Efficient Compressive Sensing with Deterministic Guarantees Using Expander Graphs

Weiyu Xu and Babak Hassibi
EE Department
California Institute of Technology
Pasadena, CA 91125
weiyu,hassibi@caltech.edu

Abstract— Compressive sensing is an emerging technology which can recover a sparse signal vector of dimension n via a much smaller number of measurements than n. However, the existing compressive sensing methods may still suffer from relatively high recovery complexity, such as $O(n^3)$, or can only work efficiently when the signal is super sparse, sometimes without deterministic performance guarantees. In this paper, we propose a compressive sensing scheme with deterministic performance guarantees using expander-graphs-based measurement matrices and show that the signal recovery can be achieved with complexity $O(n)$ even if the number of nonzero elements k grows linearly with n. We also investigate compressive sensing for approximately sparse signals using this new method. Moreover, explicit constructions of the considered expander graphs exist. Simulation results are given to show the performance and complexity of the new method.

I. INTRODUCTION

Compressive sensing has recently received a great amount of attention in the applied mathematics and signal processing community. The theory of compressive sensing, as developed over the past few years, attempts to perform sampling and compression simultaneously, thus significantly reducing the sampling rate. What allows this theory is the fact that, in many applications, signals of interest have a "sparse" representation over an appropriate basis. In fact, compressive sampling is intimately related to solving underdetermined systems of linear equations with sparseness constraints. The work of Candes, Romberg and Tao [1], [2] and Donoho [4] came as a major breakthrough in that they rigorously demonstrated, for the first time, that, under some very reasonable assumptions, the solution could be found using simple linear programming—thus rendering the solution practically feasible. The method is essentially constrained l_1 minimization, which for a long time was empirically known to perform well for finding sparse solutions and has been known in the literature as "basis pursuit" [5], [7]. Interestingly, the area of compressive sensing is closely connected to the related areas of coding [6], high-dimensional geometry [14], sparse approximation theory [15], data streaming algorithms [16], [17] and random sampling [20]. Furthermore, promising applications of compressive sensing are emerging in compressive imaging, medical imaging, sensor networks and analog-to-digital conversion [3].

While solving the linear program resulting from l_1 optimization can be done in polynomial-time (often $O(n^3)$, where n is the number of unknowns), this may

still be infeasible in applications where n is quite large (e.g., in current digital cameras the number of pixels is of the order $n = 10^6$ or more) [8]. Therefore there is a need for methods and algorithms that are more computationally efficient. Also, in many of the previous works, random measurement matrices are used where a successful signal recovery can not be always guaranteed although it succeeds with a high probability. So it is also desirable to have an explicit construction of measurement matrix for compressive sensing [3], [11].

Recently, some significant progress has been made in addressing these two problems for compressing sensing. Orthogonal Matching Pursuit (OMP) algorithms can be used as alternative recovery algrithms which require $O(nk^2)$ computations [9], where k is the number of non-zero entries in the unknown vector; however, this may also be too high a complexity. Stage-wise OMP [25] has recently been proposed that solves the problem in $O(n \log n)$ computations. In [16] a certain sparse coefficient matrix has been used, along with group testing, that yields an algorithm with $O(k \log^2 n)$ complexity; however, this comes at the expense of more measurements—$O(k \log^2 n)$ measurements, as opposed to the $O(k \log n)$ measurements required of the aforementioned methods. Chaining pursuit has been introduced in [17], which has complexity $O(k \log^2 n \log^2 k)$ and also requires $O(k \log^2 n)$ measurements. From the number of measurements needed, we can see that both the group testing methods [16] and the chaining pursuit methods only work in the "supersparse" case, i.e., when the ratio k/n is very small—when k/n is increased, an enormous number of measurements is required, as noted in [19]. Motivated by low-density parity-check codes (LDPCs) a method called sudocodes has been proposed in [18] to recover sparse signal with high probability, which requires $O(k \log n \log k)$ recovery complexity, yet only $O(k \log n)$ measurements. The Homotopy methods are able to recover the sparse solutions by reducing the computational complexity from $O(n^3)$ to $O(nk^2)$ [10]. In [21], it was shown that by using the Vandermonde measurement matrix and linear programming, one can recover k nonzero elements using approximately $2k$ measurements when the nonzero elements are restricted to positive numbers. In [22], motivated by Reed-Solomon codes rather than LDPC codes as in [18], a

scheme of recovery complexity $O(k^2)$ is proposed to recover any signal vector with k nonzero elements using the Vandermonde measurement matrix.

