Sub-linear Time Compressed Sensing for Support Recovery using Left and Right regular Sparse-Graph Codes

Avinash Vem, Nagaraj Thenkarai Janakiraman, Krishna Narayanan
Department of Electrical & Computer Engineering
Texas A&M University, College Station, TX, 77843
{vemavinash,tjnagaraj,krn}@tamu.edu

Abstract—In [1], [2], two schemes have been proposed to recover the support of a K-sparse N-dimensional signal from noisy linear measurements. Both schemes use left-regular sparse-graph code based sensing matrices and a simple peeling-based decoding algorithm. Both the schemes require $O(K \log N)$ measurements and the first scheme require $O(N \log N)$ computations whereas the second scheme requires $O(K \log N)$ computations (sub-linear time complexity when K is sub-linear in N). We show that by replacing the left-regular ensemble with left and right regular ensemble, we can reduce the number of measurements required of these schemes to the optimal order of $O\left(K \log \frac{N}{K}\right)$ with decoding complexities of $O(K \log \frac{N}{K})$ and $O(N \log \frac{N}{K})$, respectively.

I. INTRODUCTION

In [1] (and in the expanded version in [2]), Li, Pawar and Ramchandran have considered the compressed sensing problem of recovering the support of a K-sparse, N-dimensional signal from M linear and noisy measurements. Based on sparse-graph codes with a *regular left degree* and a peeling decoder, they have proposed an elegant design of the measurement matrix and a recovery algorithm. 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 $M = O(K \log N)$ measurements with a sub-linear ($O(K \log N)$) decoding complexity.

In this paper, we show that the bounds on the measurement complexity reported in [1], [2] can be improved by considering *left and right regular* sparse-graph based sensing matrices. With sub-linear time decoding complexity, we show that only $O\left(K \log \frac{N}{K}\right)$ measurements are required when $K = O(N^\delta)$, for any $0 \leq \delta < 1$. This matches the information-theoretic lower bound on the number of measurement required for asymptotically-reliable recovery [3]. Simulation results show improved performance over those in [2].

The literature on compressed sensing is vast and it is difficult to provide a comparison with several of the existing results in the literature due to page restriction

of this conference submission. Nevertheless, it should be pointed out that the use of left and right regular sensing matrices for compressed sensing has been proposed in [4] and the measurement complexity has been shown to be only $O\left(K \log \frac{N}{K}\right)$. However, the scheme and the decoding algorithm in this paper are different from those in [4] and are based directly on those in [1], [2]. In particular, unlike in [4], the sensing matrix in this paper is constructed based on a tensor-product construction and the decoding algorithm is based on identifying singletons and peeling them off.

II. SYSTEM MODEL

The classical problem of compressed sensing involves 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, let

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

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 known to be 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 the noiseless setting. It is well-known in *compressive sensing* that if $K \ll N$ 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 $\mathrm{supp}(\mathbf{x}) := |\{i : x_i \neq 0, i \in [N]\}|$ 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 probability of failure of support recovery which can be defined as

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

For the information-theoretic lower bound in [3] to hold, the non-zero elements in $\mathbf{x}$ should have a sufficiently large minimum absolute value. In view of this condition, similar to [1], [2], 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 $\mathcal{A} := \{A_{\min} + \rho l\}_{l=0}^{L_1}$, $\Theta := \{2\pi l / L_2\}_{l=0}^{L_2}$ for finite but arbitrarily large integers L_1 and L_2.

III. Prior Work

In this section, we review the construction of the measurement matrix $\mathbf{A}$ proposed by Li, Pawar and Ramchandran in [1] and [2], and also summarize 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 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 coding 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 $q_i, i \in [R]$, be the number of non-zero variable nodes connected to i^{th} check node. Assume an "oracle" that solves the 1-sparse problem by examining each check node observations and classifies it as a zero-ton($q_i = 0$), singleton($q_i = 1$) 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 single-tons. The decoding process continues until there are no more singletons. The decoding is successful if all the K non-zero elements of $\mathbf{x}$ are 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.

