Sparse Recovery With Graph Constraints

Meng Wang, *Member, IEEE*, Weiyu Xu, Enrique Mallada, *Student Member, IEEE*,
and Ao Tang, *Senior Member, IEEE*

Abstract—Sparse recovery can recover sparse signals from a set of underdetermined linear measurements. Motivated by the need to monitor the key characteristics of large-scale networks from a limited number of measurements, this paper addresses the problem of recovering sparse signals in the presence of network topological constraints. Unlike conventional sparse recovery where a measurement can contain any subset of the unknown variables, we use a graph to characterize the topological constraints and allow an additive measurement over nodes (unknown variables) only if they induce a connected subgraph. We provide explicit measurement constructions for several special graphs, and the number of measurements by our construction is less than that needed by existing random constructions. Moreover, our construction for a line network is provably optimal in the sense that it requires the minimum number of measurements. A measurement construction algorithm for general graphs is also proposed and evaluated. For any given graph G with n nodes, we derive bounds of the minimum number of measurements needed to recover any k-sparse vector over G ($M_{k,n}^G$). Using the Erdős–Rényi random graph as an example, we characterize the dependence of $M_{k,n}^G$ on the graph structure. This paper suggests that $M_{k,n}^G$ may serve as a graph connectivity metric.

Index Terms—Sparse recovery, compressed sensing, topological graph constraints, measurement construction.

I. INTRODUCTION

IN THE monitoring of engineering networks, one often needs to extract network state parameters from indirect observations. Since measuring each component (e.g., router) in the communication network directly can be operationally costly, if feasible at all, the goal of *network tomography* [10], [11], [15], [19], [27], [31], [32], [38], [46] is to infer system internal characteristics such as link bandwidth utilizations and link queueing delays from indirect aggregate measurements.

M. Wang is with the Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: wangm7@rpi.edu).

W. Xu is with the University of Iowa, Iowa, IA 52242 USA (e-mail: weiyu-xu@uiowa.edu).

E. Mallada is with the California Institute of Technology, Pasadena, CA 91125 USA (e-mail: em464@cornell.edu).

A. Tang is with Cornell University, Ithaca, NY 14850 USA (e-mail: atang@ece.cornell.edu).

In many cases, it is desirable to reduce the number of measurements without sacrificing the monitoring performance. For example, when different paths experience the same delay on the same link, network kriging [17] can recover delays on all n links in the network from only n linearly independent paths and thus, identify the delays on possibly exponential number of paths. Moreover, the number of path delay measurements needed to recover n link delays can be further reduced by exploiting the fact that only a small number of bottleneck links experience large delays, while the delay is approximately zero elsewhere. *Sparse Recovery* theory promises that if the signal of interest is sparse, i.e., its most entries are zero, m measurements are sufficient to correctly recover the signal, even though m is much smaller than the signal dimension. Since many network parameters are sparse, e.g., link delays, these network tomography problems can be formulated as a sparse recovery problem with the goal of minimizing the number of indirect observations.

Sparse recovery has two different but closely related problem formulations. One is *Compressed Sensing* [6], [12], [13], [23], [24], 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* [25], [26], where the high-dimensional signal is binary and a measurement is a logical disjunction (**OR**) on the corresponding binary values.

One key question in sparse recovery 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 [12], [23], which is not realistic in network monitoring problems where only objects that form a path or a cycle on the graph [1], [32], or induce a connected subgraph can be aggregated together in the same measurement. Only a few recent works consider graph topological constraints, either in group testing [16] setup, especially motivated by link failure localization in all-optical networks [3], [16], [34], [39], [43], or in compressed sensing setup, with applications in estimation of network parameters [20], [35], [44].

We design measurements for recovering sparse signals in the presence of graph topological constraints, and characterize the minimum number of measurements required to recover sparse signals when the possible measurements should satisfy graph constraints. Though motivated by network applications, graph constraints abstractly model 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. Our construction for line networks is optimal in the sense that it requires the minimum number of measurements. For other special graphs, the number of measurements by our construction is less than the existing estimates (see [16], [44]) of the measurement requirement. (Section III)

(2) For general graphs, we propose a measurement design guideline based on *r-partition*, and further propose a simple measurement design algorithm. (Section IV)

(3) Using Erdős-Rényi random graphs as an example, we characterize the dependence of the number of measurements for sparse recovery on the graph structure. (Section V)

Moreover, we also propose measurement construction methods under additional practical constraints such that the length of a measurement is bounded, or each measurement should pass one of a fixed set of nodes. The issue of measurement error is also addressed. (Sections VI, VII)

II. MODEL AND PROBLEM FORMULATION

We use a graph $G = (V, E)$ to represent the topological constraints, where V denotes the set of nodes with cardinality $|V| = n$, and E denotes the set of edges. Each node i is associated with a real number x_i, and we say vector $\mathbf{x} = (x_i, i = 1, \ldots, n)$ is associated with G. $\mathbf{x}$ is the unknown signal to recover. We say $\mathbf{x}$ is a k-sparse vector if $\|\mathbf{x}\|_0 = k$,[1] i.e., the number of non-zero entries of $\mathbf{x}$ is k.

Let $S \subseteq V$ denote a subset of nodes in G. Let E_S denote the subset of edges with both ends in S, then $G_S = (S, E_S)$ is the induced subgraph of G. We have the following two assumptions on graph topological constraints:

(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.

Given a unknown vector $\mathbf{x}$ associated with G, we take m measurements ($m \ll n$) that satisfy (A1) and (A2). Let vector $\mathbf{y} \in \mathcal{R}^m$ denote m measurements. Let A denote the $m \times n$ measurement matrix with $A_{ij} = 1$ ($i = 1, \ldots, m$, $j = 1, \ldots, n$) if and only if node j is included in the ith measurement and $A_{ij} = 0$ otherwise. We can write it in the compact form that $\mathbf{y} = A\mathbf{x}$. With the requirements (A1) and (A2), A must be a 0-1 matrix, and for each row of A, the set of nodes that correspond to '1' must form a connected induced subgraph of G. For the graph in Fig. 1, we can measure the sum of nodes in S_1 and S_2 by two separate measurements, and the measurement matrix is

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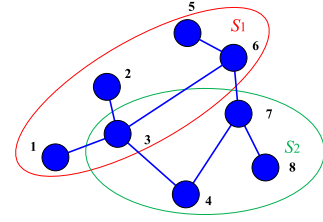

Fig. 1. Graph example.

(A1) and (A2) represent an abstraction of topological constraints. One motivation is the monitoring of the link characteristics in a communication network. (A2) follows from the additive property of many network characteristics,[2] e.g., delays and packet loss rates [32]. If we use graph $\hat{G}$ to represent the communication network where nodes in $\hat{G}$ represent routers and edges represent transmission links, then the graph model G considered in this paper is the line graph [33] (also known as interchange graph or edge graph) $L(\hat{G})$ of graph $\hat{G}$. According to the definition of a line graph, every node in $G = L(\hat{G})$ corresponds to a link in network $\hat{G}$, and the node value corresponds to the link delay. Two nodes in G are connected with an edge if and only if the corresponding links in network $\hat{G}$ are connected to the same router. See Fig. 2 (a) (b) as an example of a network $\hat{G}$ and its line graph G considered in this paper.

A connected subgraph G_S in G corresponds to a set of connected links in the communication network $\hat{G}$. From the Eulerian property there exists a cycle that traverses each link in the connected set of links exactly twice.[3] One router in this cycle sends a packet along the cycle and measures the total transmission delay, which is twice the sum of link delays on this set of links. For example, Fig. 2 shows the correspondence between assumptions (A1) (A2) in the line graph model G and the monitoring of the original network $\hat{G}$. Since large delays only occur at a small number of bottleneck links, the link delays in a network can be represented by a sparse vector $\mathbf{x}$ associated with G.

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 k-sparse vector $\mathbf{x}$ is the unique solution to the following ℓ_0-minimization problem

$$\min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad A\mathbf{z} = A\mathbf{x}. \tag{1}$$

Note (1) is a combinatorial problem in general.

Then, given topological constraints represented by G, we want to design non-adaptive measurements satisfying (A1) and (A2) such that one can identify all k-sparse vector $\mathbf{x}$, and the total number of measurements is minimized. Given a graph G with n nodes, let $M_{k,n}^G$ denote the minimum number of measurements satisfying (A1) and (A2) to identify all

---

[1] The ℓ_p-norm ($p \geq 1$) of $\mathbf{x}$ is $\|\mathbf{x}\|_p = (\sum_i |x_i|^p)^{1/p}$, $\|\mathbf{x}\|_\infty = \max_i |x_i|$, and $\|\mathbf{x}\|_0 = |\{i : x_i \neq 0\}|$.

[2] Compressed sensing can also be applied to cases where (A2) does not hold, e.g., the measurements can be nonlinear as in [7] and [40].

[3] The argument is as follows. Suppose we replace each link with two copies of itself. The resulting multigraph is Eulerian since it is connected and each node has even degree. Then there exists an Eulerian walk which visits every link in the multigraph exactly once. That corresponds to visiting each link of the original communication network twice.
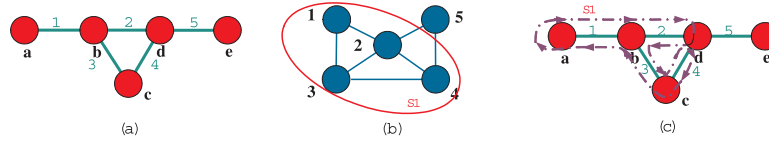
Fig. 2. (a) Network $\hat{G}$ with five links. (b) Its corresponding line graph $L(\hat{G})$ that we consider in this paper. Since the links 1, 2, 3, and 4 are connected in $\hat{G}$, the induced subgraph of nodes 1, 2, 3, and 4 in $L(\hat{G})$ is connected. (c) There exists a cycle passing each of links 1, 2, 3, and 4 in network $\hat{G}$ exactly twice.

k-sparse vectors associated with G. The questions we would like to address in the paper are:

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

Though motivated by network applications, we use graph G to characterize the topological constraints and study a general problem of recovering sparse signals from measurements satisfying graph constraints. For the majority of this paper, we assume a measurement is feasible as long as (A1) and (A2) are satisfied, and we attempt to minimize the total number of measurements for identifying sparse signals. Some additional constraints on the measurements such as bounded measurement length will be discussed in Section VI.

