Sparse Recovery with Graph Constraints: Fundamental Limits and Measurement Construction

Meng Wang   Weiyu Xu   Enrique Mallada   Ao Tang
School of ECE, Cornell University, Ithaca, NY 14853, USA

Abstract—**This paper addresses the problem of sparse recovery with graph constraints in the sense that we can take additive measurements over nodes only if they induce a connected subgraph. We provide explicit measurement constructions for several special graphs. A general measurement construction algorithm is also proposed and evaluated. For any given graph G with n nodes, we derive order optimal upper bounds of the minimum number of measurements needed to recover any k-sparse vector over G ($M_{k,n}^G$). Our study suggests that $M_{k,n}^G$ may serve as a graph connectivity metric.**

I. INTRODUCTION

Network monitoring is an essential module in the operation and management of communication networks, where one keeps track of network status parameters, such as bandwidth utilization and queueing delay. Since measuring each component (e.g., router) in the network directly can be operationally costly, if feasible at all, the topic of inferring system internal characteristics from indirect end-to-end (aggregate) measurements becomes important. This area is known as network tomography, and has been extensively studied during the last decade or so [8], [11], [13], [21], [23], [27], [34].

Ideally, with the total number of aggregate measurements much smaller than the number of nodes in a network, we still hope to extract the status of each individual node. This is possible if we have prior knowledge of the status (the unknown signal to be recovered). For example, if the signal is sparse, i.e. most entries are zero, we can recover it exactly even though the number of measurements is much smaller than the dimension of the signal. One practical example is that only a small number of bottleneck links in the communication networks experience large delays. *Sparse Recovery* addresses the problem of recovering sparse signals from a smaller number of measurements, and has two different but closely related problem formulations. One is *Compressed Sensing* [4], [9], [10], [17], [18], [22], where the signal is represented by a high-dimensional real vector, and an aggregate measurement is the arithmetical sum of the corresponding real entries. The other is *Group Testing* [19], [20], where the high-dimensional vector is logical, and a measurement is a logical disjunction (**OR**) on the corresponding logical values.

One key question in both compressed sensing and group testing is to design a small number of non-adaptive measurements (either real or logical) such that all the vectors (either real or logical) up to certain sparsity (the support size of a vector) can be correctly recovered. Most existing results, however, rely critically on the assumption that any subset of the values can be aggregated together [9], [17], which is not realistic in network monitoring problems where only objects that form a path or a cycle on the graph [23], or induce a connected subgraph can be aggregated together in the same measurement. Only a few recent works consider graph topological constraints in compressed sensing [14], [22], [25], [33] and group testing [2], [12], [24], [28], [31].

Though directly motivated by network monitoring problems, sparse recovery with graph constraints abstractly models scenarios when certain elements cannot be measured together in a complex system. These constraints can result from various reasons, not necessarily lack of connectivity. Therefore, our results can be potentially useful to other applications besides network tomography.

Here are the main contributions of this paper.
(1) We provide explicit measurement constructions for various graphs. The number of our measurements is less than the existing estimates (e.g. [12], [33]) of the minimum number of measurements required to recover sparse vectors over graphs. (Section III)
(2) We propose a measurement design guideline based on *r-partition* for general graphs and further show some of its properties. (Section IV-A)
(3) A simple measurement design algorithm is proposed for general graphs, and we evaluate its performance both theoretically and numerically. (Section IV-B and V)

We now start with Section II to introduce the model and problem formulation.

II. MODEL AND PROBLEM FORMULATION

Consider a graph $G = (V, E)$, where V denotes the set of nodes with cardinality $|V| = n$ and E denotes the set of links. We assume G is fixed and known throughout the paper. Each node i is associated with a real number x_i, and we say vector $\mathbf{x} = (x_i, i = 1, ..., n)$ is associated with G. Let $T = \{i \mid x_i \neq 0\}$ denote the support of $\mathbf{x}$, and let $\|\mathbf{x}\|_0 = |T|$ denote the number of non-zero entries of $\mathbf{x}$, we say $\mathbf{x}$ is a k-sparse vector if $\|\mathbf{x}\|_0 = k$.

Let $S \subseteq V$ denote a subset of nodes in G. Let E_S denote the subset of links with both ends in S, then $G_S = (S, E_S)$ is the induced subgraph of G. We have the following two assumptions throughout the paper:
(A1): A set S of nodes can be measured together in one measurement if and only if G_S is connected.
(A2): The measurement is an additive sum of values at the corresponding nodes.

(A1) captures the graph constraints. One practical example is a sensor network where the nodes represent sensors and the links represent feasible communication between sensors. For the set S of nodes that induce a connected subgraph, one node u in S monitors the total values corresponding to nodes in S. Every node in S obtains values from its children, if any, on the spanning tree rooted at u, aggregates them with its own value and sends the sum to its parent. Then the fusion center can obtain the sum of values corresponding to all the nodes in S by only communicating with u. (A2) follows from the additive property of many network characteristics, e.g. delays and packet loss rates [23]. However, compressed sensing can also be applied to cases where (A2) does not hold, e.g., the measurements can be nonlinear as in [5], [29].

Let $\mathbf{y} \in \mathcal{R}^m$ ($m \ll n$) denote the vector of m measurements. Let A be an $m \times n$ measurement matrix with its ith row corresponds to the ith measurement, and $A_{ij} = 1$ ($i = 1, ..., m$, $j = 1, ..., n$) if and only if node j is included in the ith measurement and $A_{ij} = 0$ otherwise. We can write in the compact form that $\mathbf{y} = A\mathbf{x}$. We say a measurement matrix A can ***identify all k-sparse vectors*** if and only if $A\mathbf{x}_1 \neq A\mathbf{x}_2$ for every two different vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ that are at most k-sparse. This definition indicates that every vector $\mathbf{x}$ that is at most k-sparse can be recovered from $A\mathbf{x}$ via ℓ_0-minimization, which returns the sparsest vector among all that can produce the measurement vector $A\mathbf{x}$. Sparse recovery theory indicates that one can identify n-dimensional vectors from m ($m \ll n$) measurements as long as the vectors are sparse enough.
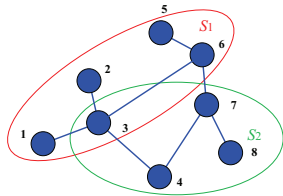

Fig. 1. Network Example

With the above assumptions, A is a 0-1 matrix, and for each row of A, the set of nodes that correspond to '1' should form a connected induced subgraph of G. In Fig. 1, we can measure nodes in S_1 and S_2 separately, and the measurement matrix is

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

We remark here that in group testing with graph constraints, the requirements for the measurement matrix A are the same, while group testing differs from compressed sensing only in that (1) $\mathbf{x}$ is a logical vector, and (2) the operations used in each group testing measurement are the logical "AND" and "OR". All arguments and results in this paper are in the compressed sensing setup if not otherwise specified, and we also compare our results with group testing results for special networks. Note that for recovering 1-sparse vectors, the numbers of measurements required by compressed sensing and group testing are the same.