For the noisy setting, they have proposed three designs for $\mathbf{S}$ which essentially performs the role of the oracle in identifying a single-ton at each check node. These designs are *RandomNoisy* with near-linear decoding complexity, *BinaryNoisy* and *FourierNoisy* each with sublinear decoding complexity. The bin-detection matrix $\mathbf{S}$ for the three settings are 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:* $\mathbf{S} = [\mathbf{S_0} \mathbf{S_1} \cdots \mathbf{S_{P-1}}]^{\mathbf{T}}$, 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:* $\mathbf{S} = f(\mathbf{C})$ where $\mathbf{C}_{P \times N}$ is a binary codebook(or subset of a codebook) of a linear code with block length P, $f : \{0,1\}^q \to \mathcal{M}$ is a modulation scheme that maps $C_{P \times N}$ to $S_{\frac{P}{q} \times N}$. For e.g., for QAM $q = 2$ and $\mathcal{M} = \{\pm 1 \pm i\}$.

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

Theorem 1 ([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	$O(K \log N)$	$O(K \log N)$

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

Theorem 2 ([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 3 (Left and right regular graph ensemble). *Let $\mathcal{G}_{\text{reg,reg}}^N \left(R, l, \frac{lN}{R} \right)$ 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.*

In the design considerations 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 $\mathbf{S}$ with $O(\frac{N}{K})$ columns. For the bin detection matrix designs in Sec. III we know from [2] that to differentiate between a zero-ton, singleton 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 to be similar to the *RandomNoisy, FourierNoisy, BinaryNoisy* designs but with dimensions $P' \times r$ where $P' = O\left(\log(\frac{N}{K})\right)$.

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 $\mathcal{G}_{\text{reg,reg}}^N \left(\eta K, l, \frac{lN}{\eta K} \right)$ and used as the coding matrix.

The measurement matrix $\mathbf{A}$ for the proposed construction with $\mathbf{H_{R \times N}} = [\mathbf{h_1}, \mathbf{h_2}, \cdots, \mathbf{h_N}]^{\mathbf{T}}$ 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 $\mathbf{s_k}$, $k \in [r]$ are placed sequentially in the positions j where $h_{ij} = 1$. We illustrate the new tensoring operation $\boxplus$ via Example 4.

Example 4. *Let*

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

denote an adjacency matrix from the ensemble $\mathcal{G}^6_{\mathrm{reg,reg}}(4,2,3)$. We choose $P' = \lceil \log_2 r \rceil = 2$ and let $S_{P' \times r}$ be defined as

$$S = \begin{bmatrix} +1 & -1 & -1 \\ -1 & +1 & -1 \end{bmatrix}$$

Then, the measurement matrix $\mathbf{A}$ with $M = P'R = 8$ measurements is given by

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

V. IMPROVED BOUNDS

With our proposed construction of the measurement matrix, Theorem 1 and Theorem 2 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'. For this particular choice of P' the probability of error for the bin detection part vanishes asymptotically in $\frac{N}{K}$. Therefore for K sub-linear in N all it remains to be shown is that the $\mathcal{G}^N_{\mathrm{reg,reg}}\left(R, l, \frac{lN}{R}\right)$ ensemble with peeling process fails with a vanishing error probability $\mathbb{P}_F$ asymptotically in K and N. For choice of $l \geq 3$, Theorem 13 gives us this required result and that completes the proof. ∎

VI. PROOFS

In this section we consider a $\mathcal{G}^N_{\mathrm{reg,reg}}(R, l, \frac{lN}{R})$ 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 directly 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 variable (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) are non-zero and form 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, sublinear in N, 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 $\mathcal{G}_{\mathrm{LDPC}}(N, l, r)$ where N is the number of variable nodes

on the left and l, r are the regular left and right degrees respectively. Let $P_{BEC}^{(i)}(\mathbf{y})$ 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 $\mathcal{G}_{reg,reg}^N(R, l, \frac{lN}{R})$ be the graph corresponding to the parity check matrix in the support recovery problem and $P_{SR}^{(i)}(\mathbf{z})$ 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., $\mathbf{z} = \mathbf{Hx}$ where the operations are over the real field.

Note that in the peeling decoder, we peel off one degree-1 check node and the variable node connected to it from the graph in each iteration. In the LDPC-BEC problem 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 variable nodes from the original graph and input to the decoder.

