Efficient and Robust Compressed Sensing Using Optimized Expander Graphs

Sina Jafarpour, Weiyu Xu, Babak Hassibi, *Member, IEEE*, and Robert Calderbank, *Fellow, IEEE*

***Abstract*—Expander graphs have been recently proposed to construct efficient compressed sensing algorithms. In particular, it has been shown that any n-dimensional vector that is k-sparse can be fully recovered using $O(k \log n)$ measurements and only $O(k \log n)$ simple recovery iterations. In this paper, we improve upon this result by considering expander graphs with expansion coefficient beyond $\frac{3}{4}$ and show that, with the same number of measurements, only $O(k)$ recovery iterations are required, which is a significant improvement when n is large. In fact, full recovery can be accomplished by at most $2k$ very simple iterations. The number of iterations can be reduced arbitrarily close to k, and the recovery algorithm can be implemented very efficiently using a simple priority queue with total recovery time $O(n \log(\frac{n}{k}))$. We also show that by tolerating a small penalty on the number of measurements, and not on the number of recovery iterations, one can use the efficient construction of a family of expander graphs to come up with explicit measurement matrices for this method. We compare our result with other recently developed expander-graph-based methods and argue that it compares favorably both in terms of the number of required measurements and in terms of the time complexity and the simplicity of recovery. Finally, we will show how our analysis extends to give a robust algorithm that finds the position and sign of the k significant elements of an almost k-sparse signal and then, using very simple optimization techniques, finds a k-sparse signal which is close to the best k-term approximation of the original signal.**

***Index Terms*—Compressive sensing, expander graphs, sparse recovery, sparse sensing matrices, unique neighborhood property.**

I. Introduction

THE goal of *compressive sampling* or *compressed sensing* is to replace the conventional sampling and reconstruction operations with a more general combination of linear measurement and optimization in order to acquire certain kinds of signals at a rate significantly below Nyquist. Formally, suppose we have a signal x which is sparse. We can model x as an n-dimensional vector that has at most k nonzero components. We aim to find an $m \times n$ matrix A such that m, the number of measurements, becomes as small as possible (and can be efficiently stored), and x can be recovered efficiently from $y = Ax$.

The original approach employed dense random matrices and random projections. It has been shown that if the matrix A has a certain **Restricted Isometry Property (RIP-2)**, that is, it almost preserves the Euclidean norm of all $3k$-sparse vectors, then A can be used in compressed sensing and the decoding can be accomplished using *linear programming and convex programming* methods [1]. This is a geometric approach based on linear and quadratic optimization, and Baraniuk *et al.* [2] showed that the RIP-2 property is a direct consequence of the *Johnson–Linderstrauss lemma* [3] so that many dense random matrices satisfy this property. However, the problem in practice is that the linear and quadratic programming algorithms have cubic complexity in n, which has difficulties as n becomes very large; furthermore, in order to store the whole matrix in memory we still need $O(m \times n)$ which is inefficient too. Another family of random matrices satisfying Restricted Isometry Propert (RIP) property, which do not come from the Johnson–Lindenstrauss family, are random Fourier matrices, obtained by randomly sampling rows of the Fourier matrix. These matrices require $O(m \log n)$ units of storage; however, there are two main issues with these matrices. The first is that the number of required measurements is suboptimal ($k \log^5(n)$ instead of $k \log(\frac{n}{k})$). The second is that while these matrices do require less storage than less structured random matrices, the computational cost of matrix–vector products required in many streaming and sensor network applications is significantly greater than that of the expander-based sensing matrices.

Following [4]–[8], we will show how random dense matrices can be replaced by the adjacency matrices of an optimized family of expander graphs, thereby reducing the space complexity of matrix storage and, more important, the time complexity of recovery to a few very simple iterations. The main idea is that we study expander graphs with expansion coefficient beyond the $\frac{3}{4}$ that was considered in [4], [5]. Furthermore, our results have interesting connection with the belief propagation decoding of the low-density parity-check (LDPC) codes, and the sequential decoding of the expander codes [9].

The remainder of the paper is organized as follows. In Section II, we review the previous results from [4], [5]. In Section III, following the geometric approach of [6], we establish that the adjacency matrix of the expander graphs satisfies a certain RIP for the Manhattan distance between the sparse signals. Using this property, or via a more direct alternative approach, we show how the recovery task becomes much easier. In Section IV,

Manuscript received May 25, 2008; revised March 12, 2009. Current version published August 19, 2009. The work of B. Hassibi was supported in parts by the National Science Foundation (NSF) under Grant CCF 0729203, by the David and Lucille Packard Foundation and by Caltech's Lee Center for Advanced Networking. The work of R. Calderbank was supported in part by NSF under Grant 0701226, by ONR under Grant N00173-06-1-G006, and by AFOSR under Grant FA9550-05-1-0443.

S. Jafarpour is with the Department of Computer Science, Princeton University, Princeton, NJ 08544 USA (e-mail: sina@cs.princeton.edu).

W. Xu and B. Hassibi are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: weiyu@caltech.edu; hassibi@caltech.edu).

R. Calderbank is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: calderbk@math.princeton.edu).

Communicated by J. Romberg, Associate Editor for Signal Processing.

Color versions of Figures 1–3 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2009.2025528

we generalize the algorithm of [4], [5] to expander graphs with expansion coefficient beyond $\frac{3}{4}$. The key difference is that now the progress in each iteration is proportional to $\log n$, as opposed to a constant in [4], [5], and so the time complexity is reduced from $O(k \log n)$ to $O(k)$. We then describe how the algorithm can be implemented using simple data structures very efficiently and show that explicit constructions of the expander graphs impose only a small penalty in terms of the number of measurements, and not the number of iterations, the recovery algorithm requires. We also compare our result to previous results based on random projections and to other approaches using the adjacency matrices of expander graphs. In Section V, we generalize the analysis to a family of almost k-sparse signals; (after a few very simple iterations) the robust recovery algorithm proposed in [4] empowered with optimized expander graphs finds the position and the sign of the k significant elements of an almost k-sparse signal. Given this information, we then show how *the restricted isometry property* of the expander graphs lets us use very efficient optimization methods with analytical solution to find a k-sparse signal that approximates the original signal very efficiently. Section VI concludes the paper.

II. Prior Result: $O(K \log N)$ Recovery

A. Basic Definitions

Xu and Hassibi [4] proposed a new scheme for compressed sensing with deterministic recovery guarantees based on combinatorial structures called *unbalanced expander graphs*:

Definition 1 (Bipartite Expander Graph, Informally): An expander graph [10] $G(V, E)$ is a d regular graph with $|V| = v$ vertices, such that

1) G is sparse (ideally d is much smaller than v);
2) G is *"well connected"*.

Various formal definitions of the second property define the various types of the expander graphs. The expander graph used in [4], [5] which has suitable properties for compressed sensing is the "$\frac{3}{4}$-*vertex expander graph*."

Definition 2 (Bipartite $\frac{3}{4}$-Expander Graph): A bipartite left regular graph with n variable nodes, m parity-check nodes, and left degree d will be $(\alpha n, \frac{3}{4}d)$ expander graph, for $0 < \alpha < \frac{1}{2}$, if for every subset of variable nodes S with cardinality $|S| \leq \alpha n$, the number of neighbors connected to S, denoted by $N(S)$ is strictly larger than $\frac{3}{4}d|S|$, i.e., $|N(S)| > \frac{3}{4}d|S|$.