With the exception of the method in [11], the group testing methods in [16] and the Vandermonde measurement matrix based methods in [21], [22], all the results described above hold with "high probability" either over the random measurement matrix or over some assumptions on the input signals [18]. While the methods in [11], [16] can guarantee sparse signal recovery deterministically with explicit measurement matrices, they suffer from the fact that they only work in the supersparse case where k can not be kept as a constant fraction of n. But recovering a constant fraction of n non-zero elements via a small number of measurements is of great practical interests [6]. For this reason, in this paper, we will allow k to grow linearly in n, i.e., $k = \Theta(n)$. In this sparsity regime, the complexity of the methods of [21], [22] are of order $O(n^3)$ and $O(n^2)$ respectively, which will still be unpractical for problems of large dimensions. Sometimes, it is also required that the recovery schemes are applicable to approximately sparse signals and robust to the noise in the measurements and numerical errors.

In this paper, we propose a new scheme for compressive sensing with deterministic performance guarantees based on bipartite expander graphs and show that even with $k = \Theta(n)$, the recovery complexity of our algorithm is $O(n)$ while saving a constant fraction of n measurements. The new scheme thus pushes the performance and complexity bounds of compressive sensing to be asymptotically linear in n at the same time. To the best of our knowledge, this has not been achieved by other methods. Bipartite expander graphs [26], [27] are a certain class of graphs whose existence has been known for quite some time and whose recent explicit construction is considered to be a major feat [26]. In some sense our approach is closest to that of [18], which is inspired by LDPCs, certain classes of which are related to expander graphs [27], but here we provide performance guarantees. Preliminary analytic results further show the feasibility of application of the new method to approximately sparse signals and the noisy measurement cases.

The rest of the paper is organized as follows. In the next section we review the background and give the problem formulation. We introduce expander graphs in Section III and show how they can be used to develop deterministic methods with $O(n)$ recovery complexity. The analysis of the new compressive sensing scheme for approximately sparse signals is given in Section IV. Simulation results are given in the section V.

II. BACKGROUND AND PROBLEM FORMULATION

In compressive sensing the starting point is an n-dimensional signal vector which admits a sparse representation in some particular basis. Since the basis is not of primary concern to us, we may, without loss of generality, assume that it is the standard basis. In other words, we shall assume that we have an n-dimensional vector $x \in \mathcal{R}^n$, such that no more than k entries are

non-zero. Clearly, $k < n$. Here we assume k can be up to a constant fraction of n, since this case is of great practical interest [6].

The vector x itself is not directly observable. What is observable are *measurements* of x that correspond to linear combinations of the form

$$\sum_{j=1}^{n} a_j x_j. \tag{1}$$

We often have control over what measurements to employ, and this may turn out to help us. In any event, assuming we have m ($k < m < n$) measurements of this form, we may collect them in a $m \times n$ matrix A so that

$$y = Ax, \tag{2}$$

or, in other words,

$$y_i = \sum_{j=1}^{n} A_{ij} x_j, \quad i = 1, \ldots, m. \tag{3}$$

The system of equations (2) is, of course, under-determined. However, the fact that a sparse solution exists, allows us to be able to find the solution. It was a significant result when it was rigorously shown by Candes, Romberg and Tao [1], [2] and Donoho [4] that, under the sparsity assumption, the solution could be found via solving the l_1 optimization problem

$$\min_{x, Ax=y} \|x\|_1, \tag{4}$$

where $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ is the l_1-norm of the vector x. The upshot is that something that appeared to be practically infeasible can now be potentially computed. For example, in [6], it was shown that if the measurement matrix A satisfies the restricted isometry conditions, then the l_1 minimization can recover a vector with up to k nonzero elements, where k is a constant fraction of n.

In spite of the recent developments, one unanswered question is whether we can develop compressive sensing schemes and recovery algorithms with complexity $O(n)$ even when $k = \Theta(n)$? If yes, can one *explicitly* develop constructions of measurement matrices that *deterministically* guarantee finding the optimal solution for all signal instances in such schemes, provided the vector x is sparse enough? We shall presently answer both questions in the affirmative and discuss all these developments in the next section.

III. EXPANDER GRAPHS AND EFFICIENT ALGORITHMS

A. Expander Graphs

Expander graphs can be defined for arbitrary graphs, however, here we shall restrict ourselves to *bipartite* graphs. For a bipartite graph, we have two types of nodes. Following coding theory parlance, we will call one type left *variable nodes* of which there are n and which correspond to the entries of x, and right *parity check nodes* of which there are m and which correspond to the entries of y (or the measurements). We assume that $n \geq m$. In a bipartite graph connections within the variable nodes and within the parity check nodes are not

allowed. The existence of edges between the different variables and parity check nodes are represented by a $m \times n$ matrix A. In particular,

$$A_{ij} = \begin{cases} 1 & \text{if right node } i \text{ connected to left node } j \\ 0 & \text{otherwise} \end{cases}$$
$$(5)$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. In what follows we shall use the matrix thus obtained from a suitably chosen bipartite graph as the measurement matrix for compressive sampling.
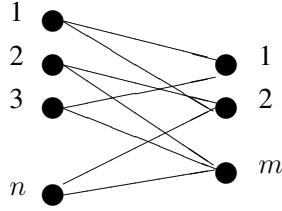

Fig. 1. A bipartite graph.