If G is a complete graph, then any subset of nodes forms a connected subgraph, and every 0-1 matrix is a feasible measurement matrix. Then the problem reduces to the conventional compressed sensing where one wants to identify sparse signals from linear measurements. Existing results [5], [6], [13], [37], [45] show that with overwhelming probability a random 0-1 A matrix with $O(k \log(n/k))$ rows[4] can identify all k-sparse vectors $\mathbf{x}$ associated with a complete graph, and $\mathbf{x}$ is the unique solution to the ℓ_1-minimization problem

$$\min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{s.t.} \quad A\mathbf{z} = A\mathbf{x}. \tag{2}$$

(2) can be recast as a linear program, and thus it is computationally more efficient to solve (2) than (1). Thus, we have

$$M_{k,n}^C = O(k \log(n/k)). \tag{3}$$

Note that $O(k \log(n/k)) \ll n$ for $k \ll n$, thus, the number of measurements can be significantly reduced for sparse signals. Explicit constructions of measurement matrices for complete graphs also exist, see [2], [6], [21], [22], [45]. We use $f(k, n)$ to denote the number of measurements to recover k-sparse vectors associated with a complete graph of n nodes by a particular measurement construction method. $f(k, n)$ varies for different construction methods, and clearly $f(k, n) \geq M_{k,n}^C$. Table I summarizes the key notations.

For a general graph G that is not complete, existing results do not hold any more. Can we still achieve a significant reduction in the number of measurements? This is the focus of this paper. 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

<hr>

[4]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.

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
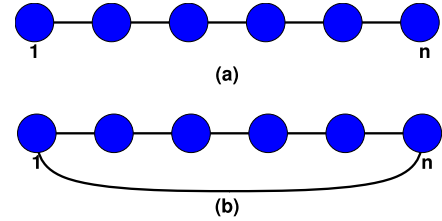
Fig. 3. (a) Line. (b) Ring.

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". Here we consider compressed sensing if not otherwise specified, and the main results are stated in theorems. We only discuss group testing for comparison (e.g., Proposition 2). Note that for recovering 1-sparse vectors, the numbers of measurements required by compressed sensing and group testing are the same.

III. SPARSE RECOVERY OVER SPECIAL GRAPHS

In this section, we consider four kinds of special graphs: one-dimensional line/ring, ring with each node connecting to its four closest neighbors, two-dimensional grid and a tree. The measurement construction method for a line/ring is different from those for the other graphs, and our construction is optimal (or near optimal) for a line (or ring) in terms of reducing the number of required measurements. For other special graphs, we construct measurements based on the "hub" idea and will later extend it to general graphs in Section IV.

A. Line and Ring

First consider a line/ring as shown in Fig. 3. Note that a line/ring is the line graph of a line/ring network. When later comparing the results here with those in Section III-B, one can see that the number of measurements required for sparse recovery can be significantly different in two graphs that only differ from each other with a small number of edges.

In a line/ring, only consecutive nodes can be measured together from (A1). Recovering 1-sparse vectors associated

with a line (or ring) with n nodes is considered in [34] and [39], which shows that $\lceil \frac{n+1}{2} \rceil$ (or $\lceil \frac{n}{2} \rceil$) measurements are both necessary and sufficient in this case. Here, we consider recovering k-sparse vectors for $k \geq 2$.

Our construction works as follows. Given k and n, let $t = \lfloor \frac{n+1}{k+1} \rfloor$. We construct $n + 1 - \lfloor \frac{n+1}{k+1} \rfloor$ measurements with the ith measurement passing all the nodes from i to $i + t - 1$. Let $A^{(n+1-t) \times n}$ be the measurement matrix, then its ith row has '1's from entry i to entry $i + t - 1$ and '0's elsewhere. For example, when $k = 3$ and $n = 11$, we have $t = 3$, and

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}. \tag{4}$$

Let $M_{k,n}^L$ and $M_{k,n}^R$ denote the minimum number of measurements required to recover k-sparse vectors in a line/ring respectively. We have the following results regarding the bounds of $M_{k,n}^L$ and $M_{k,n}^R$.

Theorem 1 (Upper Bound): Our constructed $n + 1 - \lfloor \frac{n+1}{k+1} \rfloor$ measurements can identify all k-sparse vectors associated with a line/ring with n nodes, and the sparse signals can be recovered from ℓ_1-minimization (2).

Theorem 2 (Lower Bound): The required number of measurements to recover k-sparse vectors associated with a line (or ring) with n nodes has the following lower bounds:

$$M_{k,n}^L \geq n + 1 - \lfloor \frac{n+1}{k+1} \rfloor, \tag{5}$$

and

$$M_{k,n}^R \geq n - \lfloor \frac{n}{k+1} \rfloor. \tag{6}$$

Combining Theorem 1 and 2, one can conclude that

$$M_{k,n}^L = n + 1 - \lfloor \frac{n+1}{k+1} \rfloor,$$

and

$$n - \lfloor \frac{n}{k+1} \rfloor \leq M_{k,n}^R \leq n + 1 - \lfloor \frac{n+1}{k+1} \rfloor.$$

Therefore, our construction is optimal for a line in the sense that the number of measurements $n + 1 - \lfloor \frac{n+1}{k+1} \rfloor$ by our construction is the minimum number of measurements needed to recover k-sparse vectors among all the possible measurement constructions. In a ring network, the number of measurements by our construction differs from the minimum requirement of measurements by at most one.

In our early work ([41, Th. 1]), we proved that our construction of $n + 1 - \lfloor \frac{n+1}{k+1} \rfloor$ measurements can identify k-sparse signals, and the signal can be recovered via solving ℓ_0-minimization (1). (1) is in general computationally inefficient to solve. Here we further demonstrate through Theorem 1 that with these $n + 1 - \lfloor \frac{n+1}{k+1} \rfloor$ measurements,

one can recover the signal by solving a computationally efficient ℓ_1-minimization (2).

Proof (of Theorem 1): Let A be the measurement matrix. When $t = 1$, A is the identity matrix, and the statement holds trivially. So we only consider the case $t \geq 2$. It is well known in compressed sensing (see [28]) that a k-sparse vector $\mathbf{x}$ can be recovered from ℓ_1-minimization, i.e., it is the unique solution to (2), if and only if for every vector $\mathbf{w} \neq \mathbf{0}$ such that $A\mathbf{w} = \mathbf{0}$, and for every set $T \subseteq \{1, \ldots, n\}$ with $|T| \leq k$, it holds that

$$\|\mathbf{w}_T\|_1 < \|\mathbf{w}\|_1/2, \tag{7}$$

where $\mathbf{w}_T$ is a subvector of $\mathbf{w}$ with entry indices in T. Thus, we only need to prove that (7) holds for our constructed A.

From the construction of A, one can check that for every $\mathbf{w} \neq \mathbf{0}$ such that $A\mathbf{w} = \mathbf{0}$, and for every $j \in \{1, \ldots, n\}$,

$$w_j = w_{j - \lfloor \frac{j}{t} \rfloor t} \tag{8}$$

holds. For example,

$$w_1 = w_{t+1} = w_{2t+1} = \cdots = w_{(k-1)t+1}.$$

Let $w^* := \arg \max_{j=1}^{t} |w_j|$. From (8), it also holds that

$$w^* = \arg \max_{j=1}^{n} |w_j|. \tag{9}$$

From the first row of A, we have

$$\sum_{i=1}^{t} w_i = 0, \tag{10}$$

From the definition of w^*, (10) implies

$$w^* \leq \frac{1}{2} \sum_{j=1}^{t} |w_j|. \tag{11}$$

Since $n \geq kt + t - 1$ from the definition of t, we have

$$\sum_{j=kt+1}^{n} |w_j| \geq \sum_{j=kt+1}^{kt+t-1} |w_j| = \sum_{j=1}^{t-1} |w_j| > 0, \tag{12}$$

where the equality follows from (8). The last inequality holds since $w_j \neq 0$ for at least one j in $1, \ldots, t-1$. Suppose $w_j = 0$ for all $j = 1, \ldots, t - 1$, then $w_t = 0$ from (10), which then leads to $\mathbf{w} = \mathbf{0}$ through (8), contradicting the fact that $\mathbf{w} \neq \mathbf{0}$.

Now consider any T with $|T| \leq k$, combining (8), (9), (11), and (12), we have

$$\|\mathbf{w}_T\|_1 \leq kw^* \leq \frac{k}{2} \sum_{j=1}^{t} |w_j| = \frac{1}{2} \sum_{j=1}^{kt} |w_j| < \|\mathbf{w}\|_1/2.$$

Thus, $\mathbf{x}$ can be correctly recovered via ℓ_1-minimization (2). $\square$

Proof (of Theorem 2): First, notice that if a measurement matrix $A^{m \times n}$ can be used to recover all k-sparse signals, then $A\mathbf{x}^1 \neq A\mathbf{x}^2$ should hold for every two different k-sparse signals $\mathbf{x}^1$ and $\mathbf{x}^2$. Otherwise, no method would be able to differentiate $\mathbf{x}^1$ and $\mathbf{x}^2$ from the observations $\mathbf{y} = A\mathbf{x}^1 = A\mathbf{x}^2$. This is equivalent to the requirement that every $2k$ columns of A must be linearly independent. We will prove that m should be at least $n + 1 - \lfloor \frac{n+1}{k+1} \rfloor$ for a line or ring network for this requirement to be satisfied.

Let $\bar{A}^{m \times n}$ denote a measurement matrix with which one can recover k-sparse vectors associated with a line of n nodes. Let $\boldsymbol{\alpha}^i$ denote the ith column of $\bar{A}$. Define $\boldsymbol{\beta}^1 = \boldsymbol{\alpha}^1$, $\boldsymbol{\beta}^i = \boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{i-1}$ for all $2 \le i \le n$, and $\boldsymbol{\beta}^{n+1} = -\boldsymbol{\alpha}^n$. Define matrix $P^{m \times (n+1)} = (\boldsymbol{\beta}^i, 1 \le i \le n+1)$. Since $\bar{A}$ is a measurement matrix for a line network, each row of P contains one '1' entry and one '−1' entry, and all the other entries must be '0's.

