Improved Bounds for Sub-linear Time Compressed Sensing for Support Recovery using Sparse-Graph Codes

Avinash Vem, Nagaraj Thenkarai Janakiraman, Krishna Narayanan
Dept. of Elect. & Comp. Engg.,
Texas A&M University,
College Station, TX, U.S.A

I. Introduction

In [1] (and in the expanded version in [2]), Li, Pawar and Ramchandran have considered the problem of recovering the support of a K-sparse, N-dimensional signal from M linear measurements in the presence of noise. Based on sparse-graph codes and a peeling decoder, they have proposed an elegant design of the measurement matrix and a recovery algorithm that is nearly optimal in terms of measurement complexity and computational complexity. In particular, they have proposed two designs - the first design requires $M = O(K \log N)$ measurements and a near-linear ($O(N \log N)$) decoding complexity, whereas the second design requires $M = O(K \log^{1.3} N)$ measurements with a sub-linear decoding complexity.

II. System Model

A classical problem of much interest is that of estimating a signal $\mathbf{x}$, which is sparse in some basis, from a noisy measurement signal $\mathbf{y}$ of smaller dimension compared to $\mathbf{x}$. More precisely,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w},$$

where $\mathbf{x}$ is an N-dimensional vector, $\mathbf{A}$ is a known $m \times N$ matrix commonly referred to as *measurement matrix* and $\mathbf{w}$ is additive noise. The unknown signal $\mathbf{x}$ is sparse in some basis and we denote the sparsity of $\mathbf{x}$ by K. If there is no noise then we refer to it as noiseless setting. It is well-known in *compressive sensing* that if $K << N$ we can recover the unknown signal in significantly fewer number of measurements compared to N. Particularly in this paper we focus on recovering the exact support of $\mathbf{x}$ defined as $\text{supp}(\mathbf{x}) := |\{i : x_i \neq 0, i \in [N]\}|$ where $\mathbf{x} = [x_1, \ldots, x_i \ldots, x_N]^T$ and $[N] = \{1, 2, \ldots N\}$. For a given scheme, given the reconstruction vector $\widehat{\mathbf{x}}$, we consider the metric probability of failure of support recovery which can be defined as

$$\mathbb{P}_F := \Pr(\text{supp}(\widehat{\mathbf{x}}) \neq \text{supp}(\mathbf{x})).$$

For this support recovery problem, under noisy settings, Wainwright [3] showed information theoretically that $O\left(K \log(\frac{N}{K})\right)$ number of measurements is necessary and sufficient for asymptotically reliable recovery. Note that this result is valid as long as the non-zero elements in $\mathbf{x}$ have a sufficiently large minimum absolute value. In view of this condition we set a minimum absolute value for the non-zero elements and further assume that they are from a discrete set. More precisely for the noisy setting we assume that all the non-zero elements of $\mathbf{x}$ belong to the set $\{Ae^{i\theta} : A \in \mathcal{A}, \theta \in \Theta\}$ where $\mathcal{A} := \{A_{\text{min}} + \rho l\}_{l=0}^{L_1}, \Theta := \{2\pi l / L_2\}_{l=0}^{L_2}$ for finite but arbitrarily large integers L_1 and L_2.

III. Prior Work

This section reviews the measurement matrix $\mathbf{A}$ proposed by Li, Pawar and Ramchandran in [1] and [2], and also summarizes their key results. To keep the discussion simple, we omit certain details and refer readers to the original paper [1] (and the expanded version [2]).

The main idea is to use the sparse-graph based construction to decompose the problem of K-sparse support recovery into 1-sparse problems and then use a simple peeling based decoder [4] to recover the support. The measurement matrix is constructed using a combination of a sparse-graph code defined by the $R \times N$ coding matrix $\mathbf{H} = [\mathbf{h_1}, \mathbf{h_2}, \cdots, \mathbf{h_N}] \in \{0,1\}^{R \times N}$ and a $P \times N$ bin-detection matrix $\mathbf{S} = [\mathbf{s_1}, \mathbf{s_2}, \cdots \mathbf{s_N}]$. The code matrix $\mathbf{H}$ defines a bipartite graph $\mathcal{G}$ with N left (variable) nodes, representing the N-length signal $\mathbf{x}$, and R right (check) nodes representing the measurements $\mathbf{y} = [\mathbf{y_1}, \mathbf{y_2}, \cdots, \mathbf{y_R}]$. Let $r_i, q_i, 1 \leq i \leq R$ be the degree of and the number of non-zero variable nodes connected to i^{th} check node. Consider an"oracle" that solves the 1-sparse problem by examining each check node observations and classifies it as a single-ton($q_i = 1$), zero-ton($q_i = 0$) or a multi-ton($q_i > 1$), and also identifies the position $\hat{k}$ and value $\widehat{x}_{\hat{k}}$ of the participating variable node if it is a single-ton. Once a singleton is identified the corresponding variable node's contribution is peeled off from other participating check nodes and this process creates new singletons. The decoding process continues until there are no more singletons. The decoding is successful if all the K non-zero elements of $\mathbf{x}$ is recovered at the end of decoding.

