf{A} \mathbf{B} \mathbf{x} = \lambda \mathbf{x} \\ &\Rightarrow \mathbf{B}\mathbf{B}^* \mathbf{A}^* \mathbf{A} \mathbf{B} \mathbf{x} = \lambda \mathbf{B} \mathbf{x} \\ &\Rightarrow \mathbf{A}^* \mathbf{A} (\mathbf{B} \mathbf{x}) = \lambda (\mathbf{B} \mathbf{x}) \end{aligned}$$

Thus, if $\mathbf{B} \mathbf{x} \neq 0$, then λ is an eigenvalue of $\mathbf{A}^* \mathbf{A}$ with eigenvector $\mathbf{B} \mathbf{x}$. But note that (since we assumed that \mathbf{A}_s comprised of the first s columns of \mathbf{A}) \mathbf{B} is basically \mathbf{I}_s padded below with $(n - s)$ zero row vectors, and consequently $\mathbf{B} \mathbf{x} = [\mathbf{x}^T \mathbf{0}_{n-s}^T]^T$, where $\mathbf{0}_{n-s} \in \mathbb{K}^{n-s}$ is the zero vector.

Since $\mathbf{x} \neq 0$, $\mathbf{B} \mathbf{x}$, whose first s entries are identical to \mathbf{x} ’s, can’t be zero either, and consequently $\mathbf{B} \mathbf{x}$ is a valid eigenvector for $\mathbf{A}^* \mathbf{A}$ and hence λ is an eigenvalue of $\mathbf{A}^* \mathbf{A}$.

But now we’re almost done: As in Q2, we can say that $\delta_s = \max(\lambda_{\max} - 1, 1 - \lambda_{\min})$ implying that every non-zero eigenvalue of $\Phi_{\Gamma}^* \Phi_{\Gamma}$ ($|\Gamma| = s$)⁴ for any s column sampling of Φ must lie within $[1 - \delta_s, 1 + \delta_s]$.

In particular, for some s -sampling, either the largest eigenvalue of $\Phi_{\Gamma}^* \Phi_{\Gamma}$ is $(1 + \delta_s)$ or the smallest eigenvalue is $(1 - \delta_s)$. In either case, take that Φ_{Γ} and append any $(t - s)$ new columns of Φ to it.

For the resulting $m \times t$ matrix, by the submatrix lemma we have that it either has $(1 + \delta_s)$ or $(1 - \delta_s)$ as it’s eigenvalue, and once again by the RIC property (that all eigenvalues are bounded by $(1 \pm \delta)$), we either have that $(1 - \delta_s) \geq (1 - \delta_t)$ or $(1 + \delta_s) \leq (1 + \delta_t)$, both of which yield $\delta_s \leq \delta_t$, as desired.

⁴once again, CS Theory actually talks of the eigenvalues of $\Phi_{\Gamma} \Phi_{\Gamma}^*$, but since we know that the non-zero eigenvalues of $\Phi_{\Gamma}^* \Phi_{\Gamma}$ and $\Phi_{\Gamma} \Phi_{\Gamma}^*$ are the same, we swap the $*$ according to our wishes

4 Problem 4

Paper Details	
Title of the Paper	Coded Aperture Design in Compressive Spectral Imaging Based on Side Information
Link of the paper	Click Here
Author List	Laura Galvis, Daniel Lau, Xu Ma, Henry Arguello, and Gonzalo R. Arce
Publication Date	July 2017
Publication Venue	Applied Optics 56, 6332-6340, The Optical Society 2017

4.1 Imaging System

This paper proposes a method to use side information from an RGB sensor to design the coded-aperture patterns of a CSI imager. This is done in a way such that the high frequencies are promoted, leading to more detailed reconstructed spectral images. This means that the coded apertures are designed to follow the spatial structure such that more detailed spatial images and wavelength profiles can be reconstructed.

Basically, the side information is used as input of an edge detection algorithm to approximate a version of the edges of the spectral images.

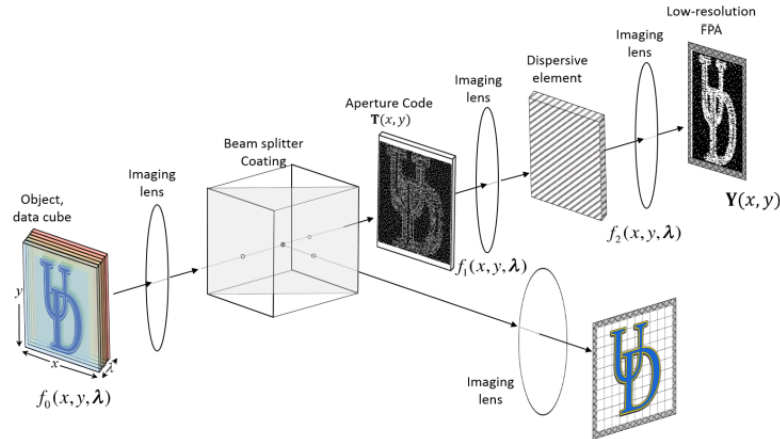


Fig. 1. Schematic architecture of CASSI with RGB side information. A beam splitter is utilized to divide the incident light into two different directions.