Using the probabilistic method, Pinsker and Bassylago [11] showed the existence of $\frac{3}{4}-$ expander graphs and they showed that any random left-regular bipartite graph, with very high probability, is an expander graph. Then Capalbo *et al.* [12] gave an explicit construction for these expander graphs.

Theorem 1: Let $0 < r < 1$ be a fixed constant. Then for large enough n there exists an $(\alpha n, \frac{3}{4}d)$ expander graph G with n variable nodes and $\frac{n}{r}$ parity-check nodes with constant left degree (not growing with n), or sublinearly grows with , and some $0 < \alpha < 1$. Furthermore, the explicit zig-zag construction can deterministically construct the expander graph.

Using Hoeffding's inequality and Chernoff bounds, Xu and Hassibi [5] showed the following theorem.

Theorem 2: For any k, if n is large enough, there exists a left regular bipartite graph with left degree d for some number d, which is $(k, \frac{3}{4}d)$ expander graph with $m = O(k \log n)$ parity-check nodes.

B. Recovery Algorithm

Suppose $\hat{x}$ is the original n-dimensional k-sparse signal, and the adjacency matrix of an $(\alpha n, \frac{3}{4}d)$ expander graph is used as the measurement matrix for the compressed sensing. We are given $y = A\hat{x}$ and we want to recover $\hat{x}$. Xu and Hassibi [4] proposed the following algorithm.

Algorithm 1 Left Degree Dependent Signal Recovery algorithm

1: Initialize $x = 0_{n \times 1}$.
2: **if** $y = Ax$ **then**
3: output x and exit.
4: **else**
5: find a variable node say x_j such that more than half of the measurements it participate in, have identical gap g.
6: set $x_j \leftarrow x_j + g$, and go to 2.
7: **end if**

In the above algorithm the gap is defined as follows.

Definition 3 (Gap): Let $\hat{x}$ be the original signal and $y = A\hat{x}$. Furthermore, let x be our estimate for $\hat{x}$. For each value y_i, we define a gap g_i as

$$g_i = y_i - \sum_{j=1}^{n} A_{ij} x_j.$$

Xu and Hassibi [4] proved the following theorem that bounds the number of steps required by the algorithm to recover $\hat{x}$.

Theorem 3: Suppose A is the adjacency matrix of an expander graph satisfying Definition 2, and $\hat{x}$ is an n-dimensional k sparse signal (with $k < \frac{\alpha n}{2}$), and $y = A\hat{x}$. Then Algorithm 1 will always find a signal x which is k sparse and for which $Ax = y$. Furthermore, the algorithm requires at most $O(kd)$ iterations, where k is the sparsity level of the signal and d is the left side degree of the expander graph.

Let us now consider the consequences of the above theorem for the expander graphs in Theorems 1 and 2. In Theorem 1, the sparsity can grow proportional to n (since $k < \frac{\alpha n}{2}$) and the algorithm will be fast; the algorithm requires $O(kd)$ iterations and since d is a constant independent of n, the number of iterations is $O(k)$. We also clearly need $O(n)$ measurements. In Theorem 2, the sparsity level k is fixed (does not grow with n) and the number of measurements needs to be $O(k \log n)$, which is desired. Once more the number of required iterations is $O(kd)$. However, in this case, Xu and Hassibi showed the following negative result for $(k, \frac{3}{4}d)$ expander graphs.

Theorem 4: Consider a bipartite graph with n variable nodes and m measurement nodes, and assume that the graph is a $(k, \frac{3}{4}d)$ expander graph with regular left degree d. Then if $m = O(k \log n)$ we have $d = \Omega(\log n)$.

This theorem implies that for a $(k, \frac{3}{4}d)$ expander graph, the recovery algorithm needs $O(k \log n)$ iterations. The main contribution of the current paper is that the number of iterations can be reduced to $O(k)$. The key idea is to use expanders with expander coefficient beyond $\frac{3}{4}$.

Remark: Theorem 3 does not imply the full recovery of the sparse signal. It only states that the output of the recovery algorithm will be a k sparse signal x such that $Ax = A\hat{x}$ where $\hat{x}$ is the original signal. However, in the next section we show how an interesting property of the expander graphs called the RIP-1 property, implies full recovery. We also give a direct proof by showing that the null-space of the adjacency matrix of an expander graph cannot be "too sparse."

III. Expander Codes, RIP-1 Property, and Full-Recovery Principle

A. Expander Codes

Compressed sensing has many properties in common with coding theory. The recovery algorithm is similar to the decoding algorithms of error correcting codes but over R^m instead of a finite field. As a result, several methods from coding theory have been generalized to derive compressed sensing algorithms. Among these methods are the generalization of Reed–Solomon codes by Akcakaya and Tarokh [13], recent results by Calderbank *et al.* [14]–[17], based the group properties of the algebraic codes, and Parvaresh *et al.* [18], based on list decoding.

In 1996, Sipser and Spielman [9] used expander graphs to build a family of linear error-correcting codes with linear decoding time complexity. These codes belong to class of error correcting codes called *low-density parity-check (LDPC) codes*. The work of Xu and Hassibi is a generalization of these expander codes to compressed sensing.

B. Norm One Restricted Isometry Property

The standard RIP is an important sufficient condition that enables compressed sensing using random projections. Intuitively, it says that the measurement almost preserves the Euclidean distance between any two sufficiently sparse vectors. This property implies that recovery using l_1 minimization is possible if a random projection is used for measurement. Berinde *et al.* in [7] showed that expander graphs satisfy a very similar property called "RIP-1" which states that if the adjacency matrix of an expander graph is used for measurement, then the Manhattan (l_1) distance between two sufficiently sparse signals is preserved by measurement. They used this property to prove that l_1-minimization is still possible in this case. However, we will show in this section how RIP-1 can guarantee that the algorithm described above will have full recovery.

Following [6], [7], we show that the RIP-1 property can be derived from the expansion property and will guarantee the uniqueness of sparse representation.

We begin with the definition of the "unbalanced lossless vertex expander graphs" with expansion coefficient $1 - \epsilon$, bearing in mind that we will be interested in $1 - \epsilon > \frac{3}{4}$.

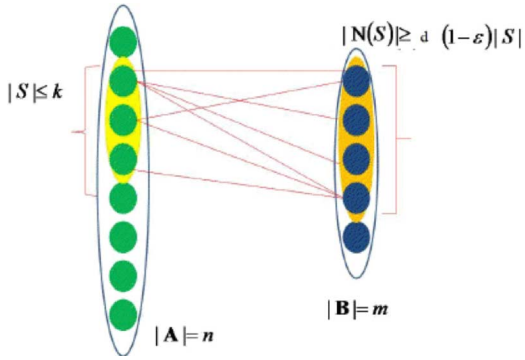

Fig. 1. (k, ϵ) vertex expander graph.

Definition 4 (Unbalanced Lossless Expander Graphs): An $(l, 1-\epsilon)$-*unbalanced bipartite expander graph* (see Fig. 1) is a bipartite graph $V = (A, B), |A| = n, |B| = m$, where A is the set of variable nodes and B is the set of parity nodes, with regular left degree d such that for any $S \subset A$, if $|S| \leq l$ then the set of neighbors $N(S)$ of S has size $N(S) > (1-\epsilon)d|S|$.