The $RP \times N$ measurement matrix $\mathbf{A}$ with $M = RP$ measurements is constructed by taking the row tensor product $\boxtimes$ of $\mathbf{H}$ and $\mathbf{S}$ given by

$$\mathbf{A} = \mathbf{H} \boxtimes \mathbf{S} := [\mathbf{h_1} \otimes \mathbf{s_1}, \mathbf{h_2} \otimes \mathbf{s_2}, \cdots, \mathbf{h_N} \otimes \mathbf{s_N}]$$

where $\otimes$ is the standard Kronecker product.

They explore both l-left regular and irregular ensembles for the $\mathbf{H}$ matrix construction and provide analysis for the left regular case. For the noisy setting, they have proposed three designs for $\mathbf{S}$ namely, *RandomNoisy* with near-linear decoding complexity, *BinaryNoisy* and *FourierNoisy* each with sub-linear decoding complexity. The only difference in these designs is the $\mathbf{S}$ matrix construction and hence the 1-sparse(bin detection) problem solving methodology.

The bin-detection matrix $\mathbf{S}$ for the three settings is as follows:

- *RandomNoisy:* Ensemble of $P \times N$ matrices $S = [S_{i,j}]_{P \times N}$ where $S_{i,j}$ s are i.i.d. sub-gaussian entries with zero mean and unit variance.
- *FourierNoisy*: $\mathbf{S} = [\mathbf{S_0 S_1} \cdots \mathbf{S_{P-1}}]^{\mathbf{T}}$, where $\mathbf{S_p}$ consists of $Q = O(\log^{1/3} N)$ consecutive 2^p-dyadically spaced rows from the $N \times N$ DFT matrix.
- *BinaryNoisy:* $\mathbf{S} = (-\mathbf{1})^{\mathbf{C}}$ where $C_{P \times N}$ is a binary codebook(or subset of a codebook) of a linear code with block length P.

The following theorems from [2] summarize their key results.

Theorem 1 ([2] Noiseless recovery)**.** *In the absence of noise, given any K-sparse signal $\mathbf{x}$ with $x_k \in \mathcal{X}$ for $k \in$ supp$(\mathbf{x})$, their noiseless recovery schemes achieve a vanishing failure probability $\mathbb{P}_F(O\left(\frac{1}{K}\right)) \to 0$ asymptotically in K and N with*

	M	T
Fourier noiseless	$2(1+\epsilon)K$	$O(K)$
Binary noiseless	$(1+\epsilon)K(\log_2 N + 1)$	$O(K \log N)$

where M and T are measurement cost and computational complexity respectively.

Theorem 2 ([2] Sub-linear Time Noisy Recovery)**.** . *In the presence of i.i.d. Gaussian noise with zero mean and variance σ^2, given any K-sparse signal $\mathbf{x}$ with $x_k \in \mathcal{X}$ for $k \in$ supp$(\mathbf{x})$, our noiseless recovery schemes achieve a vanishing failure probability $\mathbb{P}_F \to 0$ asymptotically in K and N with*

	M	T
Fourier noisy	$O(K \log^{1.3} N)$	$O(K \log^{1.3} N)$
Binary noisy	$O(K \log N)$	$O(K \log N)$

Theorem 3 ([2] Near-linear Time Noisy Recovery)**.** . *In the presence of i.i.d. Gaussian noise with zero mean and variance σ^2, given any K-sparse signal $\mathbf{x}$ with $x_k \in \mathcal{X}$ for $k \in$ supp$(\mathbf{x})$, the RandomNoisy scheme achieves a vanishing failure probability $\mathbb{P}_F \to 0$ asymptotically in K and N with a measurement complexity of $M = O(K \log N)$ and computational complexity of $T = O(N \log N)$.*

IV. Proposed Construction

The main difference between [2] and our approach is that we replace the left l-regular ensemble of graphs corresponding to the coding matrix $\mathbf{H}$ described in Sec III by left *and* right (l, r)-regular ensemble of graphs.