TABLE I
SUMMARY OF KEY NOTATIONS

Notation	Meaning
G_S	Subgraph of G induced by S
$M_{k,n}^G$	Minimum number of measurements needed to identify k-sparse vectors associated with G of n nodes.
$M_{k,n}^C$	Minimum number of measurements needed to identify k-sparse vectors associated with a complete graph of n nodes.
$f(k,n)$	Number of measurements constructed to identify k-sparse vectors associated with a complete graph of n nodes

Given a graph G with n nodes, let $M_{k,n}^G$ denote the minimum number of non-adaptive measurements needed to identify all k-sparse vectors associated with G. Let $M_{k,n}^C$ denote the minimum number of non-adaptive measurements needed in a complete graph with n nodes. Since any subset of nodes in a complete graph forms a connected subgraph, every 0-1 matrix is a feasible measurement matrix there. Existing results [4], [10], [32] show that with overwhelming probability a random 0-1 matrix with $O(k \log(n/k))$ rows[1] can identify all k-sparse vectors associated with a complete graph, and we can recover the sparse vector by ℓ_1-minimization, which returns the vector with the least ℓ_1-norm[2] among those that can produce the obtained measurements. Then we have

$$M_{k,n}^C = O(k \log(n/k)). \qquad (1)$$

We use (1) for the analysis of construction methods. Explicit constructions of measurement matrices for complete graphs also exist, e.g., [1], [4], [15], [16], [32]. We use $f(k,n)$ to denote the number of measurements to recover k-sparse vectors associated with the complete graph of n nodes by a particular measurement construction method. $f(k,n)$ varies for different construction methods, and clearly $f(k,n) \geq M_{k,n}^C$. The key notations are summarized in Table I.

The questions we would like to address in the paper are:

- Given a known graph G, what is the corresponding $M_{k,n}^G$?
- How can we explicitly design measurements such that the total number of measurements is close to $M_{k,n}^G$?

III. SPARSE RECOVERY OVER SPECIAL GRAPHS

In this section, we consider three kinds of special graphs: one-dimensional line/ring network, ring with each node connecting to four closest neighbors, and a tree. We construct measurements for each graph and later generalize the construction ideas obtained here to general graphs in Section IV.

A. Line and Ring

First consider one-dimensional line/ring network as shown in Fig. 2. When later comparing the results here with those in Section III-B one can see that the number of measurements required to recover sparse vectors can be significantly different

[1]We use the notations $g(n) \in O(h(n))$, $g(n) \in \Omega(h(n))$, or $g(n) = \Theta(h(n))$ if as n goes to infinity, $g(n) \leq ch(n)$, $g(n) \geq ch(n)$ or $c_1 h(n) \leq g(n) \leq c_2 h(n)$ eventually holds for some positive constants c, c_1 and c_2 respectively.
[2]The ℓ_p-norm ($p \geq 1$) of $\mathbf{x}$ is $\|\mathbf{x}\|_p = (\sum_i |x_i|^p)^{1/p}$, and $\|\mathbf{x}\|_\infty = \max_i |x_i|$.
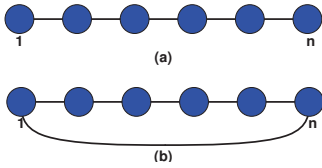
Fig. 2.   (a) line network (b) ring network

in two graphs that only differ from each other with a small number of links.

In a line/ring network, there is not much freedom in the measurement design since only consecutive nodes can be measured together from assumption (A1). In fact, [24], [28] show that $\lceil \frac{n+1}{2} \rceil$ (or $\lceil \frac{n}{2} \rceil$) measurements are both necessary and sufficient to recover 1-sparse vectors associated with a line (or ring) network with n nodes. Therefore, $\Theta(n)$ measurements are required to recover even one non-zero element associated with a line/ring network.

We next construct $k\lceil \frac{n}{k+1} \rceil + 1$ measurements to recover k-sparse vectors ($k \geq 2$) associated with the line/ring network. Let $t = \lceil \frac{n}{k+1} \rceil$. For every $1 \leq i \leq kt+1$, the ith measurement goes through all the nodes from i to $\min(i+t-1, n)$.

Theorem 1. $k\lceil \frac{n}{k+1} \rceil + 1$ *above measurements are sufficient to identify all k-sparse vectors associated with a line/ring network with n nodes.*

Proof: Consider matrix $A^{(tk+1)\times(tk+t)}$ with its ith row having '1's from entry i to entry $i+t-1$ and '0's elsewhere for all $1 \leq i \leq tk+1$. Then the first n columns of A correspond to our measurement matrix. To prove the statement, we only need to show that A can identify all k-sparse vectors in $\mathcal{R}^{tk+t}$, which happens if and only if every non-zero vector $\mathbf{z}$ such that $A\mathbf{z} = \mathbf{0}$ holds has at least $2k+1$ non-zero elements [9].

For each index $1 \leq k' \leq k$, define a submatrix $A_{k'}$, which consists of the first $tk'+1$ rows and the first $tk'+t$ columns of A. We claim that every non-zero vector $\mathbf{w}$ such that $A_{k'}\mathbf{w} = \mathbf{0}$ holds has at least $2k'+1$ non-zero elements with at least two non-zero elements in the last t entries. We prove this claim by induction over k'.

First consider A_1. Note that its first row has '1's from column 1 to t, and its last row has '1's from column $t+1$ to $2t$. Because any two columns of the submatrix A_1 are linearly independent, for any $\mathbf{w} \neq \mathbf{0}$ such that $A_1\mathbf{w} = \mathbf{0}$, $\mathbf{w}$ must have at least three non-zero elements. Let j be the index of the last non-zero element of $\mathbf{w}$. If $j \leq t$, consider the jth row of A_1 with its first '1' entry in the jth column. The inner product of the jth row and $\mathbf{w}$ is non-zero, contradicting the assumption that $A_1\mathbf{w} = \mathbf{0}$. Then $j \geq t+1$ must hold. Then since the inner product between $\mathbf{w}$ and the last row of A_1 is zero, at least two non-zero elements exist in the last t entries of $\mathbf{w}$.

Now suppose the claim holds for $A_{k'}$, consider a non-zero vector $\mathbf{w}$ such that $A_{k'+1}\mathbf{w} = \mathbf{0}$ holds. Note that the vector of the first $tk' + t$ positions of $\mathbf{w}$, denoted by $\hat{\mathbf{w}}$, satisfies $A_{k'}\hat{\mathbf{w}} = \mathbf{0}$. We remark that $\hat{\mathbf{w}} \neq \mathbf{0}$. If $\hat{\mathbf{w}} = \mathbf{0}$, let j denote the index of the first non-zero element of $\mathbf{w}$, and we have

$j \geq tk' + t + 1$. Consider the $(j + 1 - t)$th row of $A_{k'+1}$ with its last '1' entry in column j. Then the inner product of this row with $\mathbf{w}$ is non-zero, which is a contradiction.

Since $\hat{\mathbf{w}} \neq \mathbf{0}$, from the induction assumption, it has at least $2k'+1$ non-zero elements with at least two non-zero elements in its last t elements. Now consider the last $2t$ elements of $\mathbf{w}$ and the last $t + 1$ measurements in $A_{k'+1}$. From a similar argument for the case of A_1, we know that $\mathbf{w}$ must have at least two non-zero elements in the last t positions. So $\mathbf{w}$ has at least $2(k' + 1) + 1$ non-zero elements.

