Assignment 2: CS 754

Md Kamran, Samiksha Das

Due 22 February 2021

Contents

Question 1

(1)

Let h be a 2s-sparse vector. If $\delta_{2s} = 1$ then from the RIP, it holds that $0 \leq \|\phi h\|_2^2 \leq 2\|h\|_2^2$. This means that the lower bound may possibly hold which implies that $\|\phi h\|_2^2 = 0 \implies \|\phi h\|_2 = 0$. This means that h lies in the nullspace of ϕ.

Now consider the set T which contains the 2s nonzero elements from h. Let h_T be the vector with values of h from the set T and let ϕ_T be the submatrix with columns of ϕ from the set T. Then we have $\|\phi_T h_T\|_2 = \|\phi h\|_2 = 0$ which implies that h_T lies in the nullspace of ϕ_T. Hence this implies that 2s columns of ϕ may be linearly dependent.

(2)

$$\|\phi(x^* - x)\|_2 = \|(\phi x^* - y) + (y - \phi x)\|_2 \leq \|\phi x^* - y\|_2 + \|y - \phi x\|_2 \tag{1}$$

follows from triangular inequality

$\|\phi x - y\|_2 \leq \epsilon$ comes from the given constraint and $\|\phi x^* - y\|_2 \leq \epsilon$ is because x^* is a feasible solution of the P1 problem. Hence

$$\|\phi x^* - y\|_2 + \|y - \phi x\|_2 = \|\phi x^* - y\|_2 + \|\phi x - y\|_2 \leq \epsilon + \epsilon = 2\epsilon \tag{2}$$

and finally we get

$$\|\phi(x^* - x)\|_2 \leq 2\epsilon \tag{3}$$

(3)

$\|h_{T_j}\|_\infty$ means the absolute maximum element of the vector h_{T_j}. Let us denote it by μ.

$$\|h_{T_j}\|_2 = \sqrt{h_{T_j1}^2 + h_{T_j2}^2 + ... + h_{T_js}^2} \leq \sqrt{\mu^2 + \mu^2 + ... + \mu^2} = s^{1/2}\|h_{T_j}\|_\infty \tag{4}$$

because the individual elements are $\leq$ the maximum element. Therefore

$$\|h_{T_j}\|_2 \leq s^{1/2}\|h_{T_j}\|_\infty \tag{5}$$

and now we have

$$\|h_{T_{j-1}}\|_1 = |h_{T_{j-1}1}| + |h_{T_{j-1}2}| + ... + |h_{T_{j-1}s}| \geq \mu + \mu + ... + \mu = s\|h_{T_j}\|_\infty \tag{6}$$

because even the largest element of h_{T_j} is smaller than the smallest element of $h_{T_{j-1}}$. Hence

$$s^{1/2}\|h_{T_j}\|_\infty \leq s^{-1/2}\|h_{T_{j-1}}\|_1 \tag{7}$$

and finally we have

$$\|h_{T_j}\|_2 \leq s^{1/2}\|h_{T_j}\|_\infty \leq s^{-1/2}\|h_{T_{j-1}}\|_1 \tag{8}$$

(4)

The inequality above is valid $\forall j > 2$ and we do the summation on both sides which gives

$$\sum_{j\geq 2} \|h_{T_j}\|_2 \leq s^{-1/2}(\|h_{T_1}\|_1 + \|h_{T_2}\|_1 + ...) \leq s^{-1/2}\|h_{T_0^C}\|_1 \tag{9}$$

This summation is equal to the latter term because h_{T_0} has the first s absolute largest terms and $h_{T_0^C}$ has the rest $n - s$ terms which can equivalently be written as the former. It is effectively the sum of all but first s largest terms.