The following claim follows from the Chernoff bounds [6].[1]

Claim 1: for any $\frac{n}{2} \geq l \geq 1, \epsilon > 0$ there exists an $(l, 1-\epsilon)$ expander with left degree

$$d = O\left(\frac{\log\left(\frac{n}{l}\right)}{\epsilon}\right)$$

and right set size

$$m = O\left(\frac{l \log\left(\frac{n}{l}\right)}{\epsilon^2}\right).$$

Lemma 1 (RIP-1 Property of the Expander Graphs: Let $A_{m\times n}$ be the adjacency matrix of a $(k, 1-\epsilon)$ expander graph E, then for any k-sparse vector $x \in \mathbb{R}^n$ we have:

$$(1-2\epsilon)d\|x\|_1 \leq \|Ax\|_1 \leq d\|x\|_1. \tag{1}$$

Proof: The upper bound is trivial using the triangle inequality, so we only prove the lower bound.

The left side inequality is not influenced by changing the position of the coordinates of x, so we can assume that they are in a nonincreasing order: $|x_1| \geq |x_2| \geq \cdots \geq |x_n|$. Let E be the set of edges of G and $e_{ij} = (x_i, y_j)$ be the edge that connects x_i to y_j. Define

$$E_2 = \{e_{ij} : \exists k < i \text{ s.th. } e_{kj} \in E\}.$$

Intuitively, E_2 is the set of the collision edges. Let

$$T_i = \{e_{i'j} \in E_2 \text{ s.th. } i' \leq i\}$$

and $a_i = |T_i|$.

Clearly, $a_1 = 0$; moreover, by the expansion property of the graph for any k' less than or equal to $ka_{k'}$ is less than or equal to $\epsilon dk'$. Finally, since the graph is k-sparse we know that for each k greater than k, $x_{k''}$ is zero. Therefore

$$\sum_{e_{ij}\in E_2} |x_i| = \sum_{i=1}^{n} |x_i|(a_i - a_{i-1})$$
$$= \sum_{i\leq k} a_i(|x_i| - |x_{i+1}|)$$

[1]This claim is also used in the expander codes construction.

$$\leq \sum_{i\leq k} \epsilon d i(|x_i| - |x_{i+1}|)$$
$$\leq \sum_{i\leq k} |x_i| \epsilon d$$
$$= \epsilon d \, ||x||_1 .$$

Now the triangle inequality, and the definition of E_2 imply

$$\begin{aligned}
||Ax||_1 &= \sum_{j=1}^{m} \left| \sum_{e_{ij}\in E} x_i \right| = \sum_{j=1}^{m} \left| \sum_{e_{ij}\in E_2} x_i + \sum_{e_{ij}\notin E_2} x_i \right| \\
&\geq \sum_{j=1}^{m} \left(\left| \sum_{e_{ij}\notin E_2} x_i \right| - \left| \sum_{e_{ij}\in E_2} x_i \right| \right) \\
&= \sum_{j=1}^{m} \left(\left| \sum_{e_{ij}\notin E_2} x_i \right| + \left| \sum_{e_{ij}\in E_2} x_i \right| - 2 \left| \sum_{e_{ij}\in E_2} x_i \right| \right) \\
&= \sum_{e_{ij}\notin E_2} |x_i| + \sum_{e_{ij}\in E_2} |x_i| - 2 \sum_{e_{ij}\in E_2} |x_i| \\
&\geq d||x||_1 - 2\epsilon d||x||_1 \\
&= (1-2\epsilon)d||x||_1.
\end{aligned}$$

□

C. Full Recovery

The full recovery property now follows immediately from Lemma 1.

Theorem 2 (Full Recovery): Suppose $A_{m\times n}$ is the adjacency matrix of a $(3k, 1-\epsilon)$ expander graph, and suppose x_1 is a k-sparse and x_2 is a $2k$-sparse vector, such that $Ax_1 = Ax_2$. Then $x_1 = x_2$.

Proof: Let $z = x_1 - x_2$. Since x_1 is *k-sparse* and x_2 is $2k$-sparse, z is $3k$-sparse.[2] By Lemma 1 we have

$$||x_1 - x_2||_1 \leq \frac{1}{(1-2\epsilon)d} ||Ax_1 - Ax_2||_1 = 0$$

hence $x_1 = x_2$. □

Note that the proof of the above theorem essentially says that the adjacency matrix of a $(3k, 1-\epsilon)$ expander graph does not have a null vector that is $3k$ sparse. We will also give a direct proof of this result (which does not appeal to RIP-1) since it gives a flavor of the arguments to come.

Lemma 2 (Null Space of A: Suppose $A_{m\times n}$ is the adjacency matrix of a $(3k, 1-\epsilon)$ expander graph with $\epsilon \leq \frac{1}{2}$. Then any nonzero vector in the null space of A, i.e., any $z \neq 0$ such that $Az = 0$, has more than $3k$ nonzero entries.

Proof: Define $\mathcal{S}$ to be the support set of z. Suppose that z has at most $3k$ nonzero entries, i.e., that $|\mathcal{S}| \leq 3k$. Then from the expansion property we have that $|N(\mathcal{S})| > (1-\epsilon)d|\mathcal{S}|$. Partitioning the set $N(\mathcal{S})$ into the two disjoint sets $N_1(\mathcal{S})$, consisting of those nodes in $N(\mathcal{S})$ that are connected to a single node in $\mathcal{S}$, and $N_{>1}(\mathcal{S})$, consisting of those nodes in $N(\mathcal{S})$ that are connected to more than a single node in $\mathcal{S}$, we may write $N_1(\mathcal{S}) + N_{>1}(\mathcal{S}) > (1-\epsilon)d|\mathcal{S}|$. Furthermore, counting the edges connecting $\mathcal{S}$ and $N(\mathcal{S})$, we have $|N_1(\mathcal{S})| + 2|N_{>1}(\mathcal{S})| \leq d|\mathcal{S}|$. Combining these latter two inequalities yields $|N_1(\mathcal{S})| > (1-2\epsilon)d|\mathcal{S}| \geq 0$. This implies that there is at least one nonzero element in z that participates in only one equation of $y = Az$. However, this contradicts the fact that $Az = 0$ and so z must have more than $3k$ nonzero entries. □

[2] $\|z\|_0 \leq \|x_1\|_0 + \|x_2\|_0 = 3k$

IV. Our Results: Efficient Full Recovery

A. Efficient $O(k \log \frac{n}{k})$ Sensing With $O(n \log(\frac{n}{k})))$ Recovery Time Algorithm

In this section, we show the general unbalanced bipartite expander graphs introduced in Definition 4 work much better than $\frac{3}{4}$-expanders, in the sense that they give the measurement size $O(k \log \frac{n}{k})$ which is up to a constant the optimum measurement size, and simultaneously yields a simple recovery algorithm which needs only $O(k)$ simple iterations and a total running time of $O(n \log(\frac{n}{k})))$.

Before proving the result, we introduce some notations used in the recovery algorithm and in the proof.

Definition 5 (Gap): Recall the definition of the gap from Definition 3. At each iteration t, let G_t be the support[3] of the gaps vector at that iteration

$$G_t = \text{support}\,(\vec{g_t}) = \left\{ i | y_i \neq \sum_{j=1}^{n} A_{ij} x_j \right\}.$$

Definition 6: At each iteration t, we define S_t an indicator of the difference between the estimate $\hat{x}$ and x

$$S_t = \text{support}(\hat{x} - x) = \{j : \hat{x_j} \neq x_j\}.$$

Now we are ready to state the main result.

Theorem 6 (Expander Recovery Algorithm): Let $A_{m\times n}$ be the adjacency matrix of a $(2k, 1-\epsilon)$ expander graph, where $\epsilon \leq 1/4$, and $m = O(k \log(\frac{n}{k})))$. Then, for any k-sparse signal $\hat{x}$, given $y = A\hat{x}$, the expander recovery algorithm (Algorithm 2 below) recovers $\hat{x}$ successfully in at most $2k$ iterations.