Given P, we construct a graph G_{eq} with $n+1$ nodes as follows. For every row i of P, there is an edge (j, k) in G_{eq}, where $P_{ij} = 1$ and $P_{ik} = -1$. Then G_{eq} contains m edges, and P can be viewed as the transpose of an oriented incidence matrix of G_{eq}. Let S denote the set of indices of nodes in a component of G_{eq}, then one can check that

$$\sum_{i \in S} \boldsymbol{\beta}^i = \mathbf{0}. \tag{13}$$

Since every $2k$ columns of $\bar{A}$ are linearly independent, every k columns of P are linearly independent, which then implies that the sum of any k columns of P is not a zero vector. With (13), we know that any component of G_{eq} should have at least $k+1$ nodes. Since a component with r nodes contains at least $r-1$ edges, and G_{eq} has at most $\lfloor \frac{n+1}{k+1} \rfloor$ components, then G_{eq} contains at least $n+1-\lfloor \frac{n+1}{k+1} \rfloor$ edges. The (5) follows.

We next consider the ring. Let $\tilde{A}$ denote the measurement matrix with which one can recover k-sparse vectors on a ring with n nodes. Let $\tilde{\alpha}^i$ denote the ith column of $\tilde{A}$. Define $\tilde{\beta}^1 = \tilde{\alpha}^1 - \tilde{\alpha}^n$, and $\tilde{\beta}^i = \tilde{\alpha}^i - \tilde{\alpha}^{i-1}$ for all $2 \le i \le n$. Define matrix $\tilde{P}^{m \times n} = (\tilde{\beta}^i, 1 \le i \le n)$. Similarly, we construct a graph $\tilde{G}_{eq}$ with n nodes based on $\tilde{P}$, and each component of $\tilde{G}_{eq}$ should have at least $k+1$ nodes. Thus, $\tilde{G}_{eq}$ contains at most $\lfloor \frac{n}{k+1} \rfloor$ components and therefore at least $n - \lfloor \frac{n}{k+1} \rfloor$ edges. Then (6) follows. $\qquad \square$

We can save about $\lfloor \frac{n+1}{k+1} \rfloor - 1$ measurements but still be able to recover k-sparse vectors in a line/ring via compressed sensing. But for group testing on a line/ring, n measurements are necessary to recover more than one non-zero element. The arguments use the ideas in [34] and [39], and we skip the details. The key point is that every node should be the *endpoint* at least twice, where the endpoints are the nodes at the beginning and the end of a measurement. If node u is an endpoint for at most once, then it means that either it is always measured together with one of its neighbors, say v, or it is never measured at all. Then when v is '1', we cannot determine the value of u, either '1' or '0' in group testing. Therefore, in order 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

Consider a graph with each node directly connecting to its four closest neighbors as in Fig. 4(a), denoted by $\mathcal{G}^4$. $\mathcal{G}^4$ has $2n$ edges, while a ring network contains n edges. $\mathcal{G}^4$ is also the starting point for constructing small-world networks in the Watts-Strogatz model [42]. We will show that the number of measurements required by compressed sensing to recover k-sparse vectors associated with $\mathcal{G}^4$ is $O(k \log(n/k))$, which is a significant reduction from the required number $\Theta(n)$ in a ring network. We next describe our main idea in the measurement
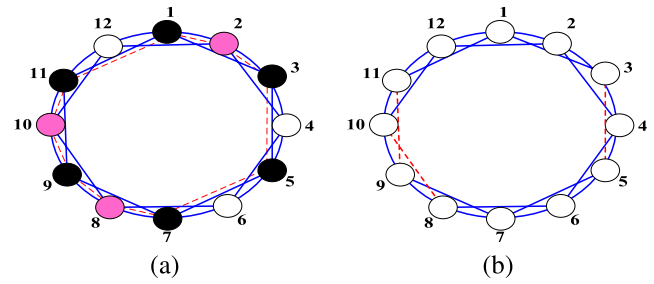

Fig. 4.   Sparse recovery on graph $\mathcal{G}^4$. (a) Measure nodes 2, 8 and 10 via hub T_o, which is the set of all odd nodes. (b) Delete h long links.
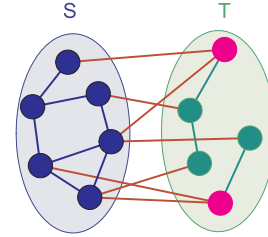

Fig. 5.   Hub S for T.

construction methods. We refer to it as "the use of a hub" in this paper.

1) The Use of a Hub:

*Definition 1: Given $G = (V, E)$ and two disjoint[5] sets $S, T \subseteq V$, we say S is a **hub**[6] for T if G_S is connected, and $\forall u \in T$, $\exists s \in S$ s.t. $(u, s) \in E$.*

If S is hub for T, We first take one measurement of the sum of nodes in S, denoted by s. Then any subset W of T, e.g., the pink nodes in Fig. 5, $S \cup W$ induces a connected subgraph from the hub definition and thus can be measured by one measurement. To measure the sum of nodes in W, we first measure nodes in $S \cup W$ and then subtract s from the sum. Therefore we can apply the measurement constructions for complete graphs on T with this simple modification, and that requires only one additional measurement for the hub S. Thus,

Theorem 3: If there exists a hub S for set T, $M_{k,|T|}^C + 1$ measurements are enough to recover k-sparse vectors associated with T.

The significance of Theorem 3 is that G_T is not necessarily a complete subgraph, i.e., a clique, and it can even be disconnected. As long as there exists a hub S, the measurement construction for a complete graph with the same number of nodes can be applied to T with simple modification. Our later results rely heavily on Theorem 3.

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

---

[5]We assume without loss of generality that S and T are disjoint. When $S \cap T \ne \varnothing$, we consider the set $T' = T \backslash S$, then S and T' are disjoint.

[6]The definition of a hub is closely related to but different from the definition of a connected dominating set in graph theory. S is a connected dominating set for a graph $G = (V, E)$, if and only if G_S is connected, and $\forall u \in V \backslash S$, $\exists s \in S$ s.t. $(u, s) \in E$. The distinction is that we only require T to be a subset of $V \backslash S$, while for S to be a connected dominating set, T must equal to $V \backslash S$.

even indices. Then $\mathbf{x}_o$ and $\mathbf{x}_e$ are both at most k-sparse. From Theorem 3, $M^C_{k,\lfloor n/2 \rfloor} + 1$ measurements are enough to recover $\mathbf{x}_e \in \mathcal{R}^{\lfloor n/2 \rfloor}$. 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.

Corollary 1: All k-sparse vectors associated with $\mathcal{G}^4$ can be recovered with $M^C_{k,\lfloor n/2 \rfloor} + M^C_{k,\lceil n/2 \rceil} + 2$ (which is $O(2k \log(n/(2k)))$) measurements.

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

2) Random Constructions: Besides the explicit measurement construction based on the hub idea, we can also recover k-sparse vectors associated with $\mathcal{G}^4$ from $O(\log n)$ random measurements. We need to point out that these random measurements do not depend on the measurement constructions for a complete graph.

Consider an n-step Markov chain $\{X_k, 1 \le k \le n\}$ with $X_1 = 1$. For any $k \le n-1$, 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. We have the following result, please refer to Appendix-A for its proof.

Theorem 4: With probability at least $1 - 1/((2k)!n)$, 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

$$g(k) = \frac{(2k+1)2^{4k^2+2k-1}}{(2k-1)!}.$$

Adding n edges in the form $(i, i+2(\text{mod } n))$ to the ring greatly reduces the number of measurements needed from $\Theta(n)$ to $O(\log n)$. Then how many edges in the form $(i, i+2(\text{mod } n))$ shall we add to the ring 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 edges in the form $(i, i+2(\text{mod } n))$ from $\mathcal{G}^4$. For example in Fig. 4(b), we delete edges $(3, 5)$, $(8, 10)$ and $(9, 11)$ in red dashed lines from $\mathcal{G}^4$. Given h, our following results do not depend on the specific choice of edges to remove. We have

Theorem 5: 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 $i \in D$, edge $(i-1, i+1)$ is removed from $\mathcal{G}^4$. The proof of the lower bound follows [39, Proof of Th. 2]. The key idea is that recovering one non-zero element in D is equivalent to recovering one non-zero element in a ring 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
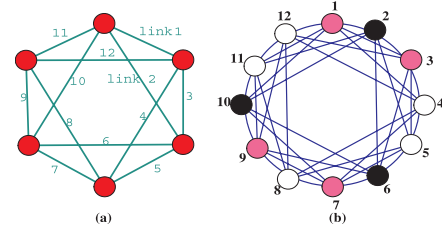

Fig. 6. (a) A communication network with $n = 12$ links, (b) the line graph of the network in (a). Measure the sum of any subset of odd nodes (e.g., 1, 3, 7, and 9) using nodes 2, 6, and 10 as a hub.

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 $M^C_{k,\lceil \frac{n}{2} \rceil} + 1$. Sum them up and the upper bound follows. $\square$

Together with (3), Theorem 5 implies that if $\Theta(\log n)$ edges in the form $(i, i+2(\text{mod } n))$ are deleted from $\mathcal{G}^4$, then $\Theta(\log n)$ measurements are necessary and sufficient to recover associated k-sparse vectors for constant k.

Since the number of measurements required by compressed sensing is greatly reduced when we add n edges to a ring, one may wonder whether the number of measurements needed by group testing can be greatly reduced or not. Our next result shows that this is not the case for group testing, please refer Appendix-B for its proof.

Proposition 1: $\lfloor n/4 \rfloor$ measurements are necessary to locate two non-zero elements associated with $\mathcal{G}^4$ by group testing.