(5)

$$\|\sum_{j\geq 2} h_{T_j}\|_2 \leq \sum_{j\geq 2} \|h_{T_j}\|_2 \tag{10}$$

follows from the extended triangular inequality and

$$\sum_{j\geq 2}\|h_{T_j}\|_2 \leq s^{-1/2}\|h_{T_0^C}\|_1 \tag{11}$$

follows from the above proven inequality and hence we have

$$\|h_{(T_0\cup T_1)^C}\|_2 \leq s^{-1/2}\|h_{T_0^C}\|_1 \tag{12}$$

(6)

$$|x_i + h_i| = |x_i - (-h_i)| \geq ||x_i| - |h_i|| \tag{13}$$

follows from reverse triangular inequality

$$||x_i| - |h_i|| \geq |x_i| - |h_i| \quad and \quad ||x_i| - |h_i|| \geq |h_i| - |x_i| \tag{14}$$

follows from the fact that $|x| \geq \pm x$. Now summation of this over $i \in T_0$ and $i \in T_0^C$ respectively and adding them both will yield

$$\sum_{i\in T_0}|x_i + h_i| + \sum_{i\in T_0^C}|x_i + h_i| \geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^C}\|_1 - \|x_{T_0^C}\|_1 \tag{15}$$

and hence we have

$$\|x\|_1 \geq \|x+h\|_1 \geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^C}\|_1 - \|x_{T_0^C}\|_1 \tag{16}$$

(7)

$$\|x_{T_0^C}\|_1 = \|x - x_{T_0}\|_1 \geq \|x\| - \|x_{T_0}\|_1 \tag{17}$$

follows from reverse triangular inequality. Now substituting the value of $\|x\|_1$ from above gives

$$\|x_{T_0^C}\|_1 \geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^C}\|_1 - \|x_{T_0^C}\|_1 - \|x_{T_0}\|_1 \tag{18}$$

which on reshuffling and cancellation gives

$$\|h_{T_0^C}\|_1 \leq \|h_{T_0}\|_1 + 2\|x_{T_0^C}\|_1 \tag{19}$$

(8) Substituting equation 19 in equation 12 gives

$$\|h_{(T_0\cup T_1)^C}\|_2 \leq s^{-1/2}(\|h_{T_0}\|_1 + 2\|x_{T_0^C}\|_1) \tag{20}$$

Now from Cauchy-Schwarz inequality we have

$$\|h_{T_0}\|_1 \leq s^{1/2}\|h_{T_0}\|_2 \tag{21}$$

and this further gives

$$\|h_{(T_0\cup T_1)^C}\|_2 \leq \|h_{T_0}\|_2 + 2e_0 \tag{22}$$

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

(9) First inequality is by definition. From RIP we have

$$\|\phi h_{(T_0\cup T_1)}\|_2 \leq \sqrt{1+\delta_{2s}}\|h_{(T_0\cup T_1)}\|_2 \tag{23}$$

Using equation 3 and equation 23 gives

$$|\langle\phi h_{(T_0\cup T_1)}, \phi h\rangle| \leq 2\epsilon\sqrt{1+\delta_{2s}}\|h_{(T_0\cup T_1)}\|_2 \tag{24}$$

(10) This inequality follows directly from the lemma 2.1 given in the paper and it follows correctly here because h_{T_0} and h_{T_j} $\forall j \geq 2$ have disjoint support sets and since both their sparsity is equal to s, this gives the factor of δ_{2s}.

$$|\langle\phi h_{T_0}, \phi h_{T_j}\rangle| \leq \delta_{2s}\|h_{T_0}\|_2\|h_{T_j}\|_2 \tag{25}$$

This also follows similarly for h_{T_1} and h_{T_j} $\forall j \geq 2$.

$$|\langle \phi h_{T_1}, \phi h_{T_j} \rangle| \leq \delta_{2s} \|h_{T_1}\|_2 \|h_{T_j}\|_2 \tag{26}$$

(11)

We know that

$$\|h_{T_1}\|_2 \leq \|h_{T_0}\|_2 \tag{27}$$

Bringing both of them on RHS and squaring them gives

$$0 \leq \|h_{T_0}\|_2^2 + \|h_{T_1}\|_2^2 - 2\|h_{T_0}\|_2 \|h_{T_1}\|_2 \tag{28}$$