A bipartite graph will be said to have *regular left degree* c if the number of edges emanating from each variable node is c.

Definition 1 (Expander). *A bipartite graph with n variable nodes, m parity check nodes and regular left degree c will be called a $(\alpha n, \beta c)$ expander, for some $0 < \alpha, \beta < 1$, if for every subset of variable nodes $\mathcal{V}$ with cardinality less than or equal to αn, i.e., $|\mathcal{V}| \le \alpha n$, the number of neighbors connected to $\mathcal{V}$ is larger than $\beta c |\mathcal{V}|$, i.e., $|\mathcal{N}(\mathcal{V})| > \beta c |\mathcal{V}|$, where $\mathcal{N}(V)$ is the set of neighbors of $\mathcal{V}$.*

Here we assume that each righthand side node also has a regular degree d, where $cn = md$. The existence of expander graphs has been known for quite some time since the work of Pinsker and Bassylago [28], who used probabilistic arguments to prove their existence. Expander graphs arise in many applications: fast, distributed routing algorithms [29], LDPC codes [27], storage schemes [30], and proof systems [31], to name a few. An explicit construction of constant regular left degree *lossless* (with β arbitrarily close to 1) expander graph is recently given in [26]. An existence result, which holds for the setting we are interested in, is the following [23]:

Theorem 1. *Let $0 < \beta < 1$ and the ratio $r = \frac{m}{n}$ be given. Then for large enough n there exists a regular left degree c and a regular right degree d bipartite expander $(\alpha n, \beta c)$ for some $0 < \alpha < 1$ and some constant (not growing with n) c.*

B. The Main Algorithm

We are now in a position to describe our main algorithm. We begin with $\beta = \frac{3}{4}$ and some fixed $r = \frac{m}{n}$. (Thus, our number of measurements is $m = nr$. We can use the construction of [26], or any other recent one, to construct an expander with some $0 < \alpha < 1$ and constant c.) Denote the resulting measurement matrix by A. In particular, assuming $x \in \mathcal{R}^n$ is sparse with at most

k nonzero entries, we perform the m measurements

$$y = Ax. \tag{6}$$

We will assume that

$$k \le \frac{\alpha n}{2}. \tag{7}$$

We need one further notation: given an estimate $\hat{x}$ of x, we define as the *gap* in the i-th equation the quantity

$$g_i = y_i - \sum_{j=1}^{n} A_{ij} \hat{x}_j. \tag{8}$$

Algorithm 1.   1) *Start with $\hat{x} = 0_{n \times 1}$.*
   2) *If $y = A\hat{x}$, declare $\hat{x}$ the solution and exit.*
      Else, find a variable node, say $\hat{x}_j$, such that of the c measurement equations it participates in $c' > \frac{c}{2}$ of them have an identical nonzero *gap g.*
   3) *Set $\hat{x}_j = \hat{x}_j + g$. Go to 2.*

Algorithm 1 is incredibly simple. What is remarkable about it is that, in step 2 of the algorithm, if $y \neq A\hat{x}$ one can always find a variable node with the property that $c' > \frac{c}{2}$ among the measurement equations it participates in has *identical* nonzero gap g. Furthermore, the algorithm terminates in at most ck steps. We proceed to establish these two claims via a series of Lemmas. At any step of the algorithm, let $\mathcal{S}$ denote the set

$$\mathcal{S} = \{j | \hat{x}_j \neq x_j\}. \tag{9}$$

Lemma 1 (Initialization). *When $\hat{x} = 0$,$y \neq A\hat{x}$ and $k \le \frac{\alpha n}{2}$, there always exists a variable node such that $c' > \frac{c}{2}$ of the measurement equations it participates in has identical nonzero gap g.*

Proof: Initially since $\hat{x}_i = 0$, the set $\mathcal{S}$ has cardinality $|\mathcal{S}| = k \le \alpha n/2$. We can therefore apply the property of the expander with $\beta = \frac{3}{4}$ to $\mathcal{S}$ to conclude that

$$|\mathcal{N}(\mathcal{S})| > \frac{3}{4} c |\mathcal{S}|. \tag{10}$$

Let us now divide the set $\mathcal{N}(\mathcal{S})$ into two disjoint sets: $\mathcal{N}_{unique}(\mathcal{S})$ comprised of those elements of $\mathcal{N}(\mathcal{S})$ that are connected to only one edge emanating from $\mathcal{S}$ and $\mathcal{N}_{>1}(\mathcal{S})$ which are the remaining elements of $\mathcal{N}(\mathcal{S})$ that are connected to more than one edges emanating from $\mathcal{S}$. Clearly, (10) implies

$$|\mathcal{N}_{unique}(\mathcal{S})| + |\mathcal{N}_{>1}(\mathcal{S})| > \frac{3}{4} c |\mathcal{S}|. \tag{11}$$