By induction over k', every $\mathbf{w} \neq \mathbf{0}$ satisfying $A\mathbf{w} = 0$ has at least $2k+1$ non-zero entries. This completes the proof. ∎

Theorem 1 implies that we can save about $\lfloor \frac{n}{k+1} \rfloor$ measurements but still be able to recover k-sparse vectors in a line/ring network via compressed sensing. However, for group testing associated with a line/ring network, one can check that n measurements are necessary to recover more than one non-zero element. The key is that every node should be an *endpoint* at least twice, where the endpoints are the nodes at the beginning and the end of a measurement. The endpoints of a measurement can be a same node. If node u is an endpoint for at most once, then it is always measured together with one of its neighbors, say v, if ever measured. Then when v is '1', we cannot determine the value of u, either '1' or '0'. Therefore, to recover more than one non-zero element, we need at least $2n$ endpoints, and thus n measurements.

B. Ring with nodes connecting to four closest neighbors

We know from Section III-A that $\lceil n/2 \rceil$ measurements are necessary to recover even one non-zero element associated with a ring network. Now consider a graph with each node directly connecting to its four closest neighbors as in Fig. 3 (a), denoted by $\mathcal{G}^4$. $\mathcal{G}^4$ is important to the study of small-world networks [30]. $\mathcal{G}^4$ has n more links than the ring network, but we will show that the number of measurements required by compressed sensing to recover k-sparse vectors associated with $\mathcal{G}^4$ significantly reduces from $\Theta(n)$ to $O(k \log(n/k))$.

Definition 1. *Given $G = (V, E)$, $S \subseteq V$ is a **hub** for $U \subseteq V$ if G_S is connected, and $\forall u \in U, \exists s \in S$ s.t. $(u, s) \in E$.*

Clearly in $\mathcal{G}^4$, if nodes are numbered consecutively around the ring, then the set of all the odd nodes, denoted by T_o, form a hub for the set of all the even nodes, denoted by T_e. Given a k-sparse vector $\mathbf{x}$, let $\mathbf{x}_o$ and $\mathbf{x}_e$ denote the subvectors of $\mathbf{x}$ with odd and even indices. Then $\mathbf{x}_o$ and $\mathbf{x}_e$ are both at most k-sparse. The sum of entries in $\mathbf{x}_o$, denoted by s_o, can be obtained by one measurement, and similarly for the sum s_e of the entries of $\mathbf{x}_e$. For any subset W of T_e, $T_o \cup W$ induces a connected subgraph and thus can be measured by one measurement. We can obtain the sum of values corresponding to nodes in W by measuring nodes in $T_o \cup W$ and then subtracting s_o from the sum. For example in Fig. 3 (b) and (c), in order to measure the sum of the pink nodes 2, 8 and 10, we measure the sum of pink nodes and all the black odd nodes, and then subtract s_o from the obtained summation. Though the subgraph induced by T_e is
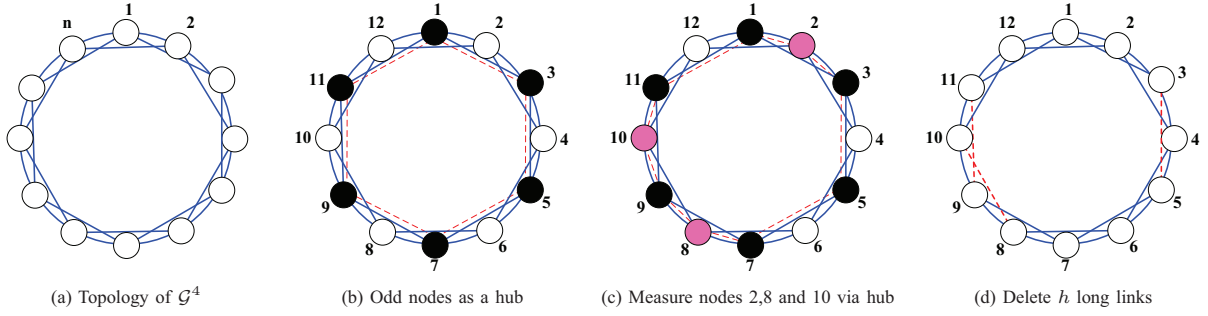
(a) Topology of $\mathcal{G}^4$ (b) Odd nodes as a hub   (c) Measure nodes 2,8 and 10 via hub   (d) Delete h long links

Fig. 3.   Sparse recovery on graph $\mathcal{G}^4$

not complete, we can still freely measure nodes in T_e with the help of the hub T_o. Therefore $M^C_{k,\lfloor n/2 \rfloor} + 1$ measurements are enough to recover $\mathbf{x}_e \in \mathcal{R}^{\lfloor n/2 \rfloor}$, where the additional one measurement measures s_o. Similarly, we can use T_e as a hub to recover the subvector $\mathbf{x}_o \in \mathcal{R}^{\lceil n/2 \rceil}$ with $M^C_{k,\lceil n/2 \rceil} + 1$ measurements, and thus $\mathbf{x}$ is recovered. From above, we have

Theorem 2. *All k-sparse vectors associated with $\mathcal{G}^4$ can be recovered with $M^C_{k,\lfloor n/2 \rfloor} + M^C_{k,\lceil n/2 \rceil} + 2$ measurements, which is $O(2k \log(n/(2k))) + 2$.*

Theorem 2 is important in the following three aspects.

Firstly, from ring network to $\mathcal{G}^4$, although the number of links only increases by n, the number of measurements required to recover k-sparse vectors significantly reduces from $\Theta(n)$ to $O(2k \log(n/(2k))) + 2$. Besides, this value is in the same order as $M^C_{k,n}$, while the number of links in $\mathcal{G}^4$ is only $2n$ compared with $n(n-1)/2$ links in a complete graph.

Secondly, the idea of using a hub to design the measurements is very important for our later results. If set S can serve as a hub for U in graph G, then the induced graph G_U is "almost equivalent" to a complete subgraph in the sense that we can measure any subset of nodes in U freely via S. The number of measurements required to recover k-sparse vectors associated with U is $M^C_{k,|U|} + 1$ with one additional measurement for the hub.

Thirdly, our estimate $O(2k \log(n/(2k))) + 2$ on the minimum number of measurements required to recover k-sparse vectors greatly improves over the existing results in [12], [33], both of which are based on the mixing time of a random walk. The mixing time $T(n)$ is the smallest t' such that a random walk of length t' starting at any node in G ends up having a distribution μ' with $\|\mu - \mu'\|_\infty \leq 1/(2cn)^2$ for some $c \geq 1$, where μ is the stationary distribution over the nodes of a standard random walk over the graph G. [33] proves that $O(kT^2(n) \log n)$ measurements can identify k-sparse vectors with overwhelming probability by compressed sensing. [12] uses $O(k^2 T^2(n) \log(n/k))$ measurements to identify k non-zero elements by group testing. In $\mathcal{G}^4$, one can easily see that $T(n)$ should be at least $n/4$. Then both results provide no saving in the number of measurements for $\mathcal{G}^4$ as the mixing time is $\Theta(n)$.