Algorithm 2 Expander Recovery Algorithm

1: Initialize $x = 0_{n\times 1}$.
2: **if** $y = Ax$ **then**
3: output x and exit.
4: **else**
5: find a variable node say x_j such that at least $(1-2\epsilon)d$ of the measurements it participate in, have identical gap g.
6: set $x_j \leftarrow x_j + g$, and go to 2.
7: **end if**

The proof is virtually identical to that of [4], except that we consider a general $(1-\epsilon)$ expander, rather than a $\frac{3}{4}$-expander, and consists of the following lemmas.

- The algorithm never gets stuck, and one can always find a coordinate j such that x_j is connected to at least $(1-2\epsilon)d$ parity nodes with identical gaps.

[3] set of nonzero elements

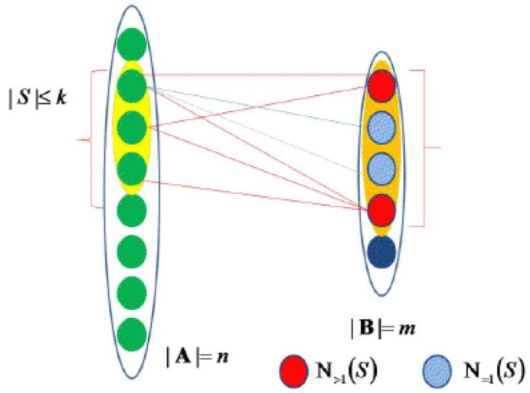

Fig. 2. Progress Lemma.

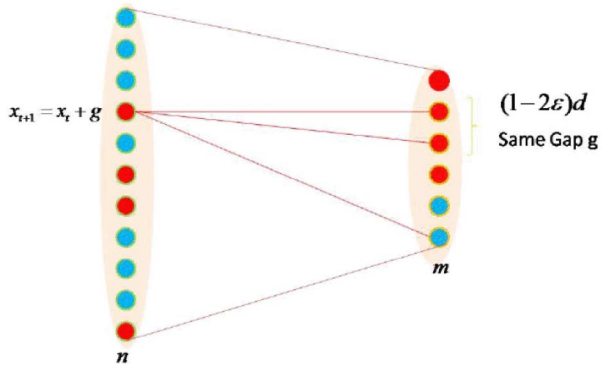

Fig. 3. Gap Elimination Lemma.

- With certainty, the algorithm will stop after at most $2k$ rounds. Furthermore, by choosing ϵ small enough, the number of iterations can be made arbitrarily close to k.

Lemma 3 (Progress): Suppose at each iteration t, $S_t = \{j : \hat{x_j} \neq x_j\}$. If $|S_t| < 2k$ then always there exists a variable node x_j such that at least $(1-2\epsilon)d$ of its neighbor check nodes have the same gap g.

Proof: We will prove that there exists a coordinate j, such that x_j is uniquely connected to at least $(1-2\epsilon)d$ check nodes, in other words, no other nonzero variable node is connected to these nodes. This immediately implies the lemma.

Since $|S_t| < 2k$ by the expansion property of the graph it follows that $N(S_t) \geq (1-\epsilon)d|S_t|$. Now we are going to count the neighbors of S_t in two ways. Fig. 2 shows the notations in the progress lemma.

We partition the set $N(S_t)$ into two disjoint sets.

- $N_1(S_t)$: The vertices in $N(S_t)$ that are connected only to one vertex in S_t.
- $N_{>1}(S_t)$: The other vertices (that are connected to more than one vertex in S_t).

By double counting the number of edges between variable nodes and check nodes we have

$$|N_1(S_t)| + |N_{>1}(S_t)| = |N(S_t)| > (1-\epsilon)d|S_t|$$
$$|N_1(S_t)| + 2|N_{>1}(S_t)| \leq \#\text{edges between } S_t, N(S_t) = d|S_t|$$

This gives

$$|N_{>1}(S_t)| < \epsilon d|S_t|$$

hence

$$|N_1(S_t)| > (1-2\epsilon)d|S_t| \tag{5}$$

so by the pigeonhole principle, at least one of the variable nodes in S_t must be connected uniquely to at least $(1-2\epsilon)d$ check nodes. □

Lemma 4 (Gap Elimination): At each step t if $|S_t| < 2k$ then $|G_{t+1}| < |G_t| - (1-4\epsilon)d$.

Proof: By the previous lemma, if $|S_t| < 2k$, there always exists a node x_j that is connected to at least $(1-2\epsilon)d$ nodes with identical nonzero gap, and hence to at most $2\epsilon d$ nodes possibly with zero gaps. Setting the value of this variable node to zero sets the gaps on these uniquely connected neighbors of x_j to zero, but it may make some zero gaps on the remaining $2\epsilon d$ neighbors nonzero. So at least $(1-2\epsilon)d$ coordinates of G_t will become zero, and at most $2\epsilon d$ its zero coordinates may become nonzero. Hence

$$|G_{t+1}| < |G_t| - (1-2\epsilon)d + 2\epsilon d = |G_t| - (1-4\epsilon)d. \tag{6}$$

Fig. 3 shows the gap elimination. □

Remark: The key to accelerating the algorithm is the preceding lemma. For a $\frac{3}{4}$ expander, $\epsilon = \frac{1}{4}$ and so $|G_{t+1}| < |G_t|$, which only guarantees that $|G_{t+1}|$ is reduced by a constant number. However, when $\epsilon < \frac{1}{4}$, we have $|G_{t+1}| < |G_t| - (1-4\epsilon)d$, which means that $|G_{t+1}|$ is guaranteed to decrease proportionally to d. Since $d = \Omega(\log n)$, we save a factor of $\log n$.

The following lemma provides a direct connection between the size of G_t and the size of S_t.

Lemma 5 (Connection): If at iteration t, $|S_t| < 2k$, then $(1-2\epsilon)d|S_t| \leq |G_t|$.

Proof: By (5), $|N_1(S_t)| > (1-2\epsilon)d|S_t|$, also each node in $N_1(S_t)$ has nonzero gap and so is a member of G_t. □

Lemma 6 (Preservation): At each step t if $|S_t| < 2k$, after running the algorithm, we have $|S_{t+1}| < 2k$.

Proof: Since at each step we are only changing one coordinate of x, we have $|S_{t+1}| = |S_t| + 1$, so we only need to prove that $S_{t+1} \neq 2k$.

Suppose for a contradiction that $|S_{t+1}| = 2k$, and partition $N(S_{t+1})$ into two disjoint sets.

1) $N_1(S_{t+1})$: The vertices in $N(S_{t+1})$ that are connected only to one vertex in S_{t+1}.
2) $N_{>1}(S_{t+1})$: The other vertices (that are connected to more than one vertex in S_{t+1}).

The argument is similar to that given above; by double counting the number of vertices in $N_1(S_{t+1}), N_{>1}(S_{t+1})$ one can show that

$$|N_1(S_{t+1})| \geq (1-2\epsilon)d2k.$$

Now we have the following facts.

- $|N_1(S_{t+1})| > |G_{t+1}|$: Coordinates in $N_1(S_{t+1})$ are connected uniquely to coordinates in S_{t+1}, hence each coordinate in $N_1(S_{t+1})$ has nonzero gap.

- $|G_{t+1}| > |G_1|$: gap elimination from Lemma 4.
- $|G_1| \leq kd$: $x, \hat{x}$ differ in at most k coordinates, so $Ax, A\hat{x}$ can differ in at most kd coordinates.