Counting the edges emanating from $\mathcal{S}$ leads to

$$|\mathcal{N}_{unique}(\mathcal{S})| + 2 |\mathcal{N}_{>1}(\mathcal{S})| \le c |\mathcal{S}|, \tag{12}$$

since the total number of edges is $c|\mathcal{S}|$ and since some of the nodes in $\mathcal{N}_{>1}(\mathcal{S})$ may have more than 2 edges connecting to $\mathcal{S}$. Eliminating $|\mathcal{N}_{>1}(\mathcal{S})|$ from the inequalities (11) and (12) yields

$$|\mathcal{N}_{unique}(\mathcal{S})| > \frac{c}{2} |\mathcal{S}|. \tag{13}$$

The above inequality implies that there must be at least one element of $\mathcal{S}$ that is connected to $c' > \frac{c}{2}$ elements of $\mathcal{N}_{unique}(\mathcal{S})$. But since this is the only element of $\mathcal{S}$ connected to these c' measurements, and since the A_{ij}'s

are all 1 for the edges connecting these nodes, they must all have the same nonzero gap g. ∎

We now need another definition. At any step of the algorithm, let $\mathcal{T}$ denote the set

$$\mathcal{T} = \left\{ i \,\middle|\, y_i \neq \sum_{j=1}^{n} A_{ij} x_j \right\}. \qquad (14)$$

Lemma 2 (Decrease in $|\mathcal{T}|$). *After the first step of the algorithm, the cardinality of the set $\mathcal{T}$ decreases at least by 1.*

Proof: According to the proof of Lemma 1, we have found a variable node with $c' > \frac{c}{2}$ measurements with identical nonzero gap g. Setting $\hat{x}_j = \hat{x}_j + g$ sets the gap on these c' equations to zero. However, it may make some zero gaps on the remaining $c - c'$ measurements nonzero. Nonetheless, since $c' - (c - c') = 2c' - c \geq 1$ (note that $c' - c/2 \geq \frac{1}{2}$) the cardinality of $\mathcal{T}$ decreases at least by one. ∎

We can now proceed to the main induction argument.

Lemma 3 (Induction). *Consider a regular left degree c bipartite graph with n variable nodes and m parity check nodes. Assume further that the graph is an $(\alpha n, \frac{3}{4}c)$ expander and consider Algorithm 1. If for all iterations of the algorithm up to step l:*

(1) $\left| \mathcal{S}^{(l')} \right| < \alpha n$, $l' = 1, \ldots, l$, *where $\mathcal{S}^{(l')}$ is the same definition as in (9), except for at the l'-th iteration.*
(2) *There always exists a variable node such that $c' > \frac{c}{2}$ of the measurement equations it participates in have identical nonzero gap g.*
(3) $\left| \mathcal{T}^{(l')} \right| \leq \left| \mathcal{T}^{(l'-1)} \right| - 1$, *for $l' = 1, \ldots, l$, where $\mathcal{T}^{(l')}$ is the same as in the definition (14), except at the l'-th iteration.*

Then at the $(l+1)$-th iteration we have

(i) $\left| \mathcal{S}^{(l+1)} \right| < \alpha n$
(ii) *If $y \neq A\hat{x}$, there always exists a variable node such that $c' > \frac{c}{2}$ of the measurement equations it participates in have identical nonzero gap g.*
(iii) $\left| \mathcal{T}^{(l+1)} \right| \leq \left| \mathcal{T}^{(l)} \right| - 1$

Proof: Let us begin with claim (ii). The argument is very similar to that of the proof of Lemma 1, which we essentially repeat here. Due to assumption (1) in the lemma, $\left| \mathcal{S}^{(l)} \right| < \alpha n$. Therefore we can apply the property of the expander with $\beta = \frac{3}{4}$ to $\mathcal{S}^{(l)}$ to conclude that

$$\left| \mathcal{N}(\mathcal{S}^{(l)}) \right| > \frac{3}{4}c \left| \mathcal{S}^{(l)} \right|. \qquad (15)$$

As before, we divide the set $\mathcal{N}(\mathcal{S}^{(l)})$ into two disjoint sets: $\mathcal{N}_{unique}(\mathcal{S}^{(l)})$ comprised of those elements of $\mathcal{N}(\mathcal{S}^{(l)})$ that are connected to only one edge of $\mathcal{S}^{(l)}$ and $\mathcal{N}_{>1}(\mathcal{S}^{(l)})$ which are the remaining elements of $\mathcal{N}(\mathcal{S}^{(l)})$ that are connected to more than one edges emanating from $\mathcal{S}^{(l)}$. Clearly, (15) implies

$$\left| \mathcal{N}_{unique}(\mathcal{S}^{(l)}) \right| + \left| \mathcal{N}_{>1}(\mathcal{S}^{(l)}) \right| > \frac{3}{4}c \left| \mathcal{S}^{(l)} \right|. \qquad (16)$$