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

$\mathbf{z} = \mathbf{Hx}$ *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*

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

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 [5, Theorem 3.107] 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:*

$$y = \lim_{\frac{N}{K} \to \infty} 1 - \left(1 - \frac{Ky^{l-1}}{N}\right)^{\frac{lN}{\eta K}}$$
$$= 1 - e^{\frac{-ly^{l-1}}{\eta}}$$

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

Lemma 10. *[5, Theorem 3.107] 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 $\mathcal{G}_{reg,reg}^N(\eta K, l, \frac{Nl}{\eta K})$, 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 $\mathcal{G}_{reg\text{-}reg}^N(\eta K, l, \frac{Nl}{\eta K})$, 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 a given SNR we generate a K-sparse signal at random and perform the support recovery for this signal over 200 sensing matrices sampled from the proposed construction. Specifically, $\text{supp}(\mathbf{x})$ 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 $\mathcal{G}_{\text{reg,reg}}^N(R = 2K, l = 4, r = \frac{2N}{K})$ for each simulation. For the bin detection matrix we consider the *BinaryNoisy* scheme and we use two classes of codes: convolutional codes and (12,24) Golay code with QAM modulation. In the case of convolutional codes we consider $(12, n)$ truncated convolutional code corresponding to rates $\frac{1}{2}, \frac{1}{4}$ and $\frac{1}{8}$ with a constraint length of 8 which results in $n = 24, 48$ and 96 respectively. This gives bin detection matrix dimensions of $12 \times r, 24 \times r$ and $48 \times r$ where $r = 4000$ is the right degree of the graph corresponding to $\mathbf{H}$. For the singleton identification Viterbi soft decision decoding is considered for convolutional codes whereas a hard decision syndrome decoding is considered for Golay code resulting in a decoding complexity of $O\left(K \log\left(\frac{N}{K}\right)\right)$. We observe from Fig. 1 that the Golay code based construction with
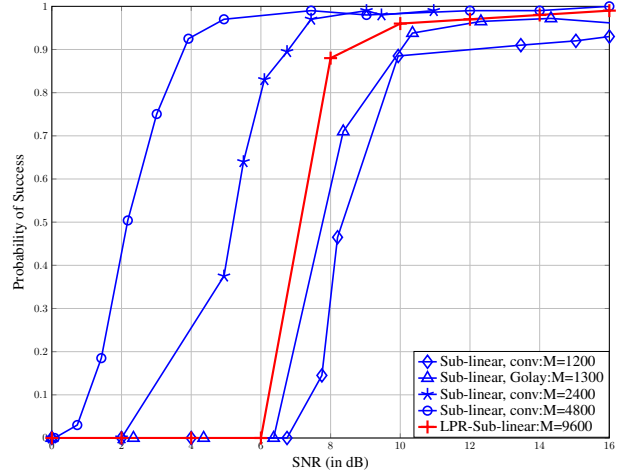

Fig. 1. Probability of Success for our construction (blue curves) with *BinaryNoisy* scheme using convolutional codes(conv) and Golay code with sub-linear time decoding complexity of $O(K \log \frac{N}{K})$.And we compare the performance with that of *BinaryNoisy* scheme by Li, Pawar and Ramachandran (LPR) (red curve) [2] with sub-linear decoding complexity of $O\left(K \log N\right)$.

$M = 1300$ has similar performance and the convolutional code based construction with $M = 2400$ has better performance when compared to that of $M = 9600$ *BinaryNoisy* scheme with sub-linear time complexity decoder of LPR[1].

VIII. Conclusion

In this paper we address the support recovery problem in compressed sensing. We use left *and* right regular sparse-graph ensemble and propose a new construction for the measurement matrix based on that of Li et al., [1]. We show that this construction achieves order optimal measurement cost and achieves asymptotically vanishing error probability under sub-linear time complexity.

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] S. Jafarpour, W. Xu, B. Hassibi, and R. Calderbank, "Efficient and robust compressed sensing using optimized expander graphs," *Proc. IEEE Symposium on Information Theory*, vol. 55, no. 9, pp. 4299–4308, 2009.

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