Besides the explicit measurement construction described before Theorem 2, we can also recover k-sparse vectors with

$O(\log n)$ random measurements with high probability. We need to point out that these random measurements do not depend on the measurements of a complete graph.

Consider an n-step Markov chain $\{X_k, 1 \leq k \leq n\}$ with $X_1 = 1$. For any $k \leq n - 1$, if $X_k = 0$, then $X_{k+1} = 1$; if $X_k = 1$, then X_{k+1} can be 0 or 1 with equal probability. Clearly any realization of this Markov chain does not contain two or more consecutive zeros, and thus is a feasible row of the measurement matrix. Moreover,

Theorem 3. *With high probability all k-sparse vectors associated with $\mathcal{G}^4$ can be recovered with $O(g(k) \log n)$ measurements obtained from the above Markov chain, where $g(k)$ is a function of k.*

Proof: See Appendix.   ∎

Adding n links in the form $(i, i + 2 (\text{mod } n))$ to the ring network greatly reduces the number of measurements needed from $\Theta(n)$ to $O(\log n)$. Then how many links in the form $(i, i + 2(\text{mod } n))$ shall we add to the ring network such that the minimum number of measurements required to recover k-sparse vectors is exactly $\Theta(\log n)$? The answer is $n - \Theta(\log n)$. To see this, let $\mathcal{G}^4_h$ denote the graph obtained by deleting h links in the form $(i, i + 2(\text{mod } n))$ from $\mathcal{G}^4$. For example in Fig. 3 (d), we delete links $(3, 5)$, $(8, 10)$ and $(9, 11)$ in red dashed lines from $\mathcal{G}^4$. Given h, our following results do not depend on the specific choice of links to remove. We have

Theorem 4. *The minimum number of measurements required to recover k-sparse vectors associated with $\mathcal{G}^4_h$ is lower bounded by $\lceil h/2 \rceil$, and upper bounded by $2M^C_{k,\lceil \frac{n}{2} \rceil} + h + 2$.*

Proof: Let D denote the set of nodes such that for every $i \in D$, link $(i - 1, i + 1)$ is removed from $\mathcal{G}^4$. The proof of the lower bound follows the proof of Theorem 2 in [28]. The key idea is that recovering one non-zero element in D is equivalent to recovering one non-zero element in a ring network with h nodes, and thus $\lceil h/2 \rceil$ measurements are necessary.

For the upper bound, we first measure nodes in D separately with h measurements. Let S contain the even nodes in D and all the odd nodes. S can be used as a hub to recover the k-sparse subvectors associated with the even nodes that are not in D, and the number of measurements used is at most $M^C_{k,\lfloor \frac{n}{2} \rfloor} + 1$. We similarly recover k-sparse subvectors associated with odd nodes that are not in D using the set of

the odd nodes in D and all the even nodes as a hub. The number of measurements is at most $M_{k,\lceil\frac{n}{2}\rceil}^C + 1$. Sum them up and the upper bound follows. ∎

Together with (1), Theorem 4 directly implies that if $\Theta(\log n)$ links in the form $(i, i + 2(\text{mod } n))$ are deleted from $\mathcal{G}^4$, then $\Theta(\log n)$ measurements are both necessary and sufficient to recover k-sparse vectors associated with $\mathcal{G}_{\Theta(\log n)}^4$ for any constant k. Moreover, the lower bound in Theorem 4 implies that if the number of links removed is $\Omega(\log n)$, then the number of measurements required for sparse recovery is also $\Omega(\log n)$. Thus, we need to add $n - \Theta(\log n)$ links to a ring network such that the number of measurements required for sparse recovery is exactly $\Theta(\log n)$.

Since the number of measurements required by compressed sensing is greatly reduced when we add n links to the ring network, one may wonder whether the number of measurements needed to locate k non-zero elements by group testing can also be greatly reduced or not. Our next result shows that this is not the case for group testing.

Theorem 5. $\lfloor n/4 \rfloor$ *measurements are necessary to locate two non-zero elements associated with $\mathcal{G}^4$ by group testing.*

Proof: Suppose two non-zero elements are on nodes $2i-1$ and $2i$ for some $1 \le i \le \lfloor\frac{n}{2}\rfloor$. We view nodes $2i-1$ and $2i$ as a group for every i $(1 \le i \le \lfloor\frac{n}{2}\rfloor)$, denoted by B_i. If both nodes in B_j are '1's for some j, then every measurement that passes either node or both nodes in B_i is always '1'. Consider a reduced graph with B_i, $\forall i$ as nodes, and link (B_i, B_j) $(i \ne j)$ exists only if in $\mathcal{G}^4$ there is a path from a node in B_i to a node in B_j without going though any other node not in B_i or B_j. B_i is '1' if both node $2i - 1$ and node $2i$ in $\mathcal{G}^4$ are '1's and is '0' otherwise. The reduced network is a ring with $\lfloor\frac{n}{2}\rfloor$ nodes, and thus $\lfloor n/4 \rfloor$ measurements are required to locate one non-zero element in the reduced network. Then we need at least $\lfloor n/4 \rfloor$ measurements to locate two consecutive non-zero elements associated with $\mathcal{R}^4$, and thus, $\lfloor n/4 \rfloor$ is also a lower bound for locating two general non-zero elements. ∎

By Theorem 2 and Theorem 5, we observe that in $\mathcal{G}^4$, with compressed sensing the number of measurements needed to recover k-sparse vectors is $O(2k \log(n/(2k)))$, while with group testing, $\Theta(n)$ measurements are required if $k \ge 2$.

C. Tree

Next we consider a tree topology as in Fig. 4. For a given tree, the root is treated as the only node in layer 0. The nodes that are t steps away from the root are in layer t. We say the tree has depth h if the farthest node is h steps away from the root. Let n_i denote the number of nodes on layer i, and $n_0 = 1$. We construct measurements to recover vectors associated with a tree by the following *tree approach*.

We recover the nodes layer by layer starting from the root, and recovering nodes in layer i requires that all the nodes above layer i should already be recovered. First measure the root separately. When recovering the subvector associated with nodes in layer i $(2 \le i \le h)$, we can measure the sum of any subset of nodes in layer i using some nodes in the upper layers
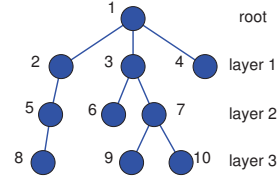

Fig. 4. Tree topology

as hub and then delete the value of the hub from the obtained sum. One simple way to find a hub is to trace back from nodes to be measured on the tree simultaneously until they reach one same node. For example in Fig. 4, in order to measure nodes 5 and 7 together, we will trace back to the root and measure nodes 1, 2, 3, 5, and 7 together and then subtract the values of nodes 1, 2, and 3, which are already identified when we recover nodes in the upper layers. With this approach, we have,

Theorem 6. $\sum_{i=0}^h M_{k,n_i}^C$ *measurements are enough to recover k-sparse vectors associated with a tree with depth h, where n_i is the number of nodes in layer i.*

IV. SPARSE RECOVERY OVER GENERAL GRAPHS

In this section we consider recovering k-sparse vectors over general graphs. The graph is assumed to be connected. If not, we simply treat each component as a connected subgraph and design measurements to recover k-sparse subvectors associated with each subgraph separately.