Counting the edges emanating from $\mathcal{N}(S^{(l)})$ leads to

$$\left| \mathcal{N}_{unique}(\mathcal{S}^{(l)}) \right| + 2 \left| \mathcal{N}_{>1}(\mathcal{S}^{(l)}) \right| \leq c \left| \mathcal{S}^{(l)} \right|, \qquad (17)$$

since the total number of edges is $c \left| \mathcal{S}^{(l)} \right|$ and since some of the nodes in $\mathcal{N}_{>1}(\mathcal{S}^{(l)})$ may have more than 2 nodes emanating from them. Eliminating $\left| \mathcal{N}_{>1}(\mathcal{S}^{(l)}) \right|$ from the inequalities (16) and (17) yields

$$\left| \mathcal{N}_{unique}(\mathcal{S}^{(l)}) \right| > \frac{c}{2} \left| \mathcal{S}^{(l)} \right|, \qquad (18)$$

which implies that there must be at least one element of $\mathcal{S}^{(l)}$ that is connected to $c' > \frac{c}{2}$ elements of $\mathcal{N}_{unique}(\mathcal{S}^{(l)})$. But since this is the only element of $\mathcal{S}^{(l)}$ connected to these c' nodes, and since the A_{ij}'s are all 1 for the edges connecting these nodes, they must all have the same nonzero gap g.

This establishes (ii). Establishing (iii) is similar to the proof of Lemma 2. We have already found a variable node with $c' > \frac{c}{2}$ measurements with identical nonzero gap g. Setting $\hat{x}_j^{(l+1)} = \hat{x}_j^{(l)} + g$ sets the gap on these c' equations to zero. However, it may make some zero gaps on the remaining $c - c'$ measurements nonzero. Nonetheless, since $c' - (c - c') = 2c' - c \geq 1$ (note that $c' - c/2 \geq \frac{1}{2}$), the cardinality of $\mathcal{T}^{(l+1)}$ decreases at least by one compared to $\mathcal{T}^{(l)}$.

This establishes (iii). We finally turn to (i). Note that, since in each iteration of Algorithm 1 we change the value of only one entry of $\hat{x}$, the cardinality of the set $\mathcal{S}^{(l')}$ can change at most by one. Since, due to assumption (1) of the lemma we have $\mathcal{S}^{(l)} < \alpha n$, (iii) can only be violated if $\mathcal{S}^{(l+1)} = \alpha n$. Let us assume this and arrive at a contradiction. Note that we can apply the property of the expander with $\beta = \frac{3}{4}$ to the set $\mathcal{S}^{(l+1)}$ to obtain

$$\left| \mathcal{N}(\mathcal{S}^{(l+1)}) \right| > \frac{3}{4}c\alpha n. \qquad (19)$$

Once again, we divide the set $\mathcal{N}(\mathcal{S}^{(l+1)})$ into two disjoint sets: $\mathcal{N}_{unique}(\mathcal{S}^{(l+1)})$ and $\mathcal{N}_{>1}(\mathcal{S}^{(l+1)})$. Clearly, (19) implies

$$\left| \mathcal{N}_{unique}(\mathcal{S}^{(l+1)}) \right| + \left| \mathcal{N}_{>1}(\mathcal{S}^{(l+1)}) \right| > \frac{3}{4}c\alpha n. \qquad (20)$$

Counting the edges emanating from $\mathcal{N}(S^{(l+1)})$ leads to

$$\left| \mathcal{N}_{unique}(\mathcal{S}^{(l+1)}) \right| + 2 \left| \mathcal{N}_{>1}(\mathcal{S}^{(l+1)}) \right| \leq c\alpha n \qquad (21)$$

since the total number of edges is $c\alpha n$ and since some of the nodes in $\mathcal{N}_{>1}(\mathcal{S}^{(l)})$ may have more than 2 nodes emanating from them. (20) and (21) imply

$$\left| \mathcal{N}_{unique}(\mathcal{S}^{(l+1)}) \right| > \frac{c}{2}\alpha n. \qquad (22)$$

Since the nodes in $\mathcal{N}_{unique}(\mathcal{S}^{(l+1)})$ are connected to unique elements in $\mathcal{S}^{(l+1)}$, we conclude that $\mathcal{N}_{unique}(\mathcal{S}^{(l+1)}) \subseteq \mathcal{T}^{(l+1)}$. This in turn implies that

$$\left| \mathcal{T}^{(l+1)} \right| > \frac{c}{2}\alpha n. \qquad (23)$$

Note, however, that since $k \leq \alpha n/2$ and the left degree of the graph is c, at the beginning of the algorithm we have $\left| \mathcal{T}^{(0)} \right| \leq \frac{c}{2}\alpha n$. However, from assumption (3) and property (iii), which we just established, we know that $\left| \mathcal{T}^{(l')} \right|$ is a decreasing function for all $l' \leq l + 1$. Therefore,

$$\left| \mathcal{T}^{(l+1)} \right| < \left| \mathcal{T}^{(0)} \right| \leq \frac{c}{2}\alpha n, \qquad (24)$$

which contradicts (23). This establishes (i) and hence all claims of the lemma. ∎

The above sequence of Lemmas establishes the following main result regarding Algorithm 1.