Now we add $\|h_{T_0}\|_2^2 + \|h_{T_1}\|_2^2$ on both the sides and rearrange the terms which gives

$$\|h_{T_0}\|_2^2 + \|h_{T_1}\|_2^2 + 2\|h_{T_0}\|_2 \|h_{T_1}\|_2 \leq 2(\|h_{T_0}\|_2^2 + \|h_{T_1}\|_2^2) \tag{29}$$

The added term above is same as $\|h_{(T_0 \cup T_1)}\|_2^2$ and on applying square root both sides gives

$$\|h_{T_0}\|_2 + \|h_{T_1}\|_2 \leq \sqrt{2} \|h_{(T_0 \cup T_1)}\|_2 \tag{30}$$

(12)

The first inequality follows directly from RIP. For proving the second inequality we are given that

$$\|\phi h_{(T_0 \cup T_1)}\|_2^2 = \langle \phi h_{(T_0 \cup T_1)}, \phi h \rangle - \langle \phi h_{(T_0 \cup T_1)}, \sum_{j \geq 2} \phi h_{T_j} \rangle \tag{31}$$

We add equation 25 and equation 26 and then use equation 30 which gives

$$|\langle \phi h_{(T_0 \cup T_1)}, \phi h_{T_j} \rangle| \leq \sqrt{2} \delta_{2s} \|h_{(T_0 \cup T_1)}\|_2 \|h_{T_j}\|_2 \tag{32}$$

We substitute terms in equation 31 by equation 24 and equation 32 which gives

$$\|\phi h_{(T_0 \cup T_1)}\|_2^2 \leq \|h_{(T_0 \cup T_1)}\|_2 (2\epsilon \sqrt{1 + \delta_{2s}} + \sqrt{2} \delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_2) \tag{33}$$

In the above, we also used the fact that if $A \leq B - C$ then $A \leq B + C$ where A,B,C all are positive and now along with the inequality given by RIP, we finally get

$$(1 - \delta_{2s}) \|h_{(T_0 \cup T_1)}\|_2^2 \leq \|h_{(T_0 \cup T_1)}\|_2 (2\epsilon \sqrt{1 + \delta_{2s}} + \sqrt{2} \delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_2) \tag{34}$$

(13)

We substitute terms in equation 34 by equation 9 to get

$$(1 - \delta_{2s}) \|h_{(T_0 \cup T_1)}\|_2^2 \leq \|h_{(T_0 \cup T_1)}\|_2 (2\epsilon \sqrt{1 + \delta_{2s}} + \sqrt{2} \delta_{2s} s^{-1/2} \|h_{T_0^C}\|_1) \tag{35}$$

Cancel the term $\|h_{(T_0 \cup T_1)}\|_2$ on both sides and divide by $(1 - \delta_{2s})$ to finally get

$$\|h_{(T_0 \cup T_1)}\|_2 \leq \alpha \epsilon + \rho s^{-1/2} \|h_{T_0^C}\|_1 \tag{36}$$

where $\alpha = 2\sqrt{1 + \delta_{2s}}/(1 - \delta_{2s})$ and $\rho = \sqrt{2} \delta_{2s}/(1 - \delta_{2s})$

(14)

Using equation 19 and equation 21, we have

$$\|h_{T_0^C}\|_1 \leq s^{1/2} \|h_{T_0}\|_2 + 2\|x_{T_0^C}\|_1 \tag{37}$$

Now by using the following fact

$$\|h_{T_0}\|_2 \leq \|h_{(T_0 \cup T_1)}\|_2 \tag{38}$$

and substituting equation 37 into equation 36 gives us

$$\|h_{(T_0 \cup T_1)}\|_2 \leq \alpha\epsilon + \rho\|h_{(T_0 \cup T_1)}\|_2 + 2\rho e_0 \tag{39}$$

and on shuffling the terms further gives

$$\|h_{(T_0 \cup T_1)}\|_2 \leq (1-\rho)^{-1}(\alpha\epsilon + 2\rho e_0) \tag{40}$$

(15)

$$\|h\|_2 \leq \|h_{(T_0 \cup T_1)}\|_2 + \|h_{(T_0 \cup T_1)^C}\|_2 \tag{41}$$

follows from the fact that $\sqrt{A^2 + B^2} \leq A + B$ where both A and B are positive.