Definition 4 (Left and right regular graph ensemble)**.** *Let $\mathcal{G}^N_{\text{reg,reg}}\left(R, l, \frac{lN}{R}\right)$ denote the ensemble of left and right regular bipartite graphs with N variable nodes and R check nodes, where each variable node $k \in [N]$ is connected to l check nodes and each check node $j \in [R]$ is connected to $\frac{lN}{R}$ left nodes.*

We know from the modern coding theory that to peel off K unknown variable nodes successfully from the bipartite graph we need ηK number of check nodes for some $\eta > 1$. So we choose the number of check nodes $R = \eta K$. A matrix $\mathbf{H}$ is chosen at random from this ensemble $\mathcal{G}^N_{\text{reg,reg}}\left(\eta K, l, \frac{lN}{\eta K}\right)$ and used as the coding matrix.

In the case of bin detection matrix we now have only $r = \frac{lN}{\eta K} = O\left(\frac{N}{K}\right)$ variable nodes connected to each check node and thus we require only a bin detection matrix $\mathbf{S}$ of $O(\frac{N}{K})$ columns. For the bin detection matrix designs in Sec. III we know from [2] that to differentiate between a zero-ton, single-ton and a multi-ton successfully with probability approaching 1 asymptotically in $\frac{N}{K}$ we only require $\log(\frac{N}{K})$ rows in $\mathbf{S}$. We choose the bin detection matrix similar to the *RandomNoisy, FourierNoisy, BinaryNoisy* designs but with dimensions $P' \times r$ where $P' = O\left(\log(\frac{N}{K})\right)$.

The measurement matrix $\mathbf{A}$ for the proposed construction with $\mathbf{H_{R \times N}} = [\mathbf{h_1}, \mathbf{h_2}, \cdots, \mathbf{h_R}]^{\mathbf{T}}$ and $\mathbf{S_{P' \times r}} = [\mathbf{s_1}, \mathbf{s_2} \cdots \mathbf{s_r}]$ is given by $\mathbf{A}_{RP' \times N} = \mathbf{H} \boxplus \mathbf{S}$, where $\boxplus$ is the new tensoring operation, which is slightly different from the row-tensor operation used in Sec III and is defined as

$$\mathbf{A}_{RP' \times N} = \mathbf{H} \boxplus \mathbf{S} = \begin{bmatrix} \mathbf{h_1} \boxtimes \mathbf{S_1} \\ \mathbf{h_2} \boxtimes \mathbf{S_2} \\ \vdots \\ \mathbf{h_R} \boxtimes \mathbf{S_R} \end{bmatrix}$$

where,

$\mathbf{S_i} = [\mathbf{0}, \cdots, \mathbf{s_1}, \mathbf{0}, \cdots, \mathbf{s_2}, \cdots, \mathbf{0}, \mathbf{s_r}, \cdots, \mathbf{0}]$, $(i \in [R])$, where $\mathbf{0}$ is an all-zero column vector of length P' placed in positions j where $h_{ij} = 0$ and the column vectors $\mathbf{s_j}$,

$j \in [r]$ are placed sequentially in the positions $j$where $h_{ij} = 1$.

The example illustrates the new tensoring operation $\boxplus$.

$$A = H \boxplus S$$

$$H = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix} \quad S = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A = H \boxplus S \; = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

V. Improved Bounds

With our proposed construction of the sensing matrix, Theorem 2 and Theorem 3 can be sharpened to the following new theorems.