Theorem 2 (Validity of Algorithm 1). *Consider a regular left degree bipartite graph with n variable nodes and m parity check nodes. Assume further that the graph is an $(\alpha n, \frac{3}{4}c)$ expander and consider its corresponding A matrix. Let $x \in \mathcal{R}^n$ be an arbitrary vector with at most $k \le \alpha n/2$ nonzero entries and consider the m measurements*

$$y = Ax. \tag{25}$$

Then Algorithm 1 finds the value of x in at most $kc \le \frac{c}{2}\alpha n$ iterations. If we assume that the bipartite graph has a regular right degree, we will have a recovery algorithm with complexity linear in n.

Proof: The theorem has essentially been proven in Lemmas 1, 2 and 3. We essentially have shown that at each iteration the cardinality of the set $\mathcal{T}^{(l)}$ decreases by at least one. Since the initial cardinality is *at most* kc, $\mathcal{T}^{(l)}$ will be empty after *at most* kc steps. But, of course, an empty $\mathcal{T}^{(l)}$ implies that the algorithm has found x (This is because in this process $\mathcal{S}$ is always smaller than αn and we can see that a non-zero vector x' satisfying $Ax' = 0$ must have larger than αn nonzero elements following essentially the same arguments as in the proof of Lemma 3). If the bipartite graph has a regular right degree, then in each iterative step of algorithm 1, we only need a fixed number of operations to update the variable nodes and its related measurements by keeping track of the list of variable nodes. ∎

Remarks: Here we can allow for $k = \Theta(n)$ nonzero entries in x since α is a constant (not going to zero as n grows) which depends on the expander graph. The number of measurements is $m = rn$, where r can take any value from $(0, 1)$ and determines the value of α.

IV. COMPRESSIVE SENSING FOR APPROXIMATELY SPARSE SIGNAL USING EXPANDER GRAPHS

In this section, we will give preliminary analytic results on expander-graph-based compressive sensing for approximately sparse signals. In an approximately sparse signal vector, only a few signal entries are significant and the remaining signal entries are near zero but possibly not exactly zero. In practice, the approximately sparse model is a more realistic model for signals. Here we use the same measurement matrix as in the previous section except that we apply it to approximately sparse signals. We also assume a two-level ('near-zero' and 'significant') signal model for the approximately sparse signal vector. (Of course, this is a coarse signal model, but it captures the nature of approximately sparse signal vectors) The entries of the 'near-zero' level in the signal vector are near-zero elements taking values from the set $[-\lambda, +\lambda]$ while the 'significant' level of entries take values from the set $\{x | (L - \Delta) \le |x| \le (L + \Delta)\}$, where $L > \Delta$ and $L > \lambda$. Let $\rho = max\{2\Delta, \lambda\}$ and d be the regular right check node degree. Now we apply the following signal recovery algorithm to y with the measurement matrix A.

Algorithm 2. 1) *Start with $\hat{x} = 0_{n \times 1}$.*
2) *If $\|y - A\hat{x}\|_\infty \le \rho d$, determine the positions and signs of the significant components in x as the positions and signs of the non-zero signal components in $\hat{x}$; exit.
Else, find one variable node, say $\hat{x}_j$, such that of the c measurement equations it participates in $c' > \frac{c}{2}$ of them are in either of the following categories:*
   a) *They have gaps which are of the same sign and have absolute values between $L - \Delta - \lambda - \rho(d-1)$ and $L + \Delta + \lambda + \rho(d-1)$. Moreover, there exists a number t in the set $\{x | x = 0, |x| = (L - \Delta), |x| = (L + \Delta)\}$ such that $|y - A\hat{x}|$ are all $\le \rho d$ over these c' measurements if we change $\hat{x}_j$ to t.*
   b) *They have gaps which are of the same sign and have absolute values between $2L - 2\Delta - \rho(d-1)$ and $2L + 2\Delta + \rho(d-1)$. Moreover, there exist a number t in the set $\{x | x = 0, |x| = (L - \Delta), |x| = (L + \Delta)\}$ such that $|y - A\hat{x}|$ are all $\le \rho d$ over these c' measurements if we change $\hat{x}_j$ to t.*
3) *Reset $\hat{x}_j = t$. Go to 2).*

The following theorem establishes the validity of Algorithm 2.