As a result we have

$$(1-2\epsilon)2dk \leq |N_1(S_{t+1})| \leq |G_{t+1}| \leq |G_1| \leq kd \quad (8)$$

This implies $\epsilon \geq \frac{1}{4}$ which contradicts the assumption $\epsilon < \frac{1}{4}$. □

Proof of the Theorem 6: Preservation (Lemma 6) and progress (Lemma 3) together immediately imply that the algorithm will never get stuck. Also, by Lemma 4 we had shown that $|G_1| \leq kd$ and $|G_{t+1}| < |G_t| - (1-4\epsilon)d$. Hence, after at most $T = \frac{k}{1-4\epsilon}$ steps we will have $|G_T| = 0$ and this together with the connection lemma implies that $|S_T| = 0$, which is the exact recovery of the original signal.

Note that we have to choose $\epsilon < \frac{1}{4}$, and as an example, by setting $\epsilon = \frac{1}{8}$ the recovery needs at most $2k$ iterations. □

Remark: The condition $\epsilon < \frac{1}{4}$ in the theorem is necessary. Even $\epsilon = \frac{1}{4}$ leads to a $\frac{3}{4}$ expander graph (Definition 2), which needs $O(k \log n)$ iterations.

B. Explicit Construction of Expander Graphs

In the definition of the expander graphs (Definition 4), we noted that probabilistic methods prove that such expander graphs exist and furthermore, that any random graph, with high probability, is an expander graph. Hence, in practice it may be sufficient to use random graphs instead of expander graphs.

Although there is no efficient explicit construction for the expander graphs of Definition 4, there exists explicit construction for a class of expander graphs which are very close to the optimum expanders of Definition 4. Recently, Guruswami *et al.* [19], based on the Parvaresh–Vardy codes [20], proved the following theorem.

Theorem 7 (Explicit Construction of Expander Graphs): For any constant $\alpha > 0$, and any $n, k, \epsilon > 0$, there exists a $(k, 1-\epsilon)$ expander graph with left degree

$$d = O\left(\left(\frac{\log n}{\epsilon}\right)^{1+\frac{1}{\alpha}}\right)$$

and number of right side vertices

$$m = O(d^2 k^{1+\alpha})$$

which has an efficient deterministic explicit construction.

Since our previous analysis was only based on the expansion property, which does not change in this case, a similar result holds if we use these expanders.

C. Efficient Implementation of the Algorithm and a Comparison With an Alternative Integrated Geometric–Combinatorial Approach

We now compare our approach with recent analysis by Berinde *et al.* [7]. This paper integrates Indyk's previous work which was based on randomness extractors [8] and a combinatorial algorithm (employing an alternative approach to the RIP-1 results of Berinde–Indyk [6]) based on geometric convex optimization methods and suggests a recursive recovery algorithm which takes $m' = O(m \log m)$ sketch measurements and needs a recovery time $O(m \log^2 n)$. The recovery algorithm exploits the hashing properties of the expander graphs, and is sublinear. However, it has difficulties for practical implementation.

By comparison, our recovery algorithm is a simple iterative algorithm, that needs $O(k \log n)$ sketch measurements, and our decoding algorithm consists of at most $2k$ very simple iterations. Each iteration can be implemented very efficiently (see [4]) since the adjacency matrix of the expander graph is sparse with all entries 0 or 1. Even the very naive implementation of the algorithm as suggested in this paper works efficiently in practice. The reason is that the unique neighborhood property of the expander graphs is much stronger than what we needed to prove the accuracy of our algorithm. Indeed, it can be shown [21], [22] that most of the variable nodes have $(1-\epsilon/2)d$ unique neighbors, and hence at each of the $O(k)$ iterations, the algorithm can find one desired node efficiently. The efficiency of the algorithm can also be improved by using a priority queue data structure. The idea is to use preprocessing as follows: For each variable node v_i compute the median of its neighbors $m_i = \mathrm{Med}(N(v_i))$ and also compute n_i the number of neighbors with the same value m_i (Note that if a node has $(1-2\epsilon)d$ unique neighbors, their median should also be among them.) Then construct the priority queue based on the values n_i, and at each iteration extract the root node from the queue, perform the gap elimination on it, and then, if required, make the correction on corresponding dD variable nodes. The main computational cost of this variation of the algorithm will be the cost of building the priority queue which is $O(n \log(\frac{n}{k}))$; finding the median of d elements can be done in $O(\log \frac{n}{k})$ and building a priority queue requires linear computational time.

V. Almost Sparse Signals and Robust Recovery

In this section, we show how the analysis using optimized expander graphs that we proposed in the previous section can be used to illustrate that the robust recovery algorithm in [4] can be done more efficiently in terms of the sketch size and recovery time for a family of almost k-sparse signals. With this analysis we will show that the algorithm will only need $O(k \log n)$ measurements. Explicit constructions for the sketch matrix exist and the recovery consists of two simple steps. First, the combinatorial iterative algorithm in [4], which is now empowered with the optimized expander sketches, can be used to find the position and the sign of the k largest elements of the signal x. Using an analysis similar to the analysis in Section IV we will show that the algorithm needs only $O(k)$ iterations, and similar to the previous section, each iteration can be done efficiently using a priority queue. Then restricting to the position of the k largest elements, we will use a robust theorem in expander graphs to show that simple optimization methods that are now restricted on k-dimensional vectors can be used to recover a k sparse signal that approximates the original signal with very high precision.

Before presenting the algorithm we will define precisely what we mean for a signal to be almost k sparse.

Definition 7 (Almost k-Sparse Signal: A signal $x \in R^n$ is said to be almost k-sparse iff it has at most k large elements and the remaining elements are very close to zero and have very small magnitude. In other words, the entries of the "near-zero" level in the signal take values from the set $[-\lambda, \lambda]$ while the "significant" levels of entries take values from the set $S = \{x : |L-\Delta| \leq |x| \leq |L+\Delta\}$. By the definition of the almost sparsity we have $|S| \leq k$. The general assumption for almost sparsity is intuitively the fact that the total magnitude of the almost sparse terms should be small enough such that it does not disturb the overall structure of the signal which may make the recovery impossible or very erroneous. Since $\sum_{x \notin S} |x| \leq n\lambda$ and the total contribution of the "near-zero" elements is small we can assume that $n\lambda$ is small enough. We will use this assumption throughout this section.

In order to make the analysis for almost k-sparse signals simpler we will use a optimized expander graph which is right-regular as well.[4] The following lemma which appears as Lemma 2.3 in [23] gives us a way to construct right-regular expanders from any expander graph without disturbing its characteristics.

Lemma 7 (Right-Regular Expanders): From any left-regular $(k, 1-\epsilon)$ unbalanced expander graph G with left size n, right size m, and left degree d it is possible to efficiently construct a left-right-regular $(k, 1-\epsilon)$ unbalanced expander graph H with left size n, right size $m' \leq 2m$, left side degree $d' \leq 2d$, and right side degree $D = [\frac{nd}{m}]$

Corollary 1: There exists a $(k, 1-\epsilon)$ left-right unbalanced expander graph with left side size n, right side size $m = O(k \log \frac{n}{k})$, left side degree $d = O(\log \frac{n}{k})$, and right side degree

$$D = O\left(\frac{n \log \frac{n}{k}}{k \log \frac{n}{k}}\right) = O\left(\frac{n}{k}\right).$$

Also based on the explicit constructions of expander graphs, explicit construction for right-regular expander graphs exists.