Theorem 5 (Sub-linear Time Noisy Recovery)**.** . *In the presence of i.i.d. Gaussian noise with zero mean and variance σ^2, given any K-sparse signal $\mathbf{x}$ with $x_k \in \mathcal{X}$ for $k \in supp(\mathbf{x})$, our noisy recovery schemes achieve a vanishing failure probability $\mathbb{P}_F \to 0$ asymptotically in K and N with*

	M	T
Fourier noisy	$O\left(K \log^{1.3} \frac{N}{K}\right)$	$O\left(K \log^{1.3} \frac{N}{K}\right)$
Binary noisy	$O\left(K \log \frac{N}{K}\right)$	$O\left(K \log \frac{N}{K}\right)$

Theorem 6 (Near-linear Time Noisy Recovery)**.** . *In the presence of i.i.d. Gaussian noise with zero mean and variance σ^2, given any K-sparse signal $\mathbf{x}$ with $x_k \in \mathcal{X}$ for $k \in supp(\mathbf{x})$, the RandomNoisy scheme achieves a vanishing failure probability $\mathbb{P}_F \to 0$ asymptotically in K and N with a measurement complexity of $M = O\left(K \log \frac{N}{K}\right)$ and computational complexity of $T = O\left(N \log \frac{N}{K}\right)$.*

Proof: The bin detection matrix and the decoding methods employed to identify a singleton are identical to that of [2] except that $P = O(\log N)$ is replaced by $P' = O\left(\log(\frac{N}{K})\right)$. Hence the probability of error for the bin detection algorithm can be analyzed exactly as in [2] with P' replaced by P and thus can be shown to be exponentially decaying in P'. Therefore all it remains to be shown is that the $\mathcal{G}^N_{\text{reg,reg}}\left(R, l, \frac{lN}{R}\right)$ ensemble with peeling process fails with a vanishing error probability $\mathbb{P}_F$ asymptotically in K and N. We show this in Theorem. 13 and that completes the proof. ■

By using the left and right regular ensemble we are able to achieve the order optimal bounds for measurement cost. However, there will be a penalty to pay in terms of the minimum SNR required and this is discussed next.

A. Minimum SNR required

We will take the *RandomNoisy* case as an example and let us consider the case when $K = O(N^\delta)$ for some $0 < \delta < 1$. In the derivation of the probability of error in [2] in Proposition 3 and Proposition 4 in Appendix D, it can be seen that the probability of error includes terms of the form

$$e^{-\frac{P}{4}\frac{\gamma^2}{1+4\gamma}} + 2e^{-c_6 P\left(1-\frac{\gamma\sigma^2}{A^2_{\min}}\right)}.$$

It is mentioned that by choosing $P = O(\log N)$, the probability of error can be made to decay as $O\left(\frac{1}{N^3}\right) < O\left(\frac{1}{K^3}\right)$. First, it should be noted that when $P = O(\log N)$, in order for the probability of error to decay as $O\left(\frac{1}{K^3}\right)$, the following two conditions must be satisfied on γ

$$\frac{\gamma^2}{4(1+4\gamma)} > 2\delta, \quad \text{and } c_6\left(1 - \frac{\gamma\sigma^2}{A^2_{\min}}\right) > 2\delta$$

Hence, for the overall probability of error to decay as $O(1/K^3)$, $A^2_{\min}/\sigma^2$ should be large enough such that $0 < 2\delta < \gamma < A^2_{\min}/\sigma^2$. Hence, the theorems appear to be valid only for $A^2_{\min}/\sigma^2 > SNR_{\text{Th}}$.

B. SNR versus Measurements

It should be noted that if the objective is to get the probability of error to decay only as $O(1/K^3)$, it is not required for the probability of error to decay as $O(1/N^3)$ and hence, we could directly choose $P = O(N/K) = N^{1-\delta}$ or in fact, $P = O(N^\beta)$ for any β and that there will be a corresponding SNR_{Th}. Indeed, this appears to reasonable since the assumption that the non-zero values of the signals are from a finite set and hence, when the $\sigma^2 \to 0$, the number of samples required should smoothly go to $O(K)$. Hence, it appears that the order of measurements required should be a function of SNR.

Based on this, indeed, it appears that when using RandomNoisy measurement matrix, it is not even required to switch to the $\mathcal{G}_{\text{reg}-\text{reg}}$ ensemble and we could have simply chosen $P = O(N^\beta)$. However, if we want to choose the Random DFT ensembele (see Page 22 [2]), then we have to switch to the $\mathcal{G}_{\text{reg}-\text{reg}}$. It appears convenient to switch to this ensemble for all results

VI. Proofs

In this section we consider a $\mathcal{G}^N_{\text{reg,reg}}(R, l, \frac{lN}{R})$ ensemble and show that this ensemble with the oracle based peeling decoder fails to recover all the variable nodes with a probability of at most $O\left(\frac{1}{K}\right)$. Although it appears this can be achieved in a straight forward manner by using a capacity achieving spatially-coupled

LDPC ensemble and use the existing results, there are two main obstacles to this:

- In traditional LDPC codes and peeling decoder over binary erasure channel, the input to the decoder is the channel output corresponding to N bit nodes and the check nodes on the right are mere parity checks whose sum modulo 2 is zero. Whereas in our problem the values corresponding to the N variable nodes on the left need to be evaluated by the decoder given the values corresponding to the R check nodes (the real sum of the variable nodes connected) on the right node are non-zero and are given as input to the decoder.