Theorem 3 (Validity of Algorithm 2). *Consider a bipartite graph with n variable nodes and m parity check nodes. Assume further that the graph is an $(\alpha n, \frac{3}{4}c)$ expander with regular right degree d and regular left degree c. Denote the corresponding measurement matrix as A. Let $x \in \mathcal{R}^n$ be an arbitrary vector with at most $k \le \alpha n/2$ significant signal components and assume that $max\{\rho(2d-1) + \Delta + \lambda, \rho(2d-2) + 3\Delta + \lambda\} < L$. Consider the m measurements*

$$y = Ax. \tag{26}$$

Then Algorithm 2 correctly finds the sign and positions of the significant components of x in at most $kc \le \frac{c}{2}\alpha n$ iterations with complexity linear in n.

Proof: The arguments here basically follow the same reasoning as in the proof of Lemma 1, Lemma 2, Lemma 3 and Theorem 3. But now we define the set $\mathcal{S}$ as the set of variable nodes j's such that x_j and $\hat{x}_j$ are on different signal levels or have opposite signs while both being on the 'significant' signal level. If a variable node $j \in S$, then $L - \Delta - \lambda \le |x_j - \hat{x}_j| \le L + \Delta + \lambda$ or $2(L - \Delta) \le |x_j - \hat{x}_j| \le 2(L + \Delta)$. Also notice that $|x_j - \hat{x}_j| \le \rho$ if x_j and $\hat{x}_j$ are both in the near-zero signal level or have the same sign while both being on the 'significant' signal level. Define the set $\mathcal{T}$ as the set of measurements where $\|y - A\hat{x}\|$ have values larger than ρd. Notice that after each iteration, we can always decrease the cardinality of $\mathcal{T}$ by at least 1. ∎

Remarks: This analysis can be extended to the noisy measurements cases. Also, after knowing the signs and locations of significant components, the estimation for their amplitudes can be further refined.

V. SIMULATION RESULTS

In this section, we give simulation results of the proposed schemes for $n = 1000$, $m = 500$ and different sparsity levels. Although there exist explicit constructions of the required bipartite expander graphs as given
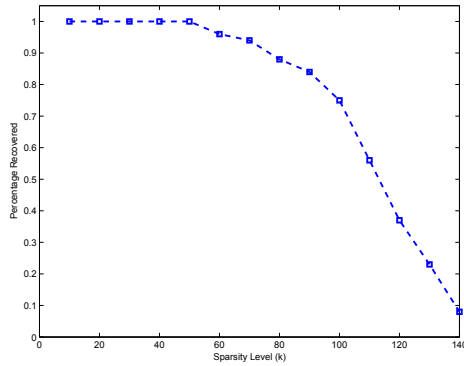
Fig. 2. The Probability of Recovering a k-sparse Signal

in [26], we will simulate Algorithm 1 using randomly generated bipartite graphs since it is easier to implement for evaluation at the current stage. Compressive sensing using explicitly constructed expander graphs for large n is an important topic of future study. We set the regular left degree c as 5 and generate the random bipartite graphs as in [23] using random permutations of size $n \times c = 5000$ since a randomly generated graph will be an expander graph with high probability. After randomly generating bipartite graphs (thus the measurement matrix A), uniformly select the support set for the k non-zero elements of the signal vector x. The nonzero entries for x are sampled as i.i.d. Gaussian random variables. We repeat the experiments for 1000 independent trials for each k in the simulation figure. As the simulation results showed, we can recover up to the sparsity level of $k = 50$, 5 percent of the signal vector length. Although the performance is not comparable with l_1 minimization method in [6], we should notice that the randomly generated bi-partite graphs are not optimized expander graphs and that the signal recovery for each instance works instantly, taking only several hundredth seconds (approximately 0.06 seconds for $k = 50$) using Matlab 6.5 on a Windows Platform with 1.7G Hz Intel Pentium CPU and 512M memory. It is much faster than the various linear programming solvers which will usually take more than one second.

VI. CONCLUSIONS

We propose to use bipartite expander graphs for compressive sensing of sparse signals and show that we can perform compressive sensing with deterministic performance guarantees at a cost of $O(n)$ signal recovery complexity even when the number of non-zero elements grows linearly in n. Also, explicit constructions of the considered expander graphs exist [26]. We give preliminary analysis results of applying the new methods to approximately sparse signals. Simulation results verified the effectiveness and efficiency of the new methods. Explicitly constructing more powerful expander graphs and further studying the applications of expander-graph-based approach to approximately sparse signal in noisy environments are two important future research topics.

REFERENCES

[1] E. Candes, J. Romberg and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information", *IEEE Transactions on Information Theory*, 52(2) pp. 489 - 509, Feb. 2006.

[2] E. Candes and T .Tao, "Near Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?", *IEEE Trans. on Information Theory*, 52(12) pp. 5406 - 5425, Dec. 2006.

[3] E. Candes. "Compressive sampling" *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, 2006.

[4] D .Donoho, "Compressed Sensing", *IEEE Transactions on Information Theory, 52(4), pp. 1289 - 1306, April 2006*

[5] S.S. Chen, "Basis Pursuit", *PhD Thesis, Stanford University*, 1995.