By Corollary 1 and Proposition 1, 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. Line Graph of a Ring Network With Each Router Connecting to Four Routers

Here we compare our construction methods with those in [16] and [44] on recovering link quantities in a network with each router connecting to four closest routers in the ring. Fig. 6(a)[7] shows such a network with n links with $n = 12$. As discussed in Section II, we analyze the line graph of the communication network in Fig. 6(a). In its line graph in Fig. 6(b), node i (representing the delay on link i in Fig. 6(a)) is connected to nodes $i-3$, $i-2$, $i-1$, $i+1$, $i+2$, and $i+3$ (all mod n) for all odd i; and node i is connected to nodes $i-4$, $i-3$, $i-1$, $i+1$, $i+3$, and $i+4$ (all mod n) for all even i.

With the hub idea, we can recover k-sparse link delays in this network from $O(2k \log(n/(2k)))$ measurements. Specifically, we use the set of all the odd nodes as a hub to recover the values associated with the even nodes, and it takes $O(k \log(n/(2k)))$ measurements. We then use the set

---

[7]Fig. 6(a) is a communication network with nodes representing routers and edges representing links, while Fig. 4(a) is a graph model capturing topological constraints with nodes representing the quantities to recover.
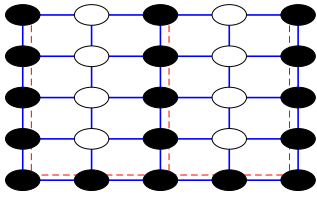
Fig. 7.   Two-dimensional grid.

Fig. 8.   Tree topology.

of nodes $\{4j + 2, j = 0, \ldots, \lceil \frac{n-2}{4} \rceil\}$ as a hub to recover the values associated with the odd nodes, and it takes another $O(k \log(n/(2k)))$ measurements. See Fig. 6(b) as an example.

Our construction of $O(2k \log(n/(2k)))$ measurements to recover k-sparse link delays in the network in Fig. 6(a) greatly improves over the existing results in [16] and [44], which are based on the mixing time of a random walk. The mixing time $T(n)$ can be roughly interpreted as the minimum length of a random walk on a graph such that its distribution is close to its stationary distribution. Xu *et al.* [44] proved that $O(kT^2(n) \log n)$ measurements can identify k-sparse vectors with overwhelming probability by compressed sensing. Chen *et al.* [16] showed that $O(k^2T^2(n) \log(n/k))$ measurements are enough to identify k non-zero elements in the group testing setup. Since $T(n)$ is at least $n/8$ for the network in Fig. 6(a), the methods in [16] and [44] provide no saving in the number of measurements for this network, while our construction reduces this number to $O(2k \log(n/(2k)))$.

D. Two-Dimensional Grid

Next we consider the two-dimensional grid, denoted by $\mathcal{G}^{2d}$. $\mathcal{G}^{2d}$ has $\sqrt{n}$ rows and $\sqrt{n}$ columns. We assume $\sqrt{n}$ to be even here, and also skip '$\lceil \cdot \rceil$' and '$\lfloor \cdot \rfloor$' for notational simplicity.

The idea of measurement construction is still the use of a hub. First, Let S_1 contain the nodes in the first row and all the nodes in the odd columns, i.e., the black nodes in Fig. 7. Then S_1 can be used as a hub to measure k-sparse subvectors associated with nodes in $V \backslash S_1$. The number of measurements is $M^C_{k,(n/2 - \sqrt{n}/2)} + 1$. Then let S_2 contain the nodes in the first row and all the nodes in the even columns, and use S_2 as a hub to recover up to k-sparse subvectors associated with nodes in $V \backslash S_2$. Then number of measurements required is also $M^C_{k,(n/2 - \sqrt{n}/2)} + 1$. Finally, use nodes in the second row as a hub to recover sparse subvectors associated with nodes in the first row. Since nodes in the second row are already identified in the above two steps, then we do not need to measure the hub separately in this step. The number of measurements here is $M^C_{k,\sqrt{n}}$. Therefore,

Proposition 2: With $2M^C_{k,n/2 - \sqrt{n}/2} + M^C_{k,\sqrt{n}} + 2$ measurements one can recover k-sparse vectors associated with $\mathcal{G}^{2d}$.

E. Tree

Next we consider a tree topology as in Fig. 8. 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
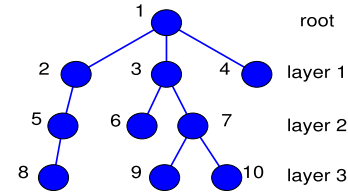
root. Let n_i denote the number of nodes on layer i, and $n_0 = 1$. We construct measurements to recover vectors associated with a tree by the following *tree approach*.

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 as a 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. 8, to measure the sum of nodes 5 and 7, we trace back to the root and measure the sum of nodes 1, 2, 3, 5, and 7 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,

Proposition 3: $\sum_{i=0}^{h} M^C_{k,n_i}$ 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 associated with general graphs. The graph is assumed to be connected. If not, we design measurements to recover k-sparse subvectors associated with each component separately.

In Section IV-A we propose a general design guideline based on "r-partition". The key idea is to divide the nodes into a small number of groups such that each group can be measured with the help of a hub. Since finding the minimum number of such groups turns out to be NP-hard in general, in Section IV-B we propose a simple algorithm to design measurements on any given graph.

A. Measurement Construction Based on r-Partition

*Definition 2 (r-Partition): Given $G = (V, E)$, disjoint subsets N_i ($i = 1, \ldots, r$) 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 Definition 2 and Theorem 3, we have

Theorem 6: If G has an r-partition N_i ($i = 1, \ldots, r$), then the number of measurements needed to recover k-sparse vectors associated with G is at most $\sum_{i=1}^{r} M^C_{k,|N_i|} + r$, which is $O(rk \log(n/k))$.

Another example of the existence of an r-partition is the Erdős-Rényi random graph $G(n, p)$ with $p > \log n/n$.

Subroutine 1 Leaves(G, u)

Initial: 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

The number of our constructed measurements on $G(n, p)$ is less than the existing estimates in [16] and [44]. Please refer to Section V for the detailed discussion.

Clearly, if an r-partition exists, the number of measurements also depends on 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 7: $\forall r \geq 3$, *r-partition problem is NP-complete.*

Please refer to Appendix-C 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.

Although finding an r-partition with the smallest r in general is NP-hard, it still provides a guideline that one can reduce the number of measurements by constructing a small number of hubs such that all the nodes are connected to at least one hub. Our measurement constructions for some special graphs in Section III are also based on this guideline. We next provide efficient measurement design methods for a general graph G based on this guideline.

B. Measurement Construction Algorithm for General Graphs

One simple way is to find the spanning tree of G and use the tree approach in Section III-E. 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 edges in the spanning tree, and the number of measurements needed is large when the radius R is large. For example, the radius of $\mathcal{G}^4$ is $n/4$, then the tree approach uses at least $n/4$ measurements, while $O(2k \log(n/2k))$ measurements are already enough if we take advantage of the additional edges not in the spanning tree.

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 original 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, u, K) deletes u from G and connects every pair of neighbors of u. Specifically, for every two neighbors v and w of u, we add an edge (v, w), if not already exists, and let $K_{(v,w)} = K_{(v,u)} \cup K_{(u,w)} \cup \{u\}$. $K_{(s,t)}$ denotes the set of nodes that connects s and t in the original graph G^*. We record K so as to translate the measurements on a reduced graph G to feasible measurements on G^*.

Subroutine 2 Reduce(G, u, K)

Initial: $G = (V, E)$, 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     $E = E \cup (v, w)$, $K_{(v,w)} = K_{(v,u)} \cup K_{(u,w)} \cup \{u\}$.
5   **end if**
6 **end for**
7 **return** G, K

Algorithm 1 Measurement Construction for Graph G^*

Initial: $G^* = (V^*, E^*)$.
1 $G = G^*$, $K_e = \varnothing$ 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** v in S **do**
6     $G = $**Reduce**($G$, v, K)
7   **end for**
8 **end while**
9 Measure the last node in V directly.
10 **Output:** All the measurements.

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-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 node v in S and connecting every pair of neighbors of v. For the 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 passes edge (v, w), then it should also include nodes in $K_{(v,w)}$ 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 8: 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 is also no greater than $Rf(k, n) + R + 1$. However, since Algorithm 1 also considers edges that are not in the spanning tree, we expect that for general graphs, it uses fewer measurements than the spanning tree approach. This is verified in Experiment 1 in Section VIII.

Algorithm 1 has at most R iterations. In each iteration, the time complexity of constructing measurements (step 4) for selected nodes depends on the specific choice of measurement

construction methods. For instance, 0-1 measurement matrices corresponding to expander graphs [6] for recovering n-dimensional sparse signals can be constructed in time poly(n) [14]. The time complexity of remaining steps in each iteration of Algorithm 1 is $O(|V|^2 \log |V| + |V||E|)$, where V and E denote the set of nodes and the set of edges respectively in the reduced graph G at the beginning of each iteration. The bound $O(|V|^2 \log |V| + |V||E|)$ comes from the computation of the radius of a graph, where we employ Johnson's algorithm [36] with time complexity $O(|V|^2 \log |V| + |V||E|)$ to compute all pairs of shortest paths.

V. Sparse Recover Over Random Graphs

Here we consider measurement constructions over the Erdős-Rényi random graph $G(n, p)$, which has n nodes and every two nodes are connected by a edge independently with probability p. The behavior of $G(n, p)$ changes significantly when p varies. We study the dependence of number of measurements needed for sparse recovery on p.

A. $np = \beta \log n$ for Some Constant $\beta > 1$

Now $G(n, p)$ is connected almost surely [29]. Moreover, we have the following lemma regarding the existence of an r-partition.

Lemma 1: When $p = \beta \log n/n$ for some constant $\beta > 1$, with probability at least $1 - O(n^{-\alpha})$ for some $\alpha > 0$, every set S of nodes with size $|S| = n/(\beta - \epsilon)$ for any $\epsilon \in (0, \beta - 1)$ forms a hub for the complementary set $T = V \backslash S$, which implies that $G(n, p)$ has a $\lceil \frac{\beta - \epsilon}{\beta - \epsilon - 1} \rceil$-partition.