- In traditional LDPC case a constant fraction ϵ of these N variable nodes are erased by the channel and usually the emphasis is on analyzing the performance of peeling decoder asymptotically in N or R when rate=$1-\frac{R}{N}$ is fixed. But in our case the fraction of the nodes erased $= 1-\frac{K}{N}$, where K is usually of the form $K = N^{\delta}$, tend to one and the rate of the code$= 1-\frac{R}{N} = 1-\frac{\eta N^{\delta}}{N}$ tend to one asymptotically in N.

Consider a left and right regular LDPC code $\mathcal{G}_{\text{LDPC}}(N,l,r)$ where N is the number of variable nodes on the left and l,r are the regular left and right degrees respectively. Let $\mathrm{P}_{\text{BEC}}^{(i)}(\mathbf{y})$ be the degree distribution of the number of check nodes after iteration i of peeling decoder given $\mathbf{y}$ is the channel output. And similarly $\mathcal{G}^{N}_{\text{reg,reg}}(R,l,\frac{lN}{R})$ be the graph corresponding to the parity check matrix in the support recovery problem and $\mathrm{P}_{\text{SR}}^{(i)}(\mathbf{z})$ be the degree distribution of the check nodes on the right after iteration i of the oracle-based peeling decoder, given $\mathbf{z}$ is the support recovery equivalent of syndrome corresponding to $\mathbf{x}$ i.e., $\mathbf{z} = \mathbf{H}\mathbf{x}$ where the operations are over the real field. Note that $\mathbf{y}$ is a vector of dimension N whereas $\mathbf{z}$ is of dimension $\frac{Nl}{r}$.

Note that in the peeling decoder we consider, we peel off one degree-1 check node and the variable node connected to it from the graph in each iteration. In the case of LDPC-BEC we remove all the variable nodes that are not erased by the channel and the resulting graph is input to the decoder. Similarly in the case of support recovery problem we consider the oracle based peeling decoder in [2] and we analyze the *pruned*-graph where we remove all the zero nodes from the original graph.

Lemma 7 (Equivalence to LDPC-BEC)**.** *Whenever* $\mathbf{y}$ *and* $\mathbf{z}$ *satisfy*

$\mathbf{z} = \mathbf{H}\mathbf{x}$ *such that* $S := |supp(\mathbf{x})| = |\{i : y_i = \mathcal{E}\}|$

where $\mathcal{E}$ *denotes erasure, then* $P_{BEC}^{(i)}(\mathbf{y}) = P_{SR}^{(i)}(\mathbf{z}) \quad \forall i.$

Proof: Define $S^c = [1:N]\backslash S$. In the case of LDPC codes on BEC we peel off all non-erased variable nodes corresponding to S^c and input the resulting graph to the peeling decoder. Similarly in the case of bipartite graph in support recovery problem we peel off all the zero nodes corresponding to S^c and we input the resulting graph to oracle based peeling decoder. From this point onward the peeling decoders are identical and thus we have our result. ■

Thus by considering a BEC of erasure probability $\epsilon = \frac{K}{N}$ we can equivalently consider peeling decoder of LDPC codes on BEC channel and use various existing results.

Lemma 8. *The evolution of the left and right degree distribution as the peeling decoder progresses can be given by*

$$\begin{aligned}
\tilde{L}_l(y) &= y^k l,\\
\tilde{R}_1(y) &= r\epsilon y^{l-1}[y-1+(1-\epsilon y^{l-1})^{r-1}]\\
\tilde{R}_i(y) &= \binom{r}{i}(\epsilon y^{l-1})^i(1-\epsilon y^{l-1})^{r-1}, \quad i \geq 2
\end{aligned}$$

where $\epsilon = \frac{K}{N}$ *and* $r = \frac{lN}{\eta K}$*. Note that the curve corresponding to* $\tilde{L}_i(y)(\tilde{R}_i(y))$ *for* $y \in [0,1]$ *gives the expected number of degree* i *variable nodes (check nodes) normalized with respect to* K (ηK)*.*