Now using the equation 22 and equation 38, we get

$$\|h_{(T_0 \cup T_1)}\|_2 + \|h_{(T_0 \cup T_1)^C}\|_2 \leq 2\|h_{(T_0 \cup T_1)}\|_2 + 2e_0 \tag{42}$$

Now finally using the equation 41, equation 42 and equation 40, we get

$$\|h\|_2 \leq 2(1-\rho)^{-1}(\alpha\epsilon + (1+\rho)e_0) \tag{43}$$

(16)

From equation 21, equation 38 and equation 36 with $\epsilon = 0$, we have

$$\|h_{T_0}\|_1 \leq \rho\|h_{T_0^C}\|_1 \tag{44}$$

And adding the equation 44 and equation 19 after shuffling the terms gives us

$$\|h_{T_0^C}\|_1 \leq 2(1-\rho)^{-1}\|x_{T_0^C}\|_1 \tag{45}$$

and finally the equation 44 and equation 45 gives us

$$\|h\|_1 \leq 2(1+\rho)(1-\rho)^{-1}\|x_{T_0^C}\|_1 \tag{46}$$

Question 2

a) Code included the folder
b)

Figure 1:

RMSE value is 0.5339.
c)

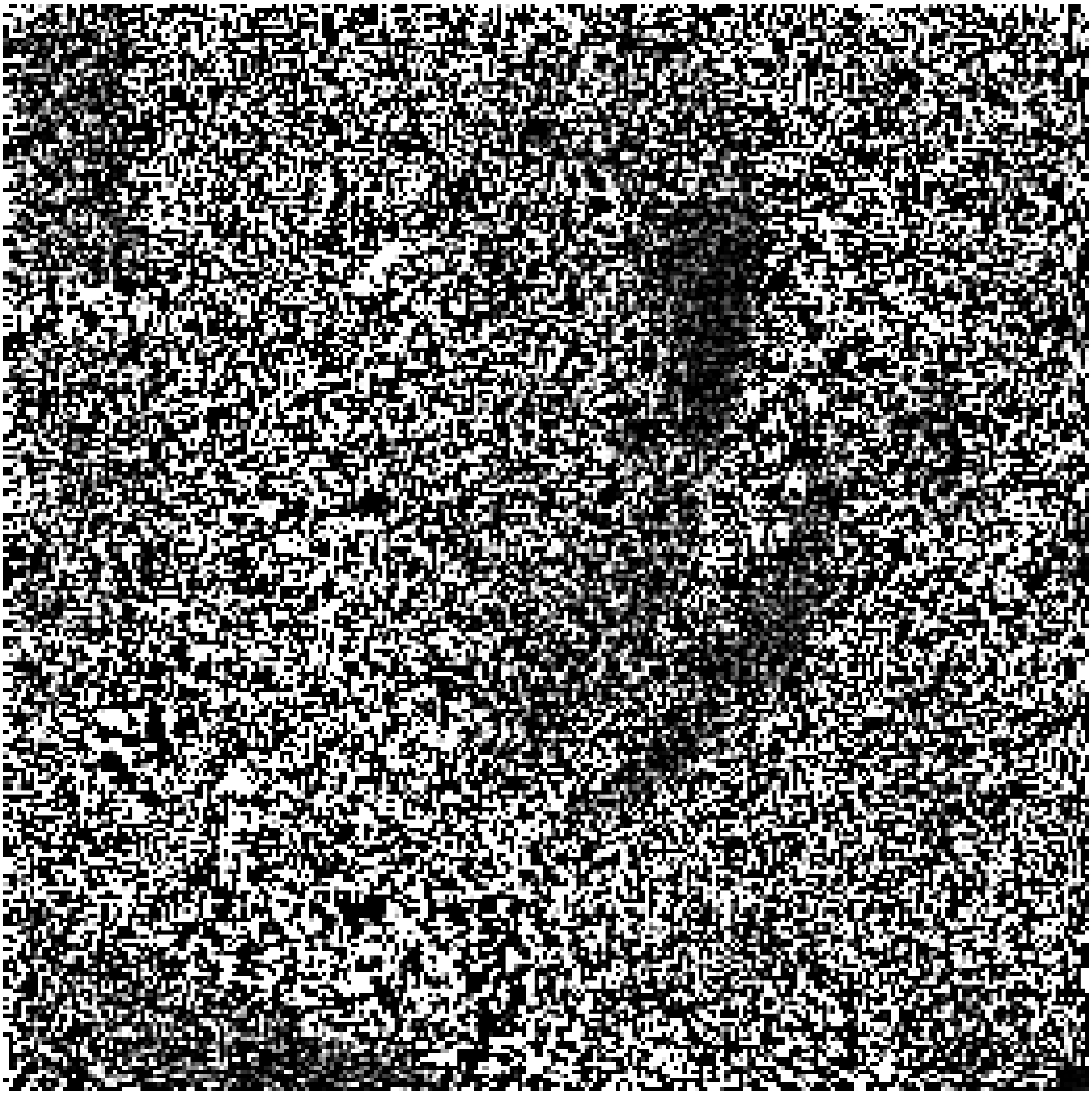

Figure 2:

RMSE value is 2.3928.

d)Code included the folder

Question 3

Here in $\Phi_s\tilde{x}$ and $\Phi_s x$ x and $tildex$ have been made of the appropriate dimensions by removing zero entries.

a) If the indices of the non-zero elements