Proof: Note that the subgraph G_S is also Erdős-Rényi random graph in $G(n/(\beta - \epsilon), p)$. Since $p = \beta \log n/n > \log(n/(\beta - \epsilon))/(n/(\beta - \epsilon))$, G_S is connected almost surely.

Let P_f denote the probability that there exists some $u \in T$ such that $(u, v) \notin E$ for every $v \in S$. Then

$$P_f = \sum_{u \in T}(1 - p)^{|S|} = (1 - \frac{1}{\beta - \epsilon})n(1 - \beta \log n/n)^{n/(\beta - \epsilon)}$$

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

$$\leq (1 - \frac{1}{\beta - \epsilon})n e^{-\frac{\beta \log n}{\beta - \epsilon}} \leq (1 - \frac{1}{\beta - \epsilon})n^{-\epsilon/(\beta - \epsilon)}.$$

Thus, S is a hub for T with probability at least $1 - O(n^{-\alpha})$ for $\alpha = \epsilon/(\beta - \epsilon) > 0$. Since the size of T is $(1 - 1/(\beta - \epsilon))n$, $G(n, p)$ has at most $\lceil \frac{\beta - \epsilon}{\beta - \epsilon - 1} \rceil$ such disjoint sets. Then by a simple union bound, one can conclude that $G(n, p)$ has a $\lceil \frac{\beta - \epsilon}{\beta - \epsilon - 1} \rceil$-partition with probability at least $1 - O(n^{-\alpha})$. $\square$

For example, when $\beta > 2$, Lemma 1 implies that any two disjoint sets N_1 and N_2 with $|N_1| = |N_2| = n/2$ form a 2-partition of $G(n, p)$ with probability $1 - O(n^{-\alpha})$. From Theorem 6 and Lemma 1, and let $\epsilon \to 0$, we have

Proposition 4: When $p = \beta \log n/n$ for some constant $\beta > 1$, all k-sparse vectors associated with $G(n, p)$ can be identified with $O(\lceil \frac{\beta}{\beta - 1} \rceil k \log(n/k))$ measurements with probability at least $1 - O(n^{-\alpha})$ for some $\alpha > 0$.

[16] 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 if it further holds that $p = \Theta(k \log^2 n/n)$. Here with compressed sensing setup and r-partition results, we can recover k-sparse vectors in $\mathcal{R}^n$ with $O(k \log(n/k))$ measurements when $p > \log n/n$. This result also improves over the previous result in [44], which requires $O(k \log^3 n)$ measurements for compressed sensing on $G(n, p)$.

B. $np - \log n \to +\infty$, and $\frac{np - \log n}{\log n} \to 0$

Roughly speaking, p is just large enough to guarantee that $G(n, p)$ is connected almost surely [29]. The diameter $D = \max_{u, v} d_{uv}$ of a connected graph is the greatest distance between any pair of nodes, and here it is concentrated around $\frac{\log n}{\log \log n}$ almost surely [8]. We design measurements on $G(n, p)$ with Algorithm 1. With Theorem 8 and the fact that the radius R is no greater than the diameter D by definition, we have

Proposition 5: When $np - \log n \to +\infty$, and $\frac{np - \log n}{\log n} \to 0$, $O(k \log n \log(n/k)/\log \log n)$ measurements can identify k-sparse vectors associated with $G(n, p)$ almost surely.

C. $1 < c = np < \log n$

Now $G(n, p)$ is disconnected and has a unique giant component containing $(\alpha + o(1))n$ nodes almost surely with α satisfying $e^{-c\alpha} = 1 - \alpha$, or equivalently,

$$\alpha = 1 - \frac{1}{c} \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(ce^{-c})^k,$$

and all the other nodes belong to small components. The expectation of the total number of components in $G(n, p)$ is $(1 - \alpha - c(1 - \alpha)^2/2 + o(1))n$ [29]. Since it is necessary to take at least one measurement for each component, $(1 - \alpha - c(1 - \alpha)^2/2 + o(1))n$ is an expected lower bound of measurements required to identify sparse vectors.

The diameter D of a disconnected graph is defined to be the largest distance between any pair of nodes that belong to the same component. Since D is now $\Theta(\log n/\log(np))$ almost surely [18], then for the radius R of the giant component, $R \leq D = O(\log n/\log(np))$, where the second equality holds almost surely. We use Algorithm 1 to design measurements on the giant component, and then measure every node in the small components directly. Thus, k-sparse vectors associated with $G(n, p)$ can be identified almost surely with $O(k \log n \log(n/k)/\log(np)) + (1 - \alpha + o(1))n$ measurements.

Note that here almost surely the size of every small component is at most $\frac{\log n + 2\sqrt{\log n}}{np - 1 - \log(np)}$ ([18, Lemma 5]). If $k = \Omega(\log n)$, almost surely $(1 - \alpha + o(1))n$ measurements are necessary to identify subvectors associated with small components, and thus necessary for identifying k-sparse vectors associated with $G(n, p)$. Combining the arguments, we have

Proposition 6: When $1 < c = np < \log n$ with constant c, we can identify k-sparse vectors associated with $G(n, p)$ almost surely with $O(k \log n \log(n/k)/\log(np)) + (1 - \alpha + o(1))n$ measurements. $(1 - \alpha - c(1 - \alpha)^2/2 + o(1))n$ is an expected lower bound of the number of measurements needed. Moreover, if $k = \Omega(\log n)$, almost surely $(1 - \alpha + o(1))n$ measurements are necessary to identify k-sparse vectors.

D. $np < 1$

Since the expectation of the total number of components in $G(n, p)$ with $np < 1$ is $n - pn^2/2 + O(1)$ [29], then $n - pn^2/2 + O(1)$ is an expected lower bound of the number of measurements required. Since almost surely all components are of size $O(\log n)$, then we need to take n measurements when $k = \Omega(\log n)$. Therefore,

Proposition 7: When $np < 1$, we need at least $n - pn^2/2 + O(1)$ measurements to identify k-sparse vectors associated with $G(n, p)$ in expectation. Moreover, when $k = \Omega(\log n)$, n measurements are necessary almost surely.

VI. ADDING ADDITIONAL GRAPH CONSTRAINTS

We addressed (A1) and (A2) for topological constraints in previous sections. Here we consider additional graph constraints brought by practical implementations. We first consider measurement construction with length constraints, since measurements with short length are preferred in practice. We then discuss the scenario that each measurement should pass at least one node in a fixed subset of nodes, since in network applications, one may want to reduce the number of routers that initiate the measurements.

A. Measurements With Short Length

We have not imposed any constraint on the maximum number of nodes in one measurement. In practice, one may want to take short measurements so as to reduce the communication cost and the measurement noise. We consider sparse recovery with additional constraint on measurement length on two special graphs.

1) Line and Ring: The construction in Section III-A is optimal for a line in terms of reducing the number of measurements needed. The length of each constructed measurement is $\lfloor \frac{n+1}{k+1} \rfloor$, which is proportional to n when k is a constant. Here we provide a different construction such that the total number of measurements needed to recover associated k-sparse vectors is $k\lceil \frac{n}{k+1} \rceil + 1$, and each measurement measures at most $k + 2$ nodes. We also remark that the number of measurements by this construction is within the minimum plus $\max(k - 1, 1)$ for a line, and within the minimum plus k for a ring.

We construct the measurements as follows. Given k, let B^k be a $k + 1$ by $k + 1$ square matrix with entries of '1' on the main diagonal and the first row, i.e., $B^k_{ii} = 1$ and $B^k_{1i} = 1$ for all i. If k is even, let $B^k_{i(i-1)} = 1$ for all $2 \leq i \leq k + 1$; if k is odd, let $B^k_{i(i-1)} = 1$ for all $2 \leq i \leq k$. $B^k_{ij} = 0$ elsewhere. Let $t = \lceil \frac{n}{k+1} \rceil$, we construct a $(kt + 1)$ by $(k+1)t$ matrix A based on B^k. Given set $S \subseteq \{1, \ldots, kt + 1\}$ and set $T \subseteq \{1, \ldots, (k+1)t\}$, A_{ST} is the submatrix of A with row indices in S and column indices in T. For all $i = 1, \ldots, t$, let $S_i = \{(i - 1)k + 1, \ldots, ik + 1\}$, and let $T_i = \{(k+1)(i - 1) + 1, \ldots, (k+1)i\}$. Define $A_{S_i T_i} = B^k$ for all i. All the other entries of A are zeros. We keep the first n columns of A as a measurement matrix for the line/ring with n nodes. Note that the last one or several rows of the reduced matrix can be all zeros. We just delete these rows, and let the resulting matrix

be the measurement matrix. For example, when $k = 2$ and $n = 9$, we have $t = 3$, and

$$B^2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

and

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \quad (14)$$

When $k = 3$, and $n = 8$, we have $t = 2$ and

$$B^3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Each measurement measures at most $k + 2$ nodes when k is even and at most $k + 1$ nodes when k is odd. We have,

Theorem 9: The above construction can recover k-sparse vectors associated with a line/ring with at most $k\lceil \frac{n}{k+1} \rceil + 1$ measurements, which is within the minimum number of measurements needed plus k. Each constructed measurement measures at most $k + 2$ nodes.

Proof: We only need to prove that all k-sparse vectors in $\mathcal{R}^{(k+1)t}$ can be identified with A, which happens if and only if for every vector $\mathbf{z} \neq \mathbf{0}$ such that $A\mathbf{z} = \mathbf{0}$, $\mathbf{z}$ has at least $2k + 1$ non-zero elements [12].

If $t = 1$, A a $k + 1$ by $k + 1$ full rank matrix, and the claim holds trivially. We next consider $t \geq 2$. We prove the case when k is even and skip the similar proof for odd k.