Proof: As we showed in Lemma. 7 the peeling decoder for an LDPC on BEC channel and oracle based peeling decoder for CS are identical upto the residual degree distributions at each iteration. Hence we can use the result for LDPC codes [4] with equivalent channel erasure probability $\epsilon = \frac{K}{N}$. ■

Definition 9 (BP Threshold)**.** *We define the BP threshold,* η^{BP} *to be the minimum value of* η *for which there is no non-zero solution for the equation:*

$$\begin{aligned}
y &= \lim_{N\to\infty} 1-\left(1-\frac{y^{l-1}}{N^{1-\delta}}\right)^{\frac{dN^{1-\delta}}{\eta}}\\
&= 1-e^{\frac{-ly^{l-1}}{\eta}}
\end{aligned}$$

in the range $y \in [0,1]$.

Lemma 10. *[4] If* $\eta > \eta^{BP}$ *then with probability at least* $1-O\left(K^{1/6}e^{-\frac{\sqrt{Kl}}{(lr)^3}}\right)$ *the peeling decoder of a specific instance progresses until the number of residual variable nodes in the graph has reached size* γK *where* γ *is an arbitrary positive constant.*

Definition 11 (Expander Graphs)**.** *A bipartite graph with K left nodes and regular left degree* l *is called a* $(\gamma, 1/2)-$ *expander if for all subsets* S *of left nodes with* $|S| \leq \gamma K$*, the right neighborhood of* S *denoted by* $\mathcal{N}(S)$ *satisfies* $|\mathcal{N}(S)| > l|S|/2$*.*

Lemma 12. *Consider a left and right regular ensemble $\mathcal{G}^N_{reg,reg}(\eta K, l, \frac{Nl}{\eta K})$, then the pruned graph resulting from any given K-sparse signal $\mathbf{x}$ is a $(\gamma, 1/2)$-expander with probability at least $1 - O\left(\frac{1}{K^{l-2}}\right)$ for a sufficiently small constant $\gamma > 0$.*

Proof: The proof is similar to the proof used in [2] with minor modifications. Let E_v denote the event that a subset S_v of variable nodes on the left with size v has at most $l|S_v|/2$ neighbors whose probability can be computed as

$$\Pr(E_v) \leq \binom{K}{v}\binom{\eta K}{lv/2}\left(\frac{vl}{2\eta K}\right)^{lv} \quad (1)$$

$$\leq c^{vl/2}\left(\frac{v}{K}\right)^{v(l/2-1)} \quad (2)$$

where $c = \frac{le^2}{2\eta}$ is a constant. In (1) we upper bound the probability of E_v via union bound over all possible size v subsets on the left and size $lv/2$ subsets on the right. In (2) we use the inequality $\binom{a}{b} \leq (ae/b)^b$ and we assume $l \geq 2$ to simplify the constant factor. Then we union bound over all subsets of size upto the remaining nodes $\gamma^* K$ where we choose $\gamma^* = \left(4c^l\right)^{\frac{-1}{l-2}}$

$$\sum_{v=2}^{\gamma^* K} \Pr(E_v) \leq \sum_{v=2}^{\gamma^* K}\left(c^l\left(\frac{v}{K}\right)^{l-2}\right)^{v/2}$$

$$= O\left(\frac{1}{K^{l-2}}\right)$$

Thus we showed that asymptotically in K, the left and right regular graphs are good expander graphs with probability atleast $1 - O(1/K^{l-2})$. ∎

Theorem 13. *Consider the ensemble $\mathcal{G}^N_{reg\text{-}reg}(\eta K, l, \frac{Nl}{\eta K})$, the oracle based peeling decoder peels off all the variable nodes in the pruned graph in ηK iterations with probability at least $1 - O\left(1/K^{l-2}\right)$.*

Proof: Lemma 10 shows us that the peeling decoder fails to peel off till the residual graph has γN variable nodes remaining with an exponentially low probability. Then in Lemma 12 we show that the left regular graphs are good expanders with a probability of atleast $1 - O(1/K^{l-2})$ and hence the remaining γN nodes can be peeled off with high probability. Thus the overall probability of failure will be dominated by small stopping sets which can be upper bounded by $O(1/K^{l-2})$. ∎