Inspired by the construction methods in Section III, in Section IV-A we propose a general design guideline based on "r-partition" which will be introduced soon. The key idea is to divide the nodes into a small number of groups such that nodes in the same group are connected to one hub, and thus can be measured freely with the help of the hub. We use the Erdős-Rényi random graph as an example to illustrate the design guideline based on r-partition. Since finding the minimum number of such groups in general turns out to be NP-hard, in Section IV-B we propose a simple algorithm to design a small number of measurements to recover k-sparse vectors associated with any given graph.

A. Measurement Construction Based on r-partition

In $\mathcal{G}^4$, we divide nodes into odd nodes T_o and even nodes T_e and use each set as a hub for the other set. In general graphs, we extend this idea and have the following definition:

Definition 2 (r-partition). *Given $G = (V, E)$, disjoint subsets N_i $(i = 1, ..., r)$ of V form an r-**partition** of G if and only if these two conditions both hold: (1) $\cup_{i=1}^r N_i = V$, and (2) $\forall i$, $V \backslash N_i$ is a hub for N_i.*

Clearly, T_o and T_e form a 2-partition of graph $\mathcal{G}^4$. With the above definition, we have

Theorem 7. *If G has an r-partition N_i $(i = 1, ..., r)$, then the number of measurements needed to recover k-sparse vectors associated with G is at most $\sum_{i=1}^r M_{k,|N_i|}^C + r$, which is $O(rk \log(n/k)) + r$.*

Proof: Note that $M^C_{k,|N_i|} + 1$ measurements (with one additional measurement for $V \backslash N_i$) are enough to recover k-sparse subvector associated with N_i via its hub $V \backslash N_i$. ∎

We remark that Theorem 2, 6, and 7 all result from hub based measurement constructions. We next apply this result to the Erdős-Rényi random graph $G(n, p)$, which contains n nodes and there exists a link between any two nodes independently with probability p. Note that if $p \geq (1 + \epsilon) \log n/n$ for some constant $\epsilon > 0$, $G(n, p)$ is connected almost surely [6].

Theorem 8. *For Erdős-Rényi random graph $G(n, p)$ with $p = \beta \log n/n$, if $\beta \geq 2 + \epsilon$ for some constant $\epsilon > 0$, then any two disjoint subsets N_1 and N_2 of nodes with $|N_1| = |N_2| = n/2$ form a 2-partition with high probability. Moreover, with high probability the number of measurements needed to recover k-sparse vectors associated with $G(n, p)$ is at most $2M^C_{k,n/2} + 2$, which is $O(2k \log(n/(2k))) + 2$.*

Proof: Let N_1 be any subset of V with $|N_1| = n/2$, and let $N_2 = V \backslash N_1$. Then G_{N_1} and G_{N_2} are both Erdős-Rényi random graphs with $n/2$ nodes, and are connected almost surely when $p \geq (2 + \epsilon) \log n/n$.

We claim that with high probability, for every $u \in N_1$, there exists $v \in N_2$ such that $(u, v) \in E$. Let P_1 denote the probability that there exists some $u \in N_1$ such that $(u, v) \notin E$ for every $v \in N_2$. Then

$$P_1 = \sum_{u \in N_1} (1 - p)^{n/2} = \frac{n}{2}(1 - \beta \log n/n)^{n/2}$$

$$= \frac{n}{2}(1 - \frac{\beta \log n}{n})^{\frac{n}{\beta \log n} \cdot \frac{\beta \log n}{2}} \leq \frac{n}{2}e^{-\frac{\beta \log n}{2}} \leq \frac{n^{-\epsilon/2}}{2},$$

where the last inequality holds from $\beta \geq 2 + \epsilon$. Then P_1 goes to zero as n goes to infinity, and the claim follows. Similarly, one can prove that with high probability for every $v \in N_2$, there exists $u \in N_1$ such that $(u, v) \in E$.

Then with high probability N_1 and N_2 form a 2-partition. The second statement follows from Theorem 7. ∎

[12] considers group testing over Erdős-Rényi random graphs and shows that $O(k^2 \log^3 n)$ measurements are enough to identify up to k non-zero entries in an n-dimensional logical vector provided that $p = \Theta(k \log^2 n/n)$. Here with compressed sensing setup and 2-partition results, we can recover k-sparse vectors in $\mathcal{R}^n$ with $O(2k \log(n/(2k))) + 2$ measurements when $p > (2+\epsilon) \log n/n$ for some $\epsilon > 0$. Note that this result also improves over the previous result in [33], which requires $O(k \log^3 n)$ measurements for compressed sensing on $G(n, p)$.

From Theorem 7, the number of measurements used is closely related to the value r. In general, one wants to reduce r so as to reduce the number of measurements. Given graph G and integer r, the question that whether or not G has an r-partition is called *r-partition problem*. In fact,

Theorem 9. $\forall r \geq 3$, *r-partition problem is NP-complete.*

Please refer to Appendix for its proof. We remark that we cannot prove the hardness of the 2-partition problem though we conjecture it is also a hard problem.

Subroutine 1 Leaves(G, u)

Input: graph G, root u
1 Find a spanning tree T of G rooted at u by breadth-first search, and let S denote the set of leaf nodes of T.
2 **Return:** S

Subroutine 2 Reduce(G, u, H)

Input: $G = (V, E)$, H_e for each $e \in E$, and node u
1 $V = V \backslash u$.
2 **for each** two different neighbors v and w of u **do**
3   **if** $(v, w) \notin E$ **then**
4        $E = E \cup (v, w)$, $H_{(v,w)} = H_{(v,u)} \cup H_{(u,w)} \cup \{u\}$.
5   **end if**
6 **end for**
7 **Return:** G, H

B. Measurement Construction Algorithm for General Graphs

Section IV-A proposes the r-partition concept as a measurement design guideline. But finding an r-partition with the smallest r in general is NP-hard. Now given a connected graph G, how shall we efficiently design a small number of measurements to recover k-sparse vectors associated with G?

One simple way is to find the spanning tree of G, and then use the tree approach in Section III-C. The depth of the spanning tree is at least R, where $R = \min_{u \in V} \max_{v \in V} d_{uv}$ is the radius of G with d_{uv} as the length of the shortest path between u and v. This approach only uses links in the spanning tree, and the number of measurements used is large when the radius R is large. For example, the radius of $\mathcal{G}^4$ in Fig. 3 is $n/4$, then the spanning tree approach uses at least $n/4$ measurements, one for each layer. However, the number of measurements can be as small as $O(2k \log(n/2k)) + 2$ if we take advantage of the additional links.

Here we propose a simple algorithm to design the measurements for general graphs. The algorithm combines the ideas of the tree approach and the r-partition. We still divide nodes into a small number of groups such that each group can be identified via some hub. Here nodes in the same group are the leaf nodes of a spanning tree of a gradually reduced graph. A leaf node has no children on the tree.