For each integer t' in $[2, t]$, define a submatrix $A_{t'}$ formed by the first $kt' + 1$ rows and the first $(k + 1)t'$ columns of A. For example, for A in (14), we define

$$A_2 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad \text{and} \quad A_3 = A.$$

We will prove by induction on t' that (*) *every non-zero vector $\mathbf{z} \in \mathcal{R}^{(k+1)t'}$ such that $A_{t'}\mathbf{z} = \mathbf{0}$ holds has at least $2k + 1$ non-zero elements for every t' in $[2, t]$.*

First consider A_2, which is a $(2k + 1) \times (2k + 2)$ matrix. Let $\mathbf{a}_i^T$ ($i = 1, \ldots, 2k + 1$) denote the ith row of A_2. Let $\mathbf{z}$ be any vector such that $A\mathbf{z} = \mathbf{0}$ holds. Then from the construction of A_2 and the property that $A\mathbf{z} = \mathbf{0}$, we have

$$\mathbf{a}_i^T \mathbf{z} = z_i + z_{i+1} = 0, \quad \forall i = k + 2, \ldots, 2k + 1. \quad (15)$$

The equations in (15) indicate that $z_{k+2}, \ldots, z_{2k+2}$, the last $k + 1$ entries of $\mathbf{z}$, are either all zeros or all non-zeros. If $z_{k+2}, \ldots, z_{2k+2}$ are all zeros, let $\mathbf{z}'$ denote the subvector containing the first $k + 1$ entries of $\mathbf{z}$. Since the submatrix that contains the first $k + 1$ rows and the first $k + 1$ columns of A is exactly B^k according to the construction of A_2, then we have

$B^k \mathbf{z}' = \mathbf{0}$. Since B^k is full rank, then $\mathbf{z}' = \mathbf{0}$, which implies that $\mathbf{z} = \mathbf{0}$.

Now consider the case that $z_{k+2}, \ldots, z_{2k+2}$ are all non-zero. Note that from (15), we have

$$z_i + z_{i+1} = 0, \quad \forall i = k+3, k+5, \ldots, 2k+1. \quad (16)$$

Summing up the equations in (16) leads to

$$\sum_{i=k+3}^{2k+2} z_i = 0. \quad (17)$$

From the $k+1$th row of A and $A\mathbf{z} = \mathbf{0}$, we have

$$\mathbf{a}_{k+1}^T \mathbf{z} = \sum_{i=k}^{2k+2} z_i = 0. \quad (18)$$

Combining (17) and (18), we know that

$$z_k + z_{k+1} + z_{k+2} = 0. \quad (19)$$

Since z_{k+2} is non-zero, then $z_k + z_{k+1}$ is also non-zero from (19). From the first row of A and $A\mathbf{z} = \mathbf{0}$, we have

$$\sum_{i=1}^{k+1} z_i = 0. \quad (20)$$

Combining (20) with the fact that $z_k + z_{k+1}$ is non-zero, we know that at least one of the first $k-1$ entries of $\mathbf{z}$ is non-zero. From rows two to $k-1$ of A and $A\mathbf{z} = \mathbf{0}$, we have

$$z_{i-1} + z_i = 0, \quad \forall i = 2, 3, \ldots, k-1. \quad (21)$$

Since one of the first $k-1$ entries of $\mathbf{z}$ is non-zero, one can see from equations (21) that all the first $k-1$ entries are non-zero. Therefore, all the first $k-1$ entries and the last $k+1$ entries of $\mathbf{z}$ are non-zero, and at least one of z_k and z_{k+1} is nonzero. Thus, $\mathbf{z}$ has at least $2k+1$ non-zero entries. Thus, (*) holds for A_2.

Now suppose (*) holds for some t' in $[2, t-1]$. We next need to prove that it holds for matrix $A_{t'+1}$. With the same arguments as those for A_2, one can show that for every $\mathbf{z} \neq \mathbf{0}$ such that $A_{t'+1}\mathbf{z} = \mathbf{0}$, its last $k+1$ entries are either all zeros or all non-zero. If the last $k+1$ entries of $\mathbf{z}$ are all zeros, let $\mathbf{z}'$ denote the subvector containing the first $(k+1)t'$ entries of $\mathbf{z}$. By induction hypothesis, $\mathbf{z}'$ has at least $2k+1$ non-zero entries, thus so does $\mathbf{z}$. Then the claim holds for $A_{t'+1}$.

If the last $k+1$ entries of $\mathbf{z}$ are all non-zero, then from the equation that

$$\mathbf{a}_{kt'+1}^T \mathbf{z} = \sum_{(k+1)t'-1}^{(k+1)(t'+1)} z_i = 0,$$

we know that $z_{(k+1)t'-1} + z_{(k+1)t'}$ is non-zero. We next claim that there exists an integer j in $[0, t'-1]$ such that the sum of all $k-1$ entries from $z_{j(k+1)+1}$ to $z_{j(k+1)+k-1}$ is non-zero. Suppose not, then from

$$\mathbf{a}_1^T \mathbf{z} = \sum_{i=1}^{k+1} z_i = 0, \quad (22)$$

and the hypothesis that the sum of entries from z_1 to z_{k-1} is zero, one can obtain that

$$z_k + z_{k+1} = 0. \quad (23)$$

Then, for all $r = 1, \ldots, t'-1$, consider the equations

$$0 = \mathbf{a}_{rk+1}^T \mathbf{z} = \sum_{i=r(k+1)-1}^{(r+1)(k+1)} z_i \quad (24)$$

$$= (z_{r(k+1)-1} + z_{r(k+1)}) + (z_{(r+1)(k+1)-1} + z_{(r+1)(k+1)}), \quad (25)$$

where the last equality follows from the hypothesis that the sum of entries from $z_{r(k+1)+1}$ to $z_{r(k+1)+k-1}$ is zero. Combining (23) and the equations (24), one can obtain that $z_{(k+1)t'-1} + z_{(k+1)t'} = 0$, which contradicts the fact that $z_{(k+1)t'-1} + z_{(k+1)t'}$ is non-zero. Therefore, there exists j in $[0, t'-1]$ such that the sum of all $k-1$ entries from $z_{j(k+1)+1}$ to $z_{j(k+1)+k-1}$ is non-zero. Then, from the equations

$$\mathbf{a}_i^T \mathbf{z} = z_{i+j-1} + z_{i+j} = 0, \quad \forall i = jk+2, \ldots, jk+k-1, \quad (26)$$

we know that if the sum of $k-1$ terms from $z_{j(k+1)+1}$ to $z_{j(k+1)+k-1}$ is non-zero, then each of these entries is non-zero. Since we obtained earlier that at least one of $z_{(k+1)t'-1}$ and $z_{(k+1)t'}$ is non-zero, and the last $k+1$ entries of $\mathbf{z}$ are all non-zero, we conclude that in this case $\mathbf{z}$ also has at least $2k+1$ non-zero entries.

By induction over t', every $\mathbf{z} \neq \mathbf{0}$ such that $A\mathbf{z} = \mathbf{0}$ has at least $2k+1$ non-zero entries, then the result follows. $\square$

This construction measures at most $k+2$ nodes in each measurement. If measurements with constant length are preferred, we provide another construction method such that every measurement only measures at most three nodes. This method requires $(2k-1)\lceil \frac{n}{2k} \rceil + 1$ measurements to recover k-sparse vectors associated with a line/ring.

Given k, let D^k be a $2k$ by $2k$ square matrix having entries of '1' on the main diagonal and the subdiagonal and '0' elsewhere, i.e. $D_{ii}^k = 1$ for all i and $D_{i(i-1)}^k = 1$ for all $i \geq 2$, and $D_{ij}^k = 0$ elsewhere. Let $t = \lceil \frac{n}{2k} \rceil$, we construct a $(2kt - t + 1)$ by $2kt$ matrix A based on D^k. Let $S_i = \{(i-1)(2k-1)+1, \ldots, i(2k-1)+1\}$, and let $T_i = \{2k(i-1)+1, \ldots, 2ki\}$. Define $A_{S_i T_i} = D^k$ for all $i = 1, \ldots, t$, and $A_{ij} = 0$ elsewhere. We keep the first n columns of A as the measurement matrix. For example, when $k = 2$ and $n = 8$, we have

$$D^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

and

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \quad (27)$$

Theorem 10: The above constructed $(2k - 1)\lceil \frac{n}{2k}\rceil + 1$ *measurements can identify k-sparse vectors associated with a line/ring of n nodes, and each measurement measures at most three nodes.*

Proof: When $t = 1$, A is a full rank square matrix. We focus on the case that $t \geq 2$. For each integer t' in $[2, t]$, define a submatrix $A_{t'}$ formed by the first $2kt' - t' + 1$ rows and the first $2kt'$ columns of A. We will prove by induction on t' that every $\mathbf{z} \neq \mathbf{0}$ such that $A_{t'}\mathbf{z} = \mathbf{0}$ holds has at least $2k + 1$ non-zero elements for every t' in $[2, t]$.

First consider A_2. For A in (27), $A_2 = A$. From the first $2k - 1$ rows of A_2, one can check that for every $\mathbf{z}$ such that $A_2\mathbf{z} = \mathbf{0}$, its first $2k - 1$ entries are zeros. From the $2k$th row of A_2, we know that z_{2k} and z_{2k+1} are either both zeros or both non-zero. In the former case, the remaining $2k-1$ entries of $\mathbf{z}$ must be zeros, thus, $\mathbf{z} = \mathbf{0}$. In the latter case, one can check that the remaining $2k - 1$ entries are all non-zero, and therefore $\mathbf{z}$ has $2k + 1$ non-zero entries.

Now suppose the claim holds for some t' in $[2, t - 1]$. Consider vector $\mathbf{z} \neq \mathbf{0}$ such that $A_{t'+1}\mathbf{z} = \mathbf{0}$. If $z_{2kt'+1} = 0$, it is easy to see that the last $2k$ entries of $\mathbf{z}$ are all zeros. Then by induction hypothesis, at least $2k + 1$ entries of the first $2kt'$ elements of $\mathbf{z}$ are non-zero. If $z_{2kt'+1} \neq 0$, one can check that the last $2k - 1$ entries of $\mathbf{z}$ are all non-zero, and at least one of $z_{2kt'-1}$ and $z_{2kt'}$ is non-zero. Thus, $\mathbf{z}$ also has at least $2k + 1$ non-zero entries in this case.