The base is the CASSI camera, which captures 2D projections of the scene using a coded-snapshot. A beam splitter is used to divide the incident light into the two optical paths, the CASSI and the RGB arms. The RGB arm also simultaneously captures the image information. The RGB arm is not only used for sensing, but also to achieve super-resolution during reconstruction. We'll elaborate how soon.

4.2 Designing the Coded Aperture

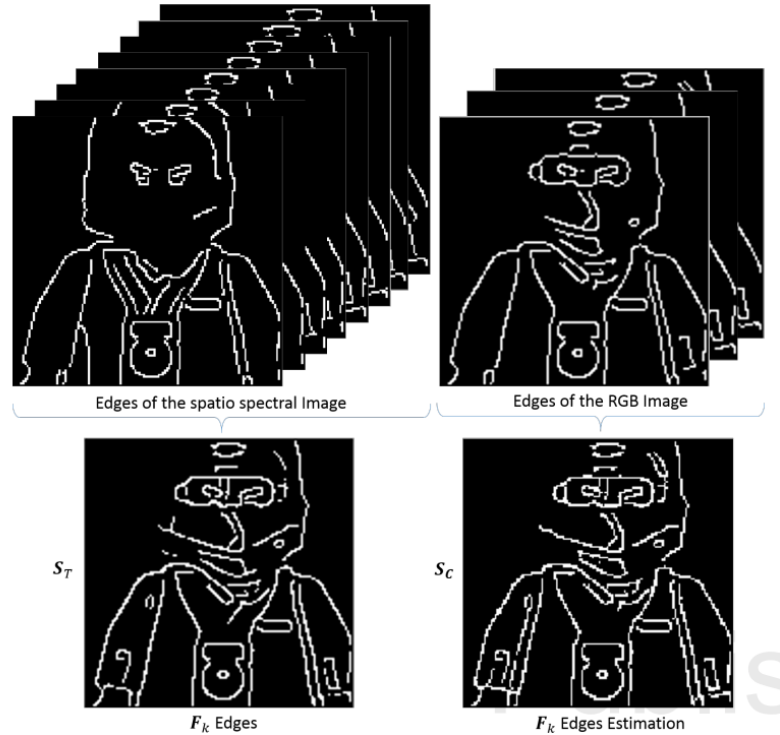
Bernoulli matrices are usually used in Compressive sensing, but they do not exploit the a-priori information of the scene. In this case, if a region dense in zero pixel values has a higher frequency image component, the reconstruction is poor. This is especially important around the edges.

Thus, this method proposes a way to improve that by using edge detectors using the side data. Using canny edge detectors, noise reduction and sobel filters to detect edges, a complete intensity gradient image is obtained for the complete data cube.

Assume that we got the edge pattern as S_C . Then the sensing matrix is designed as-

$$T = T_{b_1} \cdot S_C + T_{b_2} \cdot (1 - S_C)$$

This is where T_{b_1} and T_{b_2} are two random blue coded snapshots. Blue noise matrix is chosen as it depicts the higher frequencies well, according to the paper.



Note how the combination of the Random matrices and the edge detection matrices helps improving the density of pixels in areas of high frequencies.

Thus, by using this approach, we are optimizing the coded snapshot holes frequency to regions where it is needed to capture high frequency changes in the image. Hence, we get the sensing matrix. In the sample showed above, 87% of the edges are correctly estimated from the RGB image. The details of these regions are well-preserved.

4.3 Optimization Technique

We finally have two measurements. One by the existing RGB code and the other by the freshly designed snapshot. The two measurements are stacked together onto a single vector. The equation is written as- $\tilde{y} = \tilde{H}f + \tilde{w}$, where all three of $\tilde{y}, \tilde{H}, \tilde{w}$ are the concatenated versions from the RGB and default sensor respectively.