[6] E. Candes and T. Tao, "Decoding by Linear Programming", *IEEE Trans. on Information Theory*, 51(12), pp. 4203 - 4215, Dec. 2005.

[7] S.S. Chen, D.L. Donoho, M.A. Saunders, "Atomic Decomposition by Basis Pursuit", *SIAM Journal on Scientific Computing*, 1996.

[8] E. Candes and J. Romberg, "Practical Signal Recovery from Random Projections", (Preprint, Jan. 2005)

[9] J. Tropp and A. Gilbert, "Signal Recovery From Partial Information Via Orthogonal Matching Pursuit" (Preprint, 2005).

[10] D. Donoho and Y. Tsaig, "Fast solution of l_1-norm minimization problems when the solution may be sparse", (Preprint, 2006).

[11] R.A. DeVore, "Deterministic Constructions of Compressed Sensing Matrices", (Preprint, 2007).

[12] J. Haupt and R. Nowak, "Signal Reconstruction from Noisy Random Projections", *IEEE Trans. on Information Theory*, 52(9), pp. 4036-4048, Sep. 2006.

[13] E. Candes, J. Romberg, and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements", *Communications on Pure and Applied Mathematics*, 59(8), pp. 1207-1223, Aug. 2006.

[14] D. Donoho and J. Tanner, "Neighborliness of Randomly-Projected Simplices in High Dimensions" *Proc. National Academy of Sciences*, 102(27), pp. 9452-9457, 2005.

[15] D. Donoho, "For Most Large Underdetermined Systems of Linear Equations, The Minimal l_1-Norm Solution Is Also The Sparsest Solution", *Communications on Pure and Applied Mathematics*, 59(6), pp. 797-829, June 2006.

[16] G. Cormode and S. Muthukrishnan, "Combinatorial Algorithms for Compressed Sensing", (Technical Report DIMACS TR 2005-40, 2005).

[17] A. Gilbert, M. Strauss, J. Tropp, and R. Vershynin, "Sublinear, Small-space Approximation of Compressible Signals and Uniform Algorithmic Embeddings",(Preprint, 2005).

[18] S. Sarvotham, Dror Baron, and Richard Baraniuk, "Sudocodes - Fast Measurement and Reconstruction of Sparse Signals", *Proc. IEEE Int. Symposium on Information Theory (ISIT)*, Seattle, WA, July 2006.

[19] S. Sarvotham, D. Baron, and R. Baraniuk, "Compressed Sensing Reconstruction via Belief Propagation", (Preprint, 2006).

[20] A. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, and M. Strauss, "Near-Optimal Sparse Fourier Representations via Sampling", *Proc. ACM Symposium on Theory of Computing (STOC)*, 2002.

[21] J.J. Fuchs, "Sparsity and Uniqueness for some Specific Underdetermined Linear Systems". *IEEE ICASSP*, vol V, pp.729-732, Philadelphia, March 2005.

[22] Mehmet Akakaya and Vahid Tarokh, "Performance Bounds on Sparse Representations Using Redundant Frames". *http://arxiv.org/abs/cs.IT/0703045*, 2007

[23] D. Burshtein and G. Miller, "Expander graph arguments for message passing algorithms," *IEEE Transactions on Information Theory, volume* 47, pp. 782-790, February 2001.

[24] J. Laska, S. Kirolos, Y. Massoud, R. Baraniuk, A. Gilbert, M. Iwen, and M. Strauss, "Random Sampling for Analog-to-Information Conversion of Wideband Signals", *Proc. IEEE Dallas Circuits and Systems Workshop (DCAS)*, Dallas, TX, 2006.

[25] D. Donoho, Y. Tsaig, I. Drori and J.C. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit", (Preprint 2006).

[26] M. Capalbo, O. Reingold, S. Vadhan, A. Wigderson, "Randomness Conductors and Constant-Degree Expansion Beyond the Degree/2 Barrier", *Proc. of the 34th STOC*, pp. 659-668, 2002.

[27] M. Sipser and D. Spielman, "Expander codes", *IEEE Transactions on Information Theory*, Vol 42, No 6, pp. 1710-1722, 1996.

[28] L.A. Bassalygo and M.S. Pinsker, "Complexity of an optimum nonblocking switching network without reconnections", *Problems in Information Transmission*, vol 9 no 1, pp. 84-87, 1973.

[29] D. Peleg and E. Upfal, "Constructing disjoint paths on expander graphs", *Combinatorica*, vol 9, n0 3, pp. 289-313, 1989.

[30] E. Upfal and A. Wigderson, "How to share memory in a distributed system", *Journal of the Association of Computing Machines*, vol 34, no pp. 116-127, 1987.

[31] E. Ben-Sasson and A. Wigderson, "Short proofs are narrow—resolution made simple", *Annual ACM Symposium on Theory of Computing*, pp. 517-526, 1999.