Let $G^* = (V^*, E^*)$ denote the input graph. The algorithm is built on the following two subroutines. **Leaves**(G, u) returns the set of leaf nodes of a spanning tree of G rooted at u. **Reduce**$(G = (V, E), u, H)$ deletes u from G and fully connects all the neighbors of u. Specifically, for every two neighbors v and w of u, we add a link (v, w), if not already exist, and let $H_{(v,w)} = H_{(v,u)} \cup H_{(u,w)} \cup \{u\}$, where for each link $(s, t) \in E$, $H_{(s,t)}$ denotes the set of nodes, if any, that serves as a hub for s and t in the original graph G^*. We record H such that measurements constructed on a reduced graph G can be feasible in G^*.

Given graph G^*, let u denote the node such that $\max_{v \in V^*} d_{uv} = R$, where R is the radius of G^*. Pick u as the root and obtain a spanning tree T of G^* by breadth-

Algorithm 1 Measurement construction for graph G^*

Input: $G^* = (V^*, E^*)$.

  1 $G = G^*$, $H_e = \varnothing$ for each $e \in E$

  2 **while** $|V| > 1$ **do**

  3   Find the node u such that $\max_{v \in V} d_{uv} = R^G$, where R^G is the radius of G. $S =$ **Leaves**(G, u).

  4   Design $f(k, |S|) + 1$ measurements to recover k-sparse vectors associated with S using nodes in $V \backslash S$ as a hub.

  5   **for each** u in S **do**

  6        $G = $ **Reduce**(G, u, H)

  7   **end for**

  8 **end while**

  9 Measure the last node in V directly.

10 **Output:** All the measurements.

---

first search. Let S denote the set of leaf nodes in T. With $V^* \backslash S$ as a hub, we can design $f(k, |S|) + 1$ measurements to recover up to k-sparse vectors associated with S. We then reduce the network by deleting every u in S and fully connects all the neighbors of u. For the obtained reduced network G, we repeat the above process until all the nodes are deleted. Note that when designing the measurements in a reduced graph G, if a measurement uses link (v, w), then it should also include nodes in $H_{(v,w)}$ so as to be feasible in the original graph G^*.

In each step tree T is rooted at node u where $\max_{v \in V} d_{uv}$ equals the radius of the current graph G. Since all the leaf nodes of T are deleted in the graph reduction procedure, the radius of the new obtained graph should be reduced by at least one. Then we have at most R iterations in Algorithm 1 until only one node is left. Clearly we have,

Theorem 10. *The number of measurements designed by Algorithm 1 is at most $Rf(k, n) + R + 1$, where R is the radius of the graph.*

We remark that the number of measurements by the spanning tree approach we mentioned at the beginning of Section IV-B is also no greater than $Rf(k, n) + R + 1$. However, we expect that Algorithm 1 uses fewer measurements than the spanning tree approach for general graphs, since Algorithm 1 also considers links that are not in the spanning tree. And it is verified in Experiment 1 in Section V.

V. Simulation

Experiment 1 (Effectiveness of Algorithm 1): Given a graph G, we apply Algorithm 1 to divide the nodes into groups such that each group (except the last one) can be measured freely via some hub. The last group only contains one node and can be measured directly. We consider recovering 1-sparse vectors as an example, since in complete graphs, $M_{1,n}^C = \lceil \log(n + 1) \rceil$ can be computed exactly, and from (1) we know that the number of measurements required to recovery k-sparse vectors is within a constant times $kM_{1,n}^C$. Since a matrix with the binary expansion of interger i as column i of the measurement matrix can identify 1-sparse vectors in a complete graph [19], we construct such measurements for each group, and the total

number of measurements to recover 1-sparse vectors in G is $\sum_i^{q-1} \lceil \log(n_i + 1) \rceil + q$, where n_i is the number of nodes in group i and q is the total number of groups. To recover the sparse vectors from measurements, note that for 1-sparse vectors, the corresponding subvector of measurements of each group equals to either the zero vector or some constant times the vector of the binary expansion of some integer i. Thus, one can easily recover the location and the value of the non-zero entry from measurements.

In Fig. 6, we gradually increase the number of links in a graph with $n = 1000$ nodes. We start with a uniformly generated random tree, and in each step randomly add 25 links to the graph. All the results are averaged over one hundred realizations. The number of measurements constructed decreases from 73 to 30 when the number of links increases from $n-1$ to $2n-1$. Note that the number of measurements is already within $3M_{1,n}^C$ when the average node degree is close to 4. The radius R of the graph decreases from 13 to 7, and we also plot the upper bound $R\lceil \log n \rceil + R + 1$ provided by Theorem 10. One can see that the number of measurements actually constructed can be much less than the upper bound.

In Fig. 7, we consider the scale-free network with Barabási-Albert (BA) model [3] where the graph initially has m_0 connected nodes, and each new node connects to m existing nodes with a probability that is proportional to the degree of the existing nodes. We start with a random tree of 10 nodes and increase the total number of nodes from 64 to 1024. Every result is averaged over one hundred realizations. Since the diameter of BA model is $O(\log n / \log(\log(n)))$ [7], then by Theorem 10, the number of our constructed measurments is upper bounded by $O(\log^2 n / \log(\log(n)))$. As the mixing time of BA model is $O(\log n)$ [26], methods in [12] and [33] require $O(\log^3 n)$ random measurements.

Experiment 2 (Sparse Recovery Performance with Noise): Compressed sensing theory indicates that if A is a random 0-1 matrix, with overwhelming probability we can recover the sparse vector $\mathbf{x}_0$ though ℓ_1-minimization [9]. Here we generate a graph with $n = 500$ nodes from BA model. Algorithm 1 divides nodes into four groups with 375, 122, 2 and 1 node respectively. For each of the first two groups with size n_i ($i = 1, 2$), we generate $\lceil n_i / 2 \rceil$ random measurements each measuring a random subset of the group together with its hub. We also measure the two hubs directly. Each of the three nodes in the next two groups is measured directly by one measurement. The generated matrix A is 254 by 500. We generate a sparse vector $\mathbf{x}_0$ with i.i.d. zero-mean Gaussian entries on a randomly chosen support, and normalize $\|\mathbf{x}_0\|_2$ to 1. To recover $\mathbf{x}_0$ from $\mathbf{y} = A\mathbf{x}_0$, one can run ℓ_1-minimization to recover the subvectors associated with the first two groups, and the last three entries of $\mathbf{x}_0$ can be obtained from measurements directly. However, note that every measurement for the first two groups passes through its hub, then any error in a hub measurement will affect every measurement for the group of nodes using this hub. To address this issue, we propose to use a modified ℓ_1-minimization in which the errors in the two
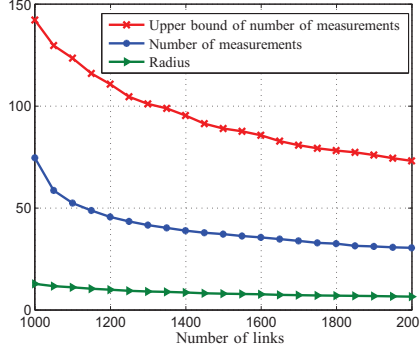
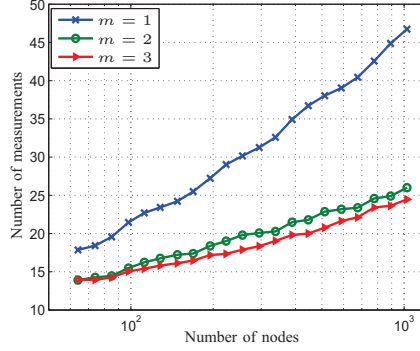Fig. 6. Random graph with $n = 1000$