VII. Numerical Results

In this section we provide the empirical performance of our scheme in the noisy setting. We fix the parameters $K = 50$ and $N = 10^5$. For each data point we generate a K-sparse signal at random and keep it fixed for all the simulations. Specifically, supp($\mathbf{x}$) is chosen uniformly at random from $[N]$ and the non-zero values in $\mathbf{x}$ are chosen uniformly at random from the set $\{+1, -1\}$. We sample the coding matrix $\mathbf{H}$ from the ensemble $\mathcal{G}^N_{\text{reg,reg}}(R = 2K, l = 4, r = \frac{2N}{K})$ for each simulation. For the bin detection matrix $\mathbf{S}$ we use a QAM-modulated codebook of $(12, n)$ truncated convolutional codes corresponding to rates $\frac{1}{2}, \frac{1}{4}$ and $\frac{1}{8}$ where $n = 24, 48$ and 96 respectively. Note that we pad an extra row of all ones for determining the sign of x_i thus resulting in the bin detection matrix dimensions of $13 \times r, 25 \times r$ and $49 \times r$ where $r = 4000$ is the right degree of the graph corresponding to $\mathbf{H}$. We use a Viterbi decoder for singleton identification.

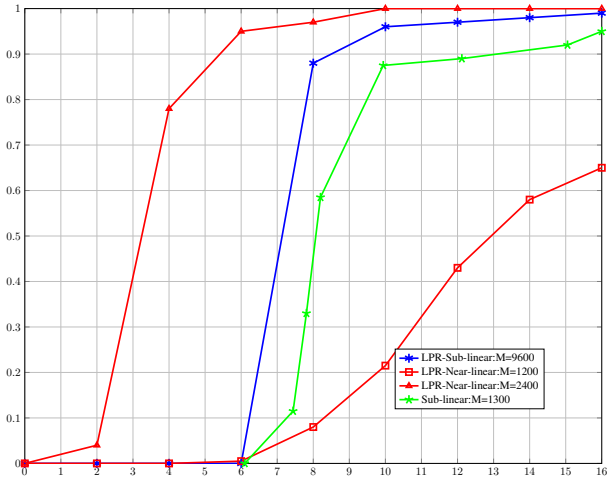

Fig. 1. Required E_b/N_0 for a target FER of 10^{-4}, Rate $\frac{1}{2}$.

VIII. Extensions

For the group testing problem addressed in [5] also, it appears we can replace the ensemble used by Lee, Pedarsani and Ramchandran by the regular-regular ensemble and obtain a sharper result, i.e., $O(K\log(N/K))$ measurements for the case of $K = N^\delta$.

We use a left and right regular ensemble for the LDPC code and we use a measurement matrix with $2\beta\log_2(N/K)$ rows and N/K columns. When the right regular ensemble is considered, then the probability of error at each measurement node will be $O\left(\frac{K^{\beta-1}}{N^{\beta-1}}\right)$ and there are K stages in the peeling process. Thus, for the overall error probability to decay with K according to $O\left(\frac{K^\beta}{N^{\beta-1}}\right) = O\left(N^{\delta\beta-\beta+1}\right)$. If we want this probability of error to be $O(N^{-\delta})$, then we can set $\beta = \frac{1+\delta}{1-\delta}$ to satisfy this. Then, the total number of measurements will be $2(1+\delta)K\log_2 N$. This is strictly an improvement over $6K\log_2 N$ measurements used in [5].

The main idea in the proof is to look at the reduced subgraph induced only by the defective items and when

$N/K \to \infty$, this graph will be a graph from the left regular, right Poisson ensemble considered in [5].

REFERENCES

[1] X. Li, S. Pawar, and K. Ramchandran, "Sub-linear time compressed sensing using sparse-graph codes," in *Proc. IEEE Symposium on Information Theory*, 2015, pp. 1645–1649.

[2] ——, "Sub-linear time compressed sensing using sparse-graph codes," http://www.eecs.berkeley.edu/~xiaoli/TR_CS_sublinear.pdf.

[3] M. J. Wainwright, "Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting," *Information Theory, IEEE Transactions on*, vol. 55, no. 12, pp. 5728–5741, 2009.

[4] T. Richardson and R. Urbanke, *Modern coding theory*. Cambridge University Press, 2008.

[5] K. Lee, R. Pedarsani, and K. Ramchandran, "Saffron: A fast, efficient, and robust framework for group testing based on sparse-graph codes," *arXiv preprint arXiv:1508.04485*, 2015.