The optimization is the following-

$$\hat{f} = \Psi\{\operatorname{argmin}_{\theta} \|\tilde{y} - \tilde{H}\Psi\theta\|_2 + \tau\|\theta\|_1\}$$

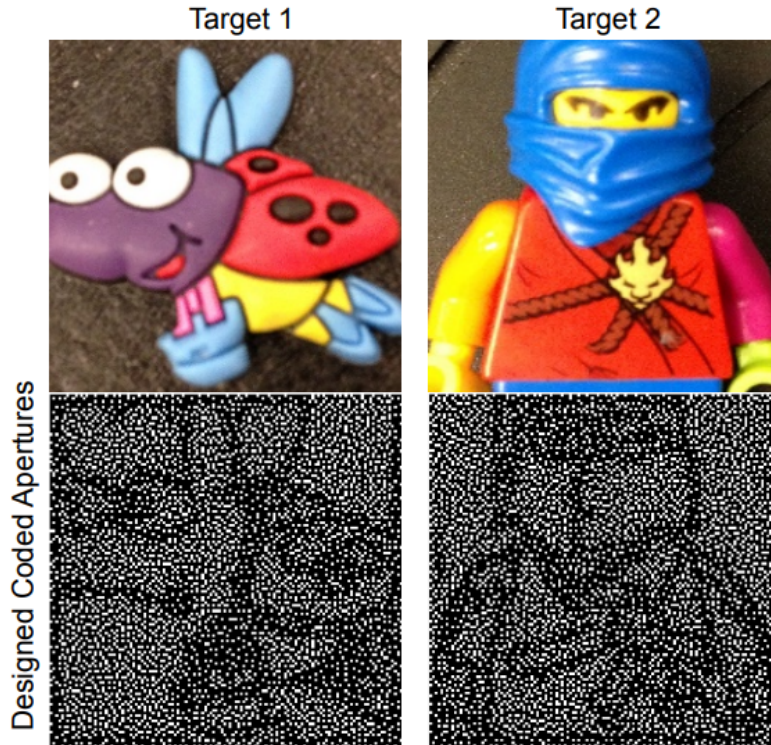
This is the familiar problem, which we have also seen in the class. The Ψ is a bit different. To quote the paper

Φ is formulated as the Kronecker product of two bases $\Phi = \Phi_1 \otimes \Phi_2$, where Φ_1 is a 2D wavelet Symmlet 8 basis and Φ_2 is playing the role of spectral sparsifier is a the 1D discrete cosine transform.

To solve this, the GSPR (gradient projection for sparse reconstruction) is used.

4.4 Results

Here is a sample sensing matrix of the following images-



The compressive sensing gradient projection for sparse reconstruction (GPSR) algorithm is used to obtain the reconstructions of the data cube. The reconstruction from this

method have lesser noise and better information around edges. Here are some samples of the same-

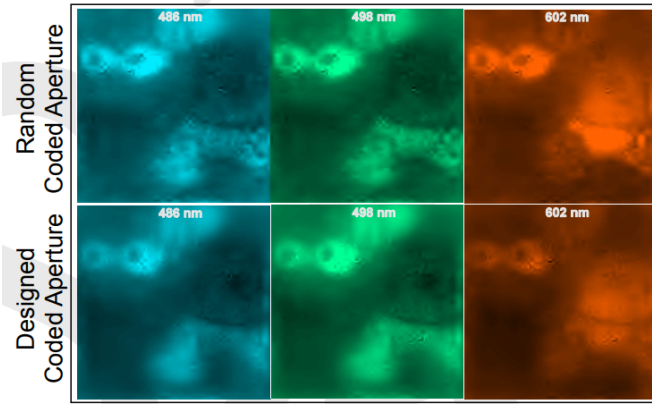


Fig. 10. Reconstruction of 3 bands for Target 1 (Top) Using random coded aperture. (Bottom) Using designed coded aperture.

We can see that the new method gives a much better Signal to Noise ratio, approximately a 3db improvement. Apart from this, due to better information of the active sights, the image is well lit as compared to random matrix.

Thus, we observe a higher intensity at high frequency points and a lower intensity at low frequency points, as is clear from the graph below. This is better for any image. Hence, this concludes our discussion of this architecture.

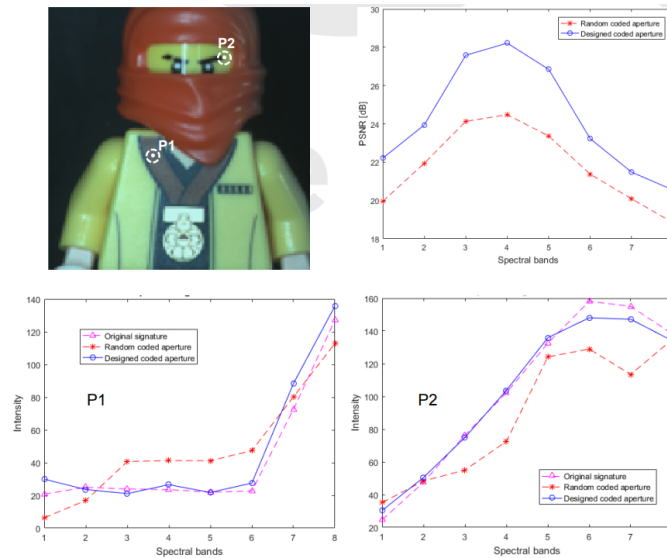


Fig. 8. Mean spectral PSNR for the eight reconstructed bands and spectral signatures for two representative spatial points, indicated as P1 and P2.