We will use the above right-regular optimized expander graphs in order to perform robust signal recovery efficiently. The following algorithm generalizes the k-sparse recovery algorithm and can be used to find the position and sign of the k largest elements of an almost k-sparse signal x from $y = Ax$. At each iteration t in the algorithm, let

$$\rho_t = 2t\Delta + (D-t-1)\lambda \quad \text{and} \quad \phi_t = 2t\Delta + (D-t)\lambda$$

where $D = O(n)$ is the right side degree of the expander graph. Throughout the algorithm we will assume that $L > 2k\Delta + D\lambda$. Hence, the algorithm is appropriate for a family of almost k-sparse signals for which the magnitude of the significant elements is large enough. We will assume that k is a small constant; when k is large with respect to n, $(k = \theta(n))$, the $(\alpha n, \frac{3}{4})$ constant degree expander sketch proposed in [4] works well.

Expander Recovery Algorithm for Almost k-sparse Signals

1: Initialize $x = 0_{n \times 1}$.
2: **if** $|y - Ax|_\infty \leq \phi_t$ **then**
3: determine the positions and signs of the significant components in $\hat{x}$ as the positions and signs of the nonzero signal components in x; go to 8.
4: **else**
5: find a variable node say x_j such that at least $(1-2\epsilon)d$ of the measurements it participate in are in either of the following categories:
   a) They have gaps which are of the same sign and have absolute values between $L - \Delta - \lambda - \rho_t$ and $L + \Delta + \lambda + \rho_t$. Moreover, there exists a number $G \in \{0, L+\Delta, L-\Delta\}$ such that $|y - Ax|$ are all $\leq \phi_t$ over these $(1-2\epsilon)d$ measurements if we change x_j to G.
   b) They have gaps which are of the same sign and have absolute values between $2L - 2\Delta - \rho_t$ and $2L + 2\Delta + \rho_t$. Moreover, there exists a number $G \in \{0, L+\Delta, L-\Delta\}$ such that $|y - Ax|$ are all $\leq \phi_t$ over these $(1-2\epsilon)d$ measurements if we change x_j to G.
6: set $x_j \leftarrow G$, and go to 2 for next iteration.
7: **end if**
8: pick the set of k significant elements of the candidate signal x_T. Let A' be the sensing matrix A restricted to these entries, output $A'^{\dagger} y$.

In order to prove the algorithm we need the following definitions which are the generalization of the similar definitions in the exactly k-sparse case.

Definition 8: At each iteration t, we define S_t an indicator of the difference between $\hat{x}$ and the estimate x :

$S_t = \{j | \hat{x_j}$ and x_j in different levels or large with different signs$\}$.

Definition 9 (Gap): At each iteration t, let G_t be the set of measurement elements in which at least one "significant" elements from x contributes

$$G_t = \{i \| y_i - \sum_{j=1}^{n} A_{ij} x_j |_\infty > \lambda D\}.$$

Theorem 8 (Validity of the Algorithm 3): The first part of the algorithm will find the position and sign of the k significant elements of the signal x (for more discussion see [4]).

Proof: This is very similar to the proof of the validity of the exactly k-sparse recovery algorithm. We will exploit the following facts.

- $\hat{x}$ is almost k-sparse so it has at most k significant elements. Initially $S_0 = k$ and $G_0 \leq kd$.
- Since at each iteration only one element x_j is selected, at each iteration t there are at most t elements x_j such that both x_j and $\hat{x}_j$ are in the significant level with the same sign.
- If $|S_t| < 2k$ then $|S_{t+1}| < 2k$ (Preservation Lemma), and by the neighborhood theorem at each round $(1-2\epsilon)|S_t|d \leq |G_t|$.
- If $S_t < 2k$ by the Neighborhood Theorem there exists a node $x_j \in S_t$ which is the unique node in S_t that is connected to at least $(1-2\epsilon)d$ parity-check nodes. This node

[4]The right-regularity assumption is just for the simplicity of the analysis and as we will show it is not mandatory.

is in S_t. It differs from its actual value in the significance level or at sign. In the first case, part a) of the recovery algorithm will detect and fix it and in the second case, part b) of the algorithm will detect and fix it. For further discussion please refer to [4].

- As a direct result, $|G_{t+1}| \leq |G_t| - (1-4\epsilon)d$. So after $T = \frac{kd}{(1-4\epsilon)d}$ iterations we have $|G_T| = 0$. Consequently, $|S_T| = 0$ after at most $2k$ iterations.

This means that after at most $2k$ iterations, the set $S_T = \{j|\hat{x}_j \text{ and } x_j \text{ in different levels or with different signs.}\}$ will be empty and hence the position of the k largest elements in $\hat{x}_T$ will be the position of the k largest elements in x. □

Knowing the position of the k largest elements of $\hat{x}$ it is easier to recover a good k-sparse approximation. If k is large, a parallel version of Algorithm 2 may be applicable. If k is small, analytical solutions are achievable. Based on the RIP-1 property of the expander graph, we propose a way to recover a good approximation for x efficiently and analytically. We need the following lemma which is a direct result of the RIP-1 property of the expander graphs and is proved in [6], [7]

Lemma 8: Consider any $u \in R^n$ such that $||Au||_1 = b$, and let S be any set of k coordinates of u. Then we have

$$||u_S||_1 \leq \frac{b}{d(1-2\epsilon)} + \frac{2\epsilon}{1-2\epsilon}||u||_1$$

and

$$\frac{1-4\epsilon}{1-2\epsilon}||u_S||_1 \leq \frac{b}{d(1-2\epsilon)} + \frac{2\epsilon}{1-2\epsilon}||u_{\bar{S}}||_1.$$

Using Lemma 8, we prove that the following minimization recovers a k-sparse signal very close to the original signal.

Theorem 9 (Final Recovery): Suppose x is an almost k-sparse signal and $y = Ax$ is given where $y \in R^m$ and $m = O(k\log(\frac{n}{k})))$. Also, suppose S is the set of the k largest elements of x. Now let A' be a submatrix of A restricted to S. Hence, the following minimization problem can be solved analytically with solution $v = A'^{\dagger}y$ (where $A'^{\dagger}$ is the pseudoinverse of A'), and recovers a k-sparse signal v with close distance to the original x in the $\ell 1$ metric:

$$\min||A'v - y||_2.$$

Proof: Suppose v is the recovered signal. Since v is k-sparse we have $Av = A'v$ and hence

$$\begin{aligned} ||Av - Ax||_1 &= ||Av - y||_1 \\ &= ||A'v - y||_1 \\ &\leq \sqrt{m}||A'v - y||_2 \\ &\leq \sqrt{m}||A'x_S - y||_2 \\ &= \sqrt{m}||Ax_S - Ax||_2 \\ &\leq \sqrt{m}\lambda D\sqrt{m} \\ &\leq \sqrt{m^2}\lambda D \\ &= mD\lambda = nd\lambda. \end{aligned} \tag{12}$$

The first two equations are only definitions. The third one is the Cauchy–Schwartz inequality. The fourth is from the definition of v, and the last is due to the almost k-sparsity of x. Since v is k-sparse and x is almost k-sparse with the same support, we may set $u = x - v$ in Lemma 8 to obtain

$$\begin{aligned} \frac{1-4\epsilon}{1-2\epsilon}||u_S||_1 &\leq \frac{||Ax - Av||_1}{d(1-2\epsilon)} + \frac{2\epsilon}{1-2\epsilon}||u_{\bar{S}}||_1 \\ &\leq \frac{n\lambda}{(1-2\epsilon)} + \frac{2\epsilon}{1-2\epsilon}||u_{\bar{S}}||_1 \\ &\leq \frac{n\lambda}{(1-2\epsilon)} + \frac{2\epsilon}{1-2\epsilon}n\lambda \\ &= O(n\lambda). \end{aligned}$$

As a result, since the signal is almost k-sparse, the value of $n\lambda$ is small, and hence the recovered k-sparse signal is close to the best k-term approximation of the original signal.