Fig. 7. BA model with increasing n and different m
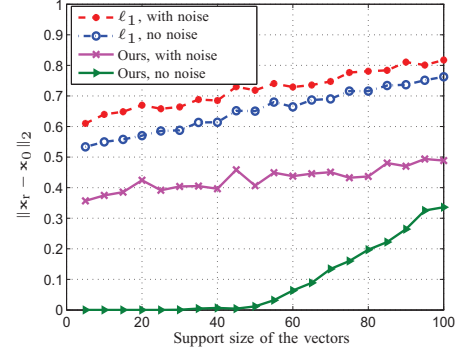

Fig. 8. Recovery performance with hub errors

hubs are treated as entries of an augmented vector to recover. Specifically, let the augmented vector $\mathbf{z} = [\mathbf{x}_0^T, e_1, e_2]^T$ and the augmented matrix $A' = [A\ \boldsymbol{\beta}\ \boldsymbol{\gamma}]$, where e_1 (or e_2) denotes the error in the measurement of the first (second) hub, and the column vector $\boldsymbol{\beta}$ (or $\boldsymbol{\gamma}$) has '1' in the row corresponding to the measurement of the first (or second) hub and '0' elsewhere. We then recover $\mathbf{z}$ (and thus $\mathbf{x}_0$) from $\mathbf{y} = A'\mathbf{z}$ via ℓ_1-minimization on each group. Fig. 8 compares the recovery performance of our modified recovering method and the traditional ℓ_1-minimization, where the hub errors e_1 and e_2 are drawn from a Gaussian distribution with zero mean and unit variance. For every support size k, we randomly generate one hundred k-sparse vectors $\mathbf{x}_0$, and let $\mathbf{x}_r$ denote the recovered vector. Even with the hub errors, the average $\|\mathbf{x}_r - \mathbf{x}_0\|_2$ is within 10^{-6} when $\mathbf{x}_0$ is at most 25-sparse by our method, while by ℓ_1-minimization, the value is at least 0.5. We also consider the case that besides errors in hub measurements, every other measurement has i.i.d. zero-mean Gaussian noise. Let $\mathbf{w}$ denote the noise vector and $\|\mathbf{w}\|_2$ is normalized to 2. The average $\|\mathbf{x}_r - \mathbf{x}_0\|_2$ here is smaller with our method than that with ℓ_1-minimization.

VI. CONCLUSION AND DISCUSSIONS

This paper addresses the sparse recovery problem with graph constraints. By providing explicit measurement constructions for different graphs, we derive upper bounds of the minimum number of measurements needed to recover vectors up to certain sparsity. It would be interesting to explore corresponding tight lower bounds. Further efforts are also needed to empirically evaluate the performance of different recovery themes, especially when the measurements are noisy.

We need to remark that this paper is only the first step towards network measurement constructions with topological constraints, and several practical concerns have not been taken into account yet. For instance, we assume full knowledge of the fixed network topology, and how to construct measurements when the topology is time-varying or partially known is an open question. We also assume that any number of nodes can be measured together as long as they form a connected subgraph, while in practice one may only measure a small number of nodes in one measurement so as to reduce the

overhead and the measurement noise.

APPENDIX

A. Proof of Theorem 3

Let $A^{m \times n}$ denote the matrix with m realizations of the n-step Markov chain. To prove the statement, from [9], we only need to show that the probability that every $2k$ columns of A are linearly independent goes to 1 as n goes to infinity.

Let A_I be a submatrix of A with columns in I, where I is an index set with $|I| = 2k$. Let $A_{S_j I}$ ($1 \leq j \leq \lfloor \frac{m}{2k} \rfloor$) be a submatrix of A_I formed by row $2k(j-1) + 1$ to row $2kj$ of A_I. Let P_d^I denote the probability that $\text{rank}(A_I) < 2k$, and let π_d^I denote the probability that $\text{rank}(A_{S_j I}) < 2k$ for given j. Note that given I, π_d^I is the same for every $A_{S_j I}$, $\forall j$. Note that $\text{rank}(A_I) < 2k$ implies that $\text{rank}(A_{S_j I}) < 2k$ for each such matrix $A_{S_j I}$, then

$$P_d^I \leq (\pi_d^I)^{\lfloor \frac{m}{2k} \rfloor}. \tag{2}$$

To characterize π_d^I, consider matrix $B^{2k \times 2k}$ with $B_{ii} = 0$ for $i = 2, 3, ..., 2k$ and $B_{ij} = 1$ for all the other elements. Since $\text{rank}(B) = 2k$, then

$$\pi_d^I \leq 1 - P(A_{S_j I} \text{ is a row permutation of } B). \tag{3}$$

One can check that in this Markov chain, for every $1 \leq i < k \leq n$, $P(X_k = 1 \mid X_i = 1) \geq 1/2$, $P(X_k = 0 \mid X_i = 1) \geq 1/4$, $P(X_k = 1 \mid X_i = 0) \geq 1/2$, and $P(X_k = 1) \geq 1/2$ by simple calculation. Since B has $(2k)!$ different row permutations, one can calculate that

$$P(A_{S_j I} \text{ is a row permutation of } B) \geq (2k)!/2^{4k^2 + 2k - 1}. \tag{4}$$

Combining (2), (3) and (4), we have

$$P(\text{every } 2k \text{ columns of } A \text{ are linearly independent})$$
$$= 1 - P(\text{rank}(A_I) < 2k \text{ for some } I \text{ with } |I| = 2k)$$
$$\geq 1 - \binom{n}{2k} P_d^I \geq 1 - \binom{n}{2k} e^{-(2k)!(\frac{1}{2})^{4k^2 + 2k - 1} \lfloor \frac{m}{2k} \rfloor}, \tag{5}$$

where the first inequality follows from the union bound. Then if $m = g(k) \log n = (2k + 1)2^{4k^2 + 2k - 1} \log n / (2k - 1)!$, from (5) we have the probability that every $2k$ columns of A are linearly independent is at least $1 - 1/((2k)!n)$. Then the statement follows.

B. Proof of Theorem 9

Since checking whether or not r given sets form an r-partition takes polynomial time, r-partition problem is NP.

We will show the r-partition problem is NP-complete for $r \geq 3$ by proving that the NP-complete r-coloring problem ($r \geq 3$) is polynomial time reducible to the r-partition problem.