are known in advance as S, and Φ_s is the submatrix with columns belonging to indices in S, then for

the solution $\tilde{x}$ we have :-

$$\Phi_s\tilde{x} = \Phi\tilde{x}$$

and

$$\Phi_s x = \Phi x$$

since the elements belonging to the columns not in Φ_s are zero. Then, the oracle solves

$$y = \Phi_s\tilde{x} \implies \Phi_s^T y = \Phi_s^T\Phi_s\tilde{x} \implies (\Phi_s^T\Phi_s)^{-1}\Phi_s^T y = \tilde{x}$$

where $(\Phi_s^T\Phi_s)^{-1}\Phi_s^T$ is pseudo inverse of Φ_s.

b) Also,

$$y = \Phi x + \eta \implies y = \Phi_s x + \eta$$

then from above ;

$$\Phi_s\tilde{x} = \Phi_s x + \eta \implies \Phi_s(\tilde{x} - x) = \eta \implies \Phi_s^T\Phi_s(\tilde{x} - x) = \Phi_s^T\eta \implies \tilde{x} - x = (\Phi_s^T\Phi_s)^{-1}\Phi_s^T\eta$$

Hence

$$\|\tilde{x} - x\| = \left\|(\Phi_s^T\Phi_s)\Phi_s^T\eta\right\| \leq \left\|(\Phi_s^T\Phi_s)\Phi_s^T\right\|\|\eta\|$$

c) the difference between two k sparse vectors is 2k-sparse. So we have by restricted isometric property of order 2k;

$$(1 - \delta_{2k})\|\tilde{x} - x\|^2 \leq \left\|\phi_s(\tilde{x} - x)\right\|^2 \leq (1 + \delta_{2k})\|\tilde{x} - x\|^2$$