Remark: Recall that the right-regularity assumption is just to make the analysis simpler. As we mentioned before, it is not necessary for the first part of the algorithm. For the second part, it is used in the inequality $|Av - Ax| \leq \sqrt{m}|Ax_S - Ax|_2$.

However, denoting the ith row of A by A_i, we have

$$||Ax_S - Ax||_2 = \sqrt{m}\sqrt{\sum_{i=1}^{m}(A_i(x_S - x))^2} \leq \sqrt{m}\sqrt{\sum_{i=1}^{m}(\lambda D_i)^2}$$

where D_i denotes the number of ones in the ith row of A. (In the right-regular case, $D_i = D$, for all i.) Therefore

$$||Ax_S - Ax||_2 \leq \sqrt{m}\lambda\sum_{i=1}^{m}D_i = \sqrt{m}\lambda nd.$$

The only difference with the constant D_i case is the extra $\sqrt{m}$ but this does not affect the end result.

VI. Conclusion

In this paper, we used a combinatorial structure called an unbalanced lossless vertex expander graph, in order to perform efficient deterministic compressed sensing and recovery. We showed how using expander graphs one needs only $O(k\log\frac{n}{k})$ measurements and the recovery needs only $O(k)$ iterations. Also, we showed how the expansion property of the expander graphs guarantees the full recovery of the original signal. Since random graphs are with high probability expander graphs and it is very easy to generate random graphs, in many cases we might use random graphs instead. However, we showed that in cases that recovery guarantees are needed, just with a little penalty on the number of measurements and without affecting the number of iterations needed for recovery, one can use another family of expander graphs for which explicit constructions exist. We also compared our result with a result by Berinde *et al.*. [7], and showed that our algorithm has advantages in terms of the number of required measurements, and the simplicity of the algorithm for practical use. Finally, we showed how the algorithm can be modified to be robust and handle almost k-sparse signals. In order to do this we slightly modified the algorithm by using right-regular optimized expander graphs to find the position of the k largest elements of an almost k-sparse signal. Then exploiting the robustness of the RIP-1 property of the expander graphs we showed how this information can be combined with efficient optimization methods to find a k-sparse approximation for x very efficiently. However, in the

TABLE I
PROPERTIES OF k-SPARSE RECONSTRUCTION ALGORITHMS THAT EMPLOY EXPANDER MATRICES WITH m ROWS AND n COLUMNS TO RECONSTRUCT A VECTOR x FROM ITS NOISY SKETCH $Ax + e$

Paper	Approach	Geometric/ /Combinatorial	Number of Measurements m	Number of Iterations	Worst Case Time Complexity	k-term Approximation	Noise Resilience	Explicit Construction
[23]	$\ell 1$ Embedding of Expander Codes	Combinatorial	$O\left(k(\log n)^{\log n}\right)$	-	$O(n^3)$	Compressible	No	[23]
[4]	Unique Neighborhood	Combinatorial	$O(n)$	O(n)	$O(n)$	Almost k-sparse	Sparse Noise	[12]
[5]	Unique Neighborhood	Combinatorial	$O(k \log n)$	$O(k \log n)$	$O(n \log n)$	Almost k-sparse	Sparse Noise	[19]
[7]	RIP-1	Geometric	$O\left(k \log\left(\frac{n}{k}\right)\right)$	-	$O(n^3)$	Compressible	Sparse Noise	[19]
[7]	Hash Functions	Combinatorial	$O(k \log \frac{n}{k} \log n)$	$O(k \log \frac{n}{k} \log n)$	$O(k \log^3 n)$	Compressible	Sparse Noise	[19]
[22]	All Unique Neighborhoods	Combinatorial	$O\left(k \log\left(\frac{n}{k}\right)\right)$	$O(n)$	$O\left(n \log\left(\frac{n}{k}\right)\right)$	Compressible	Sparse Noise	[19]
[24]	RIP-1	Geometric	$O\left(k \log\left(\frac{n}{k}\right)\right)$	$R = O\left(\log\left(\frac{d\|x\|_1}{\|e\|_1}\right)\right)$	$O\left(n \log\left(\frac{n}{k}\right) \log R\right)$	Compressible	Sparse Noise	[19]
This Paper	Unique Neighborhood	Combinatorial	$O\left(k \log\left(\frac{n}{k}\right)\right)$	$O(k)$	$O\left(n \log\left(\frac{n}{k}\right)\right)$	Almost k-sparse	Sparse Noise	[19]

almost k-sparsity model that we used nonsparse components should have "almost equal" magnitudes. This is because of the assumption that $L > k\Delta$ which restricts the degree of deviation for significant components. As a result, one important future work will be finding robust algorithms based on more general assumptions, or investigating alternative noise models in which the expander graphs are beneficial.

Table I compares our results with the other expander-based algorithms.

Remark: After submission of this paper, Indyk and Ruzic [22], and Berinde, Indyk, and Ruzic [24] proposed new compressed sensing algorithms based on the properties of the expander graphs. those algorithms are similar to the CoSaMP algorithm [25], from the orthogonal matching framework, and are designed to be robust against more general noise and compressible signals; however, this comes with a cost on complexity of the algorithm and its analysis. The algorithm that we proposed in this paper is much simpler, also the analysis why the algorithm works is only based on the unique neighborhood properties of the expander graphs. In contrast, those algorithms require a complicated preprocessing step, and their analysis uses complex combinatorial statements and is based on more complicated properties of ideal expander graphs. In practice, it will be harder to get all the required expander properties from random graphs. The explicit construction of the expander graphs satisfying those properties are also suboptimal in the number of measurements and the constants required. However, their algorithms have the advantage of being more robust against the noisy measurements, especially if the noise is also sparse.

ACKNOWLEDGMENT

The authors would like to thank Piotr Indyk, Justin Romberg, the readers and the anonymous reviewers of the paper for their insights and suggestions.

REFERENCES

[1] E. Candés and M. Wakin, "People hearing without listening: An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.

[2] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[3] W. Johnson and J. Lindenstrauss, "Extensions of Lipschitz mapping into a Hilbert space," *Contemp. Math.*, vol. 26, pp. 189–206, Sep. 1984.

[4] W. Xu and B. Hassibi, "Efficient compressive sensing with deterministic guarantees using expander graphs," in *Proc. IEEE Information Theory Workshop*, Lake Tahoe, CA, Sep. 2007, pp. 414–419.

[5] W. Xu and B. Hassibi, "Further results on performance analysis for compressive sensing using expander graphs," in *Conf. Rec. 41st Asilomar Conf. Signals, Systems and Computers (ACSSC 2007)*, Thousan Oaks, CA, Nov. 2008, pp. 621–625.

[6] R. Berinde and P. Indyk, "Sparse Recovery Using Sparse Random Matrices," MIT, Cambridge, MA, Tech. Rep. MIT-CSAIL-TR-2008-001, Jan. 2008.

[7] R. Berinde, A. Gilbert, P. Indyk, H. Karloff, and M. Strauss, "Combining geometry and combinatorics: A unified approach to sparse signal recovery," in *Proc. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2008, pp. 798–805.

[8] P. Indyk, "Explicit constructions for compressed sensing of sparse signals," in *Proc. 19th Annu. ACM-SIAM Symp. Discrete Algorithms (SODA)*, San Francisco, CA, Jan. 2008.