Let $G = (V, E)$ and an integer r be an instance of r-coloring. For every $(u, v) \in E$, add a node w and two links (w, u) and (w, v). Let W denote the set of nodes added. Add a link between every pair of nodes in V not already joined by a link. Let H denote the augmented graph and let V' denote the set of nodes in H. We claim that if there exists an r-partition of H, then we can obtain an r-coloring of G, and vice versa.

Suppose S_i $(i = 1, ..., r)$ form an r-partition of H. Note that for every $(u, v) \in E$, u and v cannot belong to the same set S_i for any i. Suppose u and v both belong to S_i for some i. Let w denote the node in W that only directly connects to u and v. If $w \in S_i$, then w has both neighbors in the same set with w, contradicting the definition of r-partition. If $w \notin S_i$, then $H_{V' \setminus S_i}$ is disconnected since w does not connect to any node in $V' \setminus S_i$. It also contradicts the definition of r-partition. Thus, for every $(u, v) \in E$, node u and v belong to two sets S_i and S_j with $i \neq j$. Then we obtain an r-coloring of G.

Let $C_i \subset V$ $(i = 1, ..., r)$ denote an r-coloring of G. We claim that $N_i = C_i$ $(i = 1, ..., r-1)$, and $N_r = C_r \cup W$ form an r-partition of H. First note $\forall u \in V$, at least one of its neighbors is not in the same set as u, since H_V is a complete subgraph. $\forall w \in W$, w is directly connected to u and v with $(u, v) \in E$. From the definition of r-coloring, u and v are in different sets C_i and C_j for some $i \neq j$. Therefore, w has at least one neighbor that is not in N_r. Second, we will show $H_{V' \setminus N_i}$ is connected for all i. $H_{V' \setminus N_r}$ is a complete subgraph, and thus connected. $\forall i < r$, let $S_i := V \setminus C_i$, then $V' \setminus N_i = S_i \cup W$. H_{S_i} is a complete subgraph, and thus connected. $\forall w \in W$, since its two neighbors cannot be both in C_i, then at least one neighbor belongs to S_i, thus $H_{V' \setminus N_r} = H_{S_i \cup W}$ is connected. N_i $(i = 1, ..., r)$ indeed forms an r-partition.

REFERENCES

[1] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Applied and Computational Harmonic Analysis*, vol. 26, no. 2, pp. 283 – 290, 2009.

[2] P. Babarczi, J. Tapolcai, and P.-H. Ho, "Adjacent link failure localization with monitoring trails in all-optical mesh networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 907 –920, 2011.

[3] A. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[4] R. Berinde, A. Gilbert, P. Indyk, H. Karloff, and M. Strauss., "Combining geometry and combinatorics: a unified approach to sparse signal recovery," *arxiv:0804.4666*, 2008.

[5] T. Blumensath, "Compressed sensing with nonlinear observations," Tech. Rep., 2010.

[6] B. Bollobas, *Random Graphs*, 2nd ed. Cambridge University Press, 2001.

[7] B. Bollobás and O. Riordan, "The diameter of a scale-free random graph," *Combinatorica*, vol. 24, pp. 5–34, 2004.

[8] T. Bu, N. Duffield, F. L. Presti, and D. Towsley, "Network tomography on general topologies," in *Proc ACM SIGMETRICS*, 2002, pp. 21–30.

[9] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[10] ——, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[11] Y. Chen, D. Bindel, H. H. Song, and R. Katz, "Algebra-based scalable overlay network monitoring: Algorithms, evaluation, and applications," *IEEE/ACM Trans. Netw.*, vol. 15, no. 5, pp. 1084 –1097, 2007.

[12] M. Cheraghchi, A. Karbasi, S. Mohajer, and V. Saligrama, "Graph-constrained group testing," *arXiv:1001.1445*, 2010.

[13] A. Coates, A. Hero III, R. Nowak, and B. Yu, "Internet tomography," *IEEE Signal Processing Magazine*, vol. 19, no. 3, pp. 47 –65, 2002.

[14] M. Coates, Y. Pointurier, and M. Rabbat, "Compressed network monitoring for ip and all-optical networks," in *Proc. ACM SIGCOMM IMC*, 2007, pp. 241–252.

[15] G. Cormode and S. Muthukrishnan, "Combinatorial algorithms for compressed sensing," ser. Lecture Notes in Computer Science, 2006, vol. 4056, pp. 280–294.

[16] R. DeVore, "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, no. 4-6, pp. 918 – 925, 2007.

[17] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[18] D. Donoho and J. Tanner, "Sparse nonnegative solution of underdetermined linear equations by linear programming," in *Proc. Natl. Acad. Sci. U.S.A.*, vol. 102, no. 27, 2005, pp. 9446–9451.

[19] R. Dorfman, "The detection of defective members of large populations," *Ann. Math. Statist.*, vol. 14, pp. 436–440, 1943.

[20] D.-Z. Du and F. K. Hwang, *Combinatorial Group Testing and Its Applications (Applied Mathematics)*, 2nd ed. World Scientific Publishing Company, 2000.

[21] N. Duffield, "Network tomography of binary network performance characteristics," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5373 –5388, 2006.

[22] M. Firooz and S. Roy, "Link delay estimation via expander graphs," *arxiv:1106.0941*, 2011.

[23] A. Gopalan and S. Ramasubramanian, "On identifying additive link metrics using linearly independent cycles and paths," 2011. [Online]. Available: http://www2.engr.arizona.edu/~srini/papers/tomography.pdf

[24] N. Harvey, M. Patrascu, Y. Wen, S. Yekhanin, and V. Chan, "Non-adaptive fault diagnosis for all-optical networks via combinatorial group testing on graphs," in *Proc. IEEE INFOCOM*, 2007, pp. 697 –705.

[25] J. Haupt, W. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 92 –101, 2008.

[26] M. Mihail, C. Papadimitriou, and A. Saberi, "On certain connectivity properties of the internet topology," *Journal of Computer and System Sciences*, vol. 72, no. 2, pp. 239–251, 2006.

[27] H. X. Nguyen and P. Thiran, "Using end-to-end data to infer lossy links in sensor networks," in *Proc. IEEE INFOCOM*, 2006, pp. 1 –12.

[28] J. Tapolcai, B. Wu, P.-H. Ho, and L. Rónyai, "A novel approach for failure localization in all-optical mesh networks," *IEEE/ACM Trans. Netw.*, vol. 19, pp. 275–285, 2011.

[29] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Towards a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, no. 99, pp. 1–14, 2011.

[30] D. Watts and S. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, 1998.

[31] B. Wu, P.-H. Ho, J. Tapolcai, and X. Jiang, "A novel framework of fast and unambiguous link failure localization via monitoring trails," in *Proc. IEEE INFOCOM*, 2010, pp. 1 –5.

[32] W. Xu and B. Hassibi, "Efficient compressive sensing with deterministic guarantees using expander graphs," in *Proc. IEEE ITW*, 2007, pp. 414 –419.

[33] W. Xu, E. Mallada, and A. Tang, "Compressive sensing over graphs," in *Proc. IEEE INFOCOM*, 2011.

[34] Y. Zhao, Y. Chen, and D. Bindel, "Towards unbiased end-to-end network diagnosis," in *Proc. SIGCOMM*, 2006, pp. 219–230.