5 Problem 5

Let

$$\begin{aligned} \mathbf{x}^* &:= \min_{\mathbf{x} \in \mathbb{K}^n} (\|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1) \\ \epsilon' &:= \|\mathbf{y} - \Phi \mathbf{x}^*\|_2 \\ \mathbf{A}_t &:= \{\mathbf{x}_0 : \mathbf{x}_0 \in \mathbb{K}^n; \|\mathbf{y} - \Phi \mathbf{x}_0\|_2 \leq t\} \end{aligned}$$

Thus the sample space of the problem P1 for any given ϵ is \mathbf{A}_ϵ . Now, choose $\epsilon = \epsilon'$. Thus, $\|\mathbf{y} - \Phi \mathbf{x}\|_2 \leq \epsilon' \forall \mathbf{x} \in \mathbf{A}_{\epsilon'}$. But since the value of $\|\mathbf{y} - \Phi \mathbf{x}\|_2$ over $\mathbf{A}_{\epsilon'}$ is lesser than or equal to ϵ' , we have that their L1 norms should be greater than or equal to $\|\mathbf{x}^*\|_1$ because otherwise if there existed $\mathbf{x}' \in \mathbf{A}_{\epsilon'}$ such that $\|\mathbf{y} - \Phi \mathbf{x}'\|_2 \leq \epsilon'$ and $\|\mathbf{x}'\|_1 < \|\mathbf{x}^*\|_1$, then $J(\mathbf{x}') < J(\mathbf{x}^*)$, contradicting the minimality of \mathbf{x}^* .

Thus $\|\mathbf{x}'\|_1 \geq \|\mathbf{x}^*\|_1 \forall \mathbf{x}' \in \mathbf{A}_{\epsilon'}$ (note that $\mathbf{x}^* \in \mathbf{A}_{\epsilon'}$ too). Thus, the set of vectors with minimum L1-norms in $\mathbf{A}_{\epsilon'}$ includes \mathbf{x}^* , and consequently \mathbf{x}^* is a minimizer for problem P1 too (although note that without additional information we can't say if \mathbf{x}^* is the unique minimizer of P1 or not).

6 Problem 6

We use the linearity of expectation to find out the number of tests. Thus, the expected number of tests required for the entire population is equal to the (Number of groups) multiplied by the expected number of tests required for a single group.

Now, let $p = \frac{k}{n}$ be the probability that a randomly chosen person is infected. Then, the expected number of tests that need to be carried out can be calculated as follows:

- Suppose none of our g pool members are infected. The probability of this happening is $(1 - p)^g$, and thus the contribution of this scenario to the expected value is $(1 - p)^g \cdot 1$, ie:- after carrying out the initial pool test itself we can declare all of the pool members to be negative.
- Suppose at least one of our pool members is infected. The probability of this happening is $1 - (1 - p)^g$, and in this case we need to carry out $(g + 1)$ tests according to the Dorfman algorithm (one initial pool test, and then g individual tests). Thus, the contribution of this scenario to the expected value is $(1 - (1 - p)^g) \cdot (g + 1)$.

Thus the total expected value of the number of tests that need to be carried out for a single pool is $(1 - p)^g + (1 - (1 - p)^g) \cdot (g + 1) = 1 + g(1 - (1 - p)^g)$, and consequently the **number of tests required for the entire population is $\frac{n}{g} \cdot (1 + g(1 - (1 - p)^g)) = \frac{n}{g} + n(1 - (1 - p)^g)$** , where $p = \frac{k}{n}$.

Now, coming to the worst case, the worst case will occur when all the k patients land up in different groups: In that case, firstly in the $\frac{n}{g}$ group tests, k groups will come out to be positive, and then all patients among those k groups have to be tested to determine the actual patients.

Thus the number of tests required in the worst case is $\frac{n}{g} + kg$.

Now, to minimize the number of tests required even for the worst case, we differentiate w.r.t g to get that $\frac{1}{(g^*)^2} = \frac{k}{n} \Rightarrow g^* = \frac{1}{\sqrt{k/n}} = \sqrt{\frac{n}{k}}$, ie:- the **optimal group size in the worst case** is $\sqrt{\frac{n}{k}}$, where k is the number of people infected and n is the population size.