[9] M. Sipser and D. Spielman, "Expander codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1710–1722, Nov. 1996.

[10] P. Sarnak, "What is an expander," *Notices Amer. Math. Soc.*, vol. 51, no. 7, Aug. 2004.

[11] L. Bassalygo and M. Pinsker, "Complexity of an optimum nonblocking switching network without reconnections," *Probl. Inf. Transm.*, vol. 9, no. 1, pp. 289–313, 1973.

[12] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson, "Randomness conductors and constant degree expansions beyond the degree 2 barrier," in *Proc. 34th Annu.ACM Symp. Theory of Computing (STOC)*, Montreal, QC, Canada, May 2002, pp. 659–668.

[13] M. Akcakaya and V. Tarokh, "A frame construction and a universal distortion bound for sparse representations," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2443–2450, Jun. 2008.

[14] L. Applebaum, S. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Appl. Comput. Harmonic Anal.*, vol. 26, no. 2, pp. 283–290, Mar. 2008.

[15] S. Howard, R. Calderbank, and S. Searle, "A fast reconstruction algorithm for deterministic compressive sensing using second order Reed-Muller codes," in *Proc. Conf. Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2008, pp. 11–15.

[16] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property", submitted for publication.

[17] R. Calderbank, S. Howard, and S. Jafarpour, "Deterministic sensing with groups of random variables", submitted for publication.

[18] F. Parvaresh and B. Hassibi, "List decoding for compressed sensing", unpublished.

[19] V. Guruswami, C. Umans, and S. Vadhan, "Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes," *J. ACM*, vol. 56, no. 4, Jun. 2009, Article No. 20.
[20] F. Parvaresh and A. Vardy, "Correcting errors beyond the Guruswami-Sudan radius in polynomial time," in *Proc. 46th Ann. IEEE Symp. Foundations of Computer Science*, Pittsburgh, PA, Oct. 2008, pp. 285–294.
[21] S. Hoory, N. Linial, and A. Wigderson, "Expander graphs and their applications," *Bull. (New Series) Amer. Math. Soc.*, vol. 43, pp. 439–561, 2006.
[22] P. Indyk and M. Ruzic, "Near-optimal sparse recovery in the $\ell 1$ norm," in *Proc. IEEE Symp. Foundations of Computer Science (FOCS)*, Philadelpha, PA, Oct. 2008, pp. 199–207.
[23] V. Guruswami, J. Lee, and A. Razborov, "Almost Euclidean subspaces of ℓ_1 via expander codes," in *Proc. 19th Annu. ACM-SIAM Symp. Discrete Algorithms (SODA)*, San Francisco, CA, Jan. 2008, pp. 649–658.
[24] R. Berinde, P. Indyk, and M. Ruzic, "Practical near-optimal sparse recovery in the $\ell 1$ norm," in *Proc. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2008, pp. 198–205.
[25] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comp. Harmonic Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.

Sina Jafarpour received the B.Sc. degree in computer engineering with great distinction from Sharif University of Technology, Tehran, Iran, in July 2007 and the M.A. degree in computer science from Princeton University, Princeton, NJ, in January 2009.

He is currently a Ph.D. candidate in the Department of Computer Science, Princeton University, where he has been working towards the Ph.D. degree as a member of the Machine Learning and Theoretical Computer Science groups in the Department of Computer Science, and the Van Gogh Project team in the Program in Applied and Computational Mathematics. His research interests include compressive sampling, machine learning, coding theory, and signal processing.

Mr. Jafarpour received the Princeton University Fellowship award in July 2007. In October 2008, his research on Compressive Learning and Detection received one of the five best student paper awards from Google and the New York Academy of Sciences.

Weiyu Xu was born in Rushan City, Shandong Province, China. He received the B.E. degree in information engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2002, the M.S. degree in electronic engineering from Tsinghua University, Beijing, in 2005, and the M.S. degree in electrical engineering from California Institute of Technology, Pasadena, in 2006.

Since 2005, he has been working towards the Ph.D. degree in the Electrical Engineering Department of the California Institute of Technology with a minor in applied and computational mathematics. His research interests are in signal processing, compressive sensing, communication networks, information theory and coding theory.

Mr. Xu was honored as an outstanding graduate of Tsinghua University in 2005, was a recipient of the Information Science and Technology Fellowship at California Institute of Technology and a finalist for the Best Student Paper Award at the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

Babak Hassibi (M'08) was born in Tehran, Iran, in 1967. He received the B.S. degree from the University of Tehran in 1989, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1993 and 1996, respectively, all in electrical engineering.

He has been with the California Institute of Technology, Pasadena, since January 2001, where he is currently Professor and Executive Officer of Elecrical Engineering. From October 1996 to October 1998, he was a Research Associate at the Information Systems Laboratory, Stanford University, Stanford, CA, and from November 1998 to December 2000 he was a Member of the Technical Staff in the Mathematical Sciences Research Center at Bell Laboratories, Murray Hill, NJ. He has also held short-term appointments at Ricoh California Research Center, the Indian Institute of Science, and Linköping University, Sweden. His research interests include wireless communications and networks, robust estimation and control, adaptive signal processing and linear algebra. He is the coauthor of the books (both with A.H. Sayed and T. Kailath) *Indefinite Quadratic Estimation and Control: A Unified Approach to* H^2 and H^∞ *Theories* (New York: SIAM, 1999), and *Linear Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 2000).

Prof. Hassibi is a recipient of an Alborz Foundation Fellowship, the 1999 O. Hugo Schuck best paper award of the American Automatic Control Council (with H. Hindi and S. P. Boyd), the 2002 National Science Foundation Career Award, the 2002 Okawa Foundation Research Grant for Information and Telecommunications, the 2003 David and Lucille Packard Fellowship for Science and Engineering, and the 2003 Presidential Early Career Award for Scientists and Engineers (PECASE), and was a participant in the 2004 National Academy of Engineering "Frontiers in Engineering" program. He has been a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY special issue on "Space–Time Transmission, Reception, Coding and Signal Processing" was an Associate Editor for Communications of the IEEE TRANSACTIONS ON INFORMATION THEORY during 2004–2006, and is currently an Editor for the Journal *Foundations and Trends in Information and Communication*.

Robert Calderbank (M89–SM97–F98) received the B.Sc. degree in 1975 from Warwick University, U.K., the M.Sc. degree in 1976 from Oxford University, U.K., and the Ph.D. degree in 1980 from the California Institute of Technology, Pasadena, all in mathematics.

He is Professor of Electrical Engineering and Mathematics at Princeton University, Princeton, NJ, where he directs the Program in Applied and Computational Mathematics. He joined Bell Telephone Laboratories as a Member of Technical Staff in 1980, and retired from AT&T in 2003 as Vice President of Research. He has made significant contributions to a wide range of research areas, from algebraic coding theory and quantum computing to wireless communication and active sensing.

Dr. Calderbank served as Editor in Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY from 1995 to 1998, and as Associate Editor for Coding Techniques from 1986 to 1989. He was a member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996 and from 2006 to 2008. He was honored by the IEEE Information Theory Prize Paper Award in 1995 for his work on the Z_4 linearity of Kerdock and Preparata Codes (jointly with A. R. Hammons Jr., P. V. Kumar, N. J. A. Sloane, and P. Solé), and again in 1999 for the invention of space-time codes (joint with V. Tarokh and N. Seshadri). He received the 2006 IEEE Donald G. Fink Prize Paper Award and the IEEE Millennium Medal, and was elected to the U.S. National Academy of Engineering in 2005.