Now, from above;

$$(1 - \delta_{2k})\left\|(\Phi_s^T\Phi_s)^{-1}\Phi_s^T\eta\right\|^2 \leq \|\eta\|^2 \leq (1 + \delta_{2k})\left\|(\Phi_s^T\Phi_s)^{-1}\Phi_s^T\eta\right\|^2$$

$$\implies \frac{1}{\sqrt{(1 + \delta_{2k})}} \leq \left\|(\Phi_s^T\Phi_s)^{-1}\Phi_s^T\eta\right\| / \|\eta\| \leq \frac{1}{\sqrt{(1 - \delta_{2k})}}$$

Hence the largest singular value of pseudo inverse of Φ_s lies between $\frac{1}{\sqrt{(1+\delta_{2k})}}$ and $\frac{1}{\sqrt{(1-\delta_{2k})}}$.

d)Now for the solution θ^* given by theorem 3 the error is :-

$$\|\theta^* - \theta\|_2 \, {}_0\|\theta - \theta_s\| + C_1\epsilon$$

from above the error in the oracle term is also bounded by $\frac{\epsilon}{\sqrt{(1-\delta_{2k})}}$. Then the sum of the error of the solution provided by oracle and $C_1\epsilon$ is hence bounded by a constant multiple of ϵ. So the error in the solution provided in theorem 3 and the oracle only differ by constants.

Question 4

We have by restricted isometric property ;

$$(1-\delta_s)\|\theta_s\|^2 \leq \|A\theta_s\|^2 \leq (1+\delta_s)\|\theta_s\|^2$$

Now, let the values of θ_s such that the lower and upper limits are achieved in the above equation be θ_{min} and θ_{max}.

$$(1-\delta_t)\|\theta_t\|^2 \leq \|A\theta_t\|^2 \leq (1+\delta_t)\|\theta_t\|^2$$

Now since both θ_{min} and θ_{max} are s sparse and hence t sparse;

$$(1-\delta_t)\|\theta_{min}\|^2 \leq \|A\theta_{min}\|^2 \leq (1+\delta_t)\|\theta_{min}\|^2 \implies (1-\delta_t) \leq \frac{\|A\theta_{min}\|^2}{\|\theta_{min}\|^2} \implies (1-\delta_t) \leq (1-\delta_s) \implies \delta_s \leq \delta_t$$

and

$$(1-\delta_t)\|\theta_{max}\|^2 \leq \|A\theta_{max}\|^2 \leq (1+\delta_t)\|\theta_{max}\|^2 \implies (1+\delta_t) \geq \frac{\|A\theta_{max}\|^2}{\|\theta_{max}\|^2} \implies (1+\delta_t) \geq (1+\delta_s) \implies \delta_s \leq \delta_t$$

Hence proved.

Question 5

(a)

Title - **Two-Stage Adaptive Pooling with RT-qPCR for COVID-19 Screening**

Link - https://arxiv.org/pdf/2007.02695.pdf

(b)

Given a signal x, we partition the n signal coordinates into q pools of size $s = n/q$. We denote x_l as the l^{th} pool of coordinates. We denote by 1_t or 0_t an all-one or an all-zero row vector of length t respectively.

We denote $m_l^{(1)}$ and $m_l^{(2)}$ as the number of measurements for l^{th} pool in first and second stage respectively and denote $A_l^{(1)}$ and $A_l^{(2)}$ as the sensing matrix of l^{th} pool in first and second stage respectively. Let $A_l = [(A_l^{(1)})^T, (A_l^{(2)})^T]^T$ and $m_l = m_l^{(1)} + m_l^{(2)}$ be defined.

For each pool $l \in [q]$, $y_l^{(1)} = 1_s \cdot x_l$ ie. $m_l^{(1)} = 1$ and $A_l^{(1)} = 1_s$. Let $z_l^{(1)} = y_l^{(1)} e_l^{(1)}$ be the noisy measurements for $l \in [q]$. WLOG we assume $L = [q] \setminus [t]$ where the first t pools are all positive, and the last $q - t$ pools are all negative.