By induction over t', every $\mathbf{z} \neq \mathbf{0}$ such that $A\mathbf{z} = \mathbf{0}$ has at least $2k + 1$ non-zero entries, then the theorem follows. □

The number of measurements by this construction is greater than those of the previous methods. But the advantage of this construction is that the number of nodes in each measurement is at most three despite of the values of n and k.

2) Ring With Each Node Connecting to Four Neighbors: We next consider $\mathcal{G}^4$ in Fig. 4(a). We further impose the constraint that the number of nodes in each measurement cannot exceed d for some predetermined integer d. We neglect $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ for notational simplicity.

All the even nodes are divided into n/d groups. Each group contains $d/2$ consecutive even nodes and is used as a hub to measure $d/2$ odd nodes that have direct edges with nodes in the hub. Then we can identify the values related to all the odd nodes with $nM_{k,d/2}^C/d + n/d$ measurements, and the number of nodes in each measurement does not exceed d. We then measure the even nodes with groups of odd nodes as hubs. In total, the number of measurements is $2nM_{k,d/2}^C/d + 2n/d$, which equals $O(2n(k\log(d/2) + 1)/d)$. When d equals n, the result coincides with Corollary 1. Since n/d measurements are needed to measure each node at least once, we have

Theorem 11: The number of measurements needed to recover k-sparse vectors associated with $\mathcal{G}^4$ with each measurement containing at most d nodes is no less than n/d and no more than $O(2n(k\log(d/2) + 1)/d)$.

Therefore, the ratio of the number of measurements by our construction to the minimum number needed with length constraint is within $Ck\log(d/2)$ for some positive constant C.
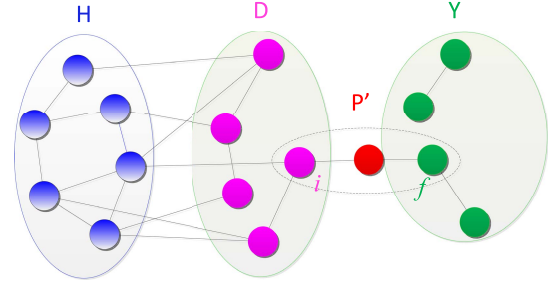

Fig. 9. When $H \cap Y = \varnothing$, use hub $H' = H \cup P'$ to measure nodes $D\backslash i$.

B. Measurements Passing at Least One Node in a Fixed Subset

Recall that in network delay monitoring, a router sends a probing packet to measure the sum of delays on links that the packet transverses. Then every measurement initiated by this router measures the delay on at least one link that is connected to the router. In order to reduce the monitoring cost, one may only employ several routers to initiate measurements, thus, each measurement would include at least one link that is connected to these routers. In the graph model $G = (V, E)$ we consider in this paper, it is equivalent to the requirement that every measurement should contain at least one node in a fixed subset of nodes $Y \subset V$. We will show that this requirement can be achieved with small modifications to Algorithm 1.

After step 3 in Algorithm 1, let H denote the currently chosen hub, and let D denote the set of nodes that one needs to design measurements via hub H. If $H \cap Y$ is not empty, since every measurement constructed to measure nodes in D should contain all the nodes in the H, then it contains at least one node in Y naturally. If $H \cap Y$ is empty, we will find a new hub that contains at least one node in Y. We consider two scenarios. If $D \cap Y$ is not empty, say node i belongs to both set D and set Y, then we let $\hat{H} := H \cup \{i\}$ be the new hub, and design measurements for nodes in $\hat{D} := D\backslash i$ using hub $\hat{H}$. Then every measurement containing hub $\hat{H}$ contains node i and therefore, at least one node in Y. We use one additional measurement to measure i directly. When $D \cap Y$ is also empty, we pick any node j in D and any node f in Y and compute the shortest path from j to f on G. Let P denote the resulting shortest path, which contains a sequence of nodes that connect j to f. Let node i be the last node on path P that belongs to D, and let P' denote the segment of P that connects i to f. Let $H' := H \cup P'$ be the hub, and let $D' := D\backslash i$ be the set of nodes that can be measured via H', see Fig. 9. Then every measurement containing hub H' contains f. We use two additional measurements to measure node i, where one measurement records path P', and the other one measures $P'\backslash i$. With this simple modification, we can measure nodes in D with each measurement containing one node in Y, and the total number of measurements is increased by at most two.

We summarize the above modification in subroutine **Agent**. For measurement design on general graphs, we first replace step 4 in Algorithm 1 in Section IV-B with subroutine Agent($V\backslash S$, S, Y, G). Then in each iteration the number of measurements is increased by at most two. We then replace step 9 with measuring the paths P^* and $P^*\backslash n_{\text{last}}$, where

---

Subroutine 3 Agent(H, D, Y, G)

Initial: hub H, set D of nodes to measure, set Y of fixed nodes, G

1 **if** $H \cap Y \neq \varnothing$ **then**

2   Design $f(k, |D|) + 1$ measurements to recover k-sparse vectors associated with D using H as a hub.

3 **else**

4   **if** $D \cap Y \neq \varnothing$ **then**

5   Pick some node i in $D \cap Y$. $\hat{D} := D \backslash i$, $\hat{H} := H \cup \{i\}$. Design $f(k, |\hat{D}|) + 1$ measurements to recover k-sparse vectors associated with $\hat{D}$ using $\hat{H}$ as a hub.

6   Measure i directly.

7   **else**

8   Pick any node j in D and any node f in Y. Find the shortest path P from j to f. Find the last node i on path P that is in D. Let P' denote the segment of P that starts from i and ends at f.

9      $D' := D \backslash i$, $H' := H \cup P'$, design $f(k, |D'|) + 1$ measurements to recover D' with H' as a hub.

10   Measure P' and $P' \backslash i$ to recover node i.

11   **end if**

12 **end if**

---

n_{last} is the last node in G, and P^* connects n_{last} to any node j in Y on the original graph. Therefore, the total number of measurements needed by the modified algorithm is upper bounded by $Rf(k, n) + 3R + 2$, and each measurement in the modified version contains at least one node in Y. The modified algorithm also has at most R iterations. In each iteration, the additional computation compared with that of Algorithm 1 takes time $O(|E| + |V| \log |V|)$, which results from the computation of the shortest path in step 8 of subroutine **Agent**.

VII. SENSITIVITY TO HUB MEASUREMENT ERRORS

When constructing measurements based on the use of a hub, in order to measure nodes in S using hub H, we first measure the sum of nodes in H and then delete it from other measurements to obtain the sum of some subset of nodes in S. This arises the issue that if the sum of H is not measured correctly, this single error would be introduced into all the measurements. Here we prove that successful recovery is still achievable when a hub measurement is erroneous.

Mathematically, let $\mathbf{x}_S$ denote the sparse vector associated with S, and let $\mathbf{x}_H$ denote the vector associated with H. Let $A^{m \times |S|}$ be a measurement matrix that can identify k-sparse vectors associated with a complete graph of $|S|$ nodes. We arrange the vector $\mathbf{x}$ such that $\mathbf{x} = [\mathbf{x}_S^T \quad \mathbf{x}_H^T]^T$, then

$$F = \begin{bmatrix} A & W^{m \times |H|} \\ \mathbf{0}_{|S|}^T & \mathbf{1}_{|H|}^T \end{bmatrix}$$

is the measurement matrix for detecting k non-zeros in S using hub H, where W is a matrix with all '1's, $\mathbf{0}$ is a column vector of all '0's, and $\mathbf{1}$ is a column vector of all '1's. Let vector $\mathbf{z}$ denote the first m measurements, and let z_0 denote the last measurement of the hub H. Then

$$\begin{bmatrix} \mathbf{z} \\ z_0 \end{bmatrix} = \begin{bmatrix} A\mathbf{x}_S + \mathbf{1}^T \mathbf{x}_H \mathbf{1}_m \\ \mathbf{1}^T \mathbf{x}_H \end{bmatrix},$$

or equivalently

$$\mathbf{z} - z_0 \mathbf{1}_m = A\mathbf{x}_S. \tag{28}$$

If there is some error e_0 in the last measurement, i.e., instead of z_0, the actual measurement we obtain is

$$\hat{z}_0 = \mathbf{1}^T \mathbf{x}_H + e_0,$$

then e_0 affects the recovery accuracy of $\mathbf{x}_S$ through (28).

To eliminate the impact of e_0, we model it as an entry of an augmented sparse signal to recover. Let $\mathbf{x}' = [\mathbf{x}^T \quad e_0]^T$, and $A' = [A \quad -\mathbf{1}_m]$, we have

$$A'\mathbf{x}' = \mathbf{z} - \hat{z}_0 \mathbf{1}_m. \tag{29}$$

Then, recovering $\mathbf{x}_S$ in the presence of hub error e_0 is equivalent to recovering $k + 1$-sparse vector $\mathbf{x}'$ from (29).

We consider one special construction of matrix $A^{m \times |S|}$ for a complete graph. A has '1' on every entry in the last row, and takes value '1' and '0' with equal probability independently for every other entry. $A' = [A \quad -\mathbf{1}_m]$, let $\hat{A}$ be the submatrix of the first $m - 1$ rows of A'. Let $\mathbf{y} = \mathbf{z} - \hat{z}_0 \mathbf{1}_m$, and let $\hat{\mathbf{y}}$ denote the first $m - 1$ entries of $\mathbf{y}$. We have,

$$(2\hat{A} - W^{(m-1) \times |S|})\mathbf{x}' = 2\hat{\mathbf{y}} - y_m.$$

We recover $\mathbf{x}'$ by solving the ℓ_1-minimization problem,

$$\min \|\mathbf{x}\|_1, \quad \text{s.t.} \ (2\hat{A} - W^{(m-1) \times |S|})\mathbf{x} = 2\hat{\mathbf{y}} - y_m. \tag{30}$$

Theorem 12: With the above construction of A, when $m \geq C(k + 1) \log |S|$ for some constant $C > 0$ and $|S|$ is large enough, with probability at least $1 - O(|S|^{-\alpha})$ for some constant $\alpha > 0$, $\mathbf{x}'$ is the unique solution to (30) for all $k + 1$-sparse vectors $\mathbf{x}'$ in $\mathcal{R}^{|S|+1}$.