Let $p(k)$ be the probability that x_l has k nonzero coordinates and $s - k$ zero coordinates and let $p(z_l^{(1)}|k)$ be the probability density of $z_l^{(1)}$ given above condition on x_l. For each pool $l \in [t]$, let the noisy measurements be $z_l^{(2)} = y_l^{(2)} e_l^{(2)}$ where $y_l^{(2)} = A_l^{(2)} x_l$ and let $z_l = [(z_l^{(1)})^T, (z_l^{(2)})^T]^T$ be the overall noisy measurement vector corresponding to the l^{th} pool.

We use COMP algorithm to recover a super-set estimate $\hat{S}_l$ of S_l from z_l given A_l. Let $s_l^* = |\hat{S}_l|$. Also let $I_l = \{i \in [m_l] : (z_l)_i = 0\}$. T is a k-subset of $\hat{S}_l$ where $k_{min} \leq k \leq k_{max}$ where $k_{min} = max\{\hat{k}_l - 1, 1\}$ and $k_{max} = min\{\hat{k}_l + 1, s_l^*\}$ where $\hat{k}_l$ is given as estimate of k_l where k_l is the number of nonzero coordinates in x_l.

We denote x_l^* as the sub-vector of x_l restricted to the coordinates indexed by $\hat{S}_l$, denote A_l^* the sub-matrix of A_l restricted to the rows indexed by $[m_l] \setminus I_l$ and the columns indexed by $\hat{S}_l$ and denote z_l^* as the sub-vector of z_l restricted to the coordinates indexed by $[m_l] \setminus I_l$.

The optimization problem in this scheme is:-

$$f(T) = max_{\hat{x}_l^*}\{p(\hat{x}_l^* | z_l^*)\} \tag{47}$$

where we find $\hat{x}_l^*$ with a support set T such that the conditional probability density of $\hat{x}_l^*$ given z_l^* is maximum.

(c)

- Tapestry is a single-stage scheme in which all measurements can be made in parallel, whereas this one is a two-stage scheme in which the measurements in the second stage can only be made after those in the first stage.
- The total number of measurements in the Tapestry scheme may vary for different signal realizations whereas this scheme is oblivious to different signal realizations and uses the same number of measurements always.
- The total number of nonzero entries in the overall sensing matrix of this scheme is much smaller than in the Tapestry scheme. Also the nonzero entries in each row of the sensing matrix in the Tapestry scheme are spread out whereas the nonzero entries in the sensing matrix of this scheme are localized in each row.

Question 6

We have the problem P1 : $min_x\|x\|_1$ such that $\|y - \phi x\|_2 \leq \epsilon$ and we also have the LASSO problem which minimizes the cost function $J(x) = \|y - \phi x\|_2^2 + \lambda\|x\|_1$.

Let the minimizer of J be x^* and let the minimizer of $P1$ be x'. Then we have the following:-

$\|x'\|_1 \leq \|x^*\|_1$ because x' is the minimizer. On multiplying both sides by λ ($\lambda > 0$), we have

$$\lambda\|x'\|_1 \leq \lambda\|x^*\|_1 \tag{48}$$

Also we have $\|y - \phi x'\|_2 \leq \epsilon'$ for some ϵ'. Let $\epsilon' = \|y - \phi x^*\|_2$, then we have $\|y - \phi x'\|_2 \leq \|y - \phi x^*\|_2$. On squaring both sides, we have

$$\|y - \phi x'\|_2^2 \leq \|y - \phi x^*\|_2^2 \tag{49}$$

Now on adding the above two equations, we get

$$\|y - \phi x'\|_2^2 + \lambda\|x'\|_1 \leq \|y - \phi x^*\|_2^2 + \lambda\|x^*\|_1 \tag{50}$$

This can equivalently be written as $J(x') \leq J(x^*)$ but we know that x^* is the minimizer of LASSO therefore we also have $J(x^*) \leq J(x')$. This means that $J(x') = J(x^*)$ which further means that $x' = x^*$ which eventually implies that the minimizer of LASSO is also the minimizer of the P1 problem.