Theorem 12 indicates that even though the hub measurement is erroneous, one can still identify k-sparse vectors associated with S with $O((k + 1) \log |S|)$ measurements.

The proof of Theorem 12 relies heavily on Lemma 2.

Lemma 2: If matrix $\Phi^{p \times n}$ takes value $-1/\sqrt{p}$ on every entry in the last column and takes value $\pm 1/\sqrt{p}$ with equal probability independently on every other entry, then for any $\delta > 0$, there exists some constant C such that when $p \geq C(k+1) \log n$ and n is large enough, with probability at least $1 - O(n^{-\alpha})$ for some constant $\alpha > 0$ it holds that for every $U \subseteq \{1, \ldots, n\}$ with $|U| \leq 2k + 2$ and for every $\mathbf{x} \in \mathcal{R}^{2k+2}$,

$$(1 - \delta)\|\mathbf{x}\|_2^2 \leq \|\Phi_U \mathbf{x}\|_2^2 \leq (1 + \delta)\|\mathbf{x}\|_2^2, \tag{31}$$

where Φ_U is the submatrix of Φ with column indices in U.

Proof: Consider matrix $\Phi'^{p \times n}$ with each entry taking value $\pm 1/\sqrt{p}$ with equal probability independently. For every realization of matrix Φ', construct a matrix $\hat{\Phi}$ as follows. For every $i \in \{1, \ldots, p\}$ such that $\Phi'_{in} = 1/\sqrt{p}$, let $\hat{\Phi}_{ij} = -\Phi'_{ij}$ for all $j = 1, \ldots, n$. Let $\hat{\Phi}_{ij} = \Phi'_{ij}$ for every other entry. One can check that $\hat{\Phi}$ and Φ follow the same probability distribution. Besides, according to the construction of $\hat{\Phi}$, for any subset $U \subseteq \{1, \ldots, n\}$,

$$\Phi_U'^T \Phi_U' = \hat{\Phi}_U^T \hat{\Phi}_U. \tag{32}$$

The Restricted Isometry Property [12] indicates that the statement in Lemma 2 holds for Φ'. From (32), and the fact
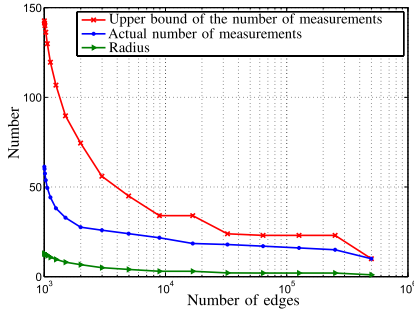
Fig. 10.   Random graph with $n = 1000$.

Fig. 11.   BA model with increasing n.

that $\|\Phi'_U \mathbf{x}\|_2^2 = \mathbf{x}^T {\Phi'_U}^T \Phi'_U \mathbf{x}$, the statement also holds for $\hat{\Phi}$. Since $\hat{\Phi}$ and Φ follow the same probability distribution, the lemma follows. $\square$

Proof (of Theorem 12): From Lemma 2, we fix some $\delta < 0.4531$ [30], then when $m \geq C(k+1)\log|S|$ for some $C > 0$ and $|S|$ is large enough, matrix $(2\hat{A} - W^{(m-1)\times|S|})/\sqrt{m-1}$ satisfies (31) with probability at least $1 - O(|S|^{-\alpha})$. Then from [13] and [30], one can recover all $k+1$-sparse vectors correctly from (30). $\square$

VIII. SIMULATION

Experiment 1 (Effectiveness of Algorithm 1): The exact number of required measurements to recover 1-sparse n-dimensional signals associated with a complete graph satisfies $M_{1,n}^C = \lceil \log(n+1) \rceil$, and the corresponding measurement matrix can be constructed by taking the binary expansion of integer i as column i [25]. Moreover, the number of measurements required to recover k-sparse vectors is within certain constant times $k M_{1,n}^C$ from (3). Therefore, the number of measurements required for recovering 1-sparse signals on a graph G is a simple indicator of the measurement requirement for general k-sparse signals.

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 via some hub. The last group contains one node and can be measured directly. The total number of constructed measurements for recovering 1-sparse signals 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.

We start with a uniformly generated random tree with $n = 1000$ nodes and gradually increase the number of edges in the graph until it becomes a complete graph of n nodes. Fig. 10 shows the number of constructed measurements for graphs versus the number of edges in the graph. All the results are averaged over one hundred realizations. When the graph evolves from a tree to a complete graph, its radius R decreases from 13 to 1, and the number of measurements constructed for recovering 1-sparse vectors associated with the graph decreases from 62 to 10. Note that when the average node degree in the graph is only four, the number of measurements is already within $3M_{1,n}^C$, where $M_{1,n}^C$ (equals 10 if n equals 1000) is the required number of measurements in a complete graph. We also plot the upper bound $R\lceil \log n \rceil + R + 1$ provided by Theorem 8. One can see that the number of measurements actually constructed is much less than the upper bound.
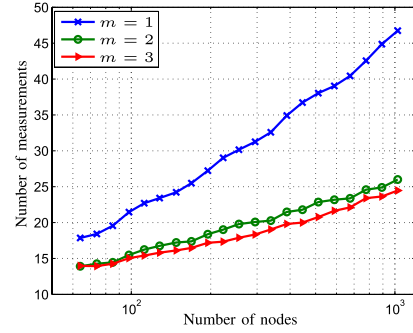
We also consider the scale-free network with Barabási-Albert (BA) model [4] where the graph initially has m_0 connected nodes, and new nodes are added to the graph gradually. Each new node connects to m existing nodes with a probability that is proportional to the degree of the existing nodes. We start with a random tree of 10 nodes and increase the total number of nodes to 1024. Every result in Fig. 11 is averaged over one hundred realizations. Since the diameter of BA model is $O(\log n / \log \log n)$ [9], then by Theorem 8, the number of our constructed measurements is upper bounded by $O(\log^2 n / \log \log n)$.

Experiment 2 (Recovery Performance With Hub Error): 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. We construct the measurements as follows. For each of the first two groups with size n_i ($i = 1, 2$), we generate $\lceil n_i/2 \rceil$ random measurements, each of which measures a random subset of the group together with its hub. Every node of the group is included in the random subset independently with probability 0.5. 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 the widely used ℓ_1-minimization [13] to recover the subvectors associated with each of the first two groups, and the last three entries of $\mathbf{x}_0$ can be obtained from measurements directly. However, as discussed in Section VII, an error in a hub measurement degrades the recovery accuracy of subvectors associated with that group. To address this issue, we use a modified ℓ_1-minimization in which the errors in the two hubs are treated as entries of an augmented vector to recover. Specifically, denote the augmented vector by $\mathbf{z} = [\mathbf{x}_0^T, e_1, e_2]^T$ and the augmented matrix by $\tilde{A} = [A \; \boldsymbol{\beta} \; \boldsymbol{\gamma}]$, 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 entry corresponding to the measurement of the first (or second) hub and '0' elsewhere. We then recover $\mathbf{z}$ (and thus $\mathbf{x}_0$) from $\mathbf{y} = \tilde{A}\mathbf{z}$ by running ℓ_1-minimization on each group separately.

Fig. 12 compares the recovery performance of our modified ℓ_1-minimization and the conventional ℓ_1-minimization, where the hub errors e_1 and e_2 are drawn from standard Gaussian distribution with zero mean and unit variance. For every
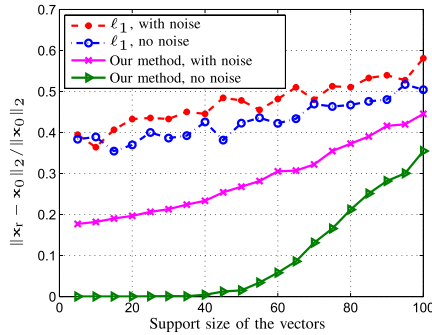
Fig. 12.   Recovery performance with hub errors.

support size k, we randomly generate two hundred k-sparse vectors $\mathbf{x}_0$, and let $\mathbf{x}_r$ denote the recovered vector. Even with the hub errors, the average $\|\mathbf{x}_r - \mathbf{x}_0\|_2/\|\mathbf{x}_0\|_2$ is within 10^{-6} when $\mathbf{x}_0$ is at most 35-sparse by our method, while by standard ℓ_1-minimization, $\|\mathbf{x}_r - \mathbf{x}_0\|_2/\|\mathbf{x}_0\|_2$ is at least 0.35. We also consider the case that besides errors in hub measurements, the other measurements all have i.i.d. Gaussian noise with zero mean and variance 0.04^2. The average $\|\mathbf{x}_r - \mathbf{x}_0\|_2/\|\mathbf{x}_0\|_2$ here is smaller by our method than that by ℓ_1-minimization.

IX. CONCLUSION

This paper addresses the sparse recovery problem with graph constraints. We provide explicit measurement constructions for special graphs, and propose measurement design algorithms for general graphs. Our construction for a line network is optimal since the number of measurements needed is minimized. Our constructions on other graphs also improve over the existing results. We characterize the relationship between the number of measurements for sparse recovery and the graph topology and develop bounds of the minimum number of measurements needed for sparse recovery on a given graph.

We remark that this paper is the first step towards network measurement constructions for sparse recovery in a compressed sensing framework with topological constraints, and several practical concerns have not been much addressed yet. We only include the length constraints on the measurements into the measurement construction for some special graphs (Section VI-A). Since measurements with short lengths are preferred in network applications, it is interesting to develop measurement construction methods for general graphs when the measurement length is bounded. We have not analyzed the effect of the measurement noise on recovery performance. Also, we assume full knowledge of the fixed network topology, and measurement construction when the topology is time-varying or partially known is an open question.

APPENDIX A
PROOF OF THEOREM 4

Let $A^{m \times n}$ denote the matrix with m realizations of the n-step Markov chain. To prove the statement, from [12], 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. Given I, the probability that $\mathrm{rank}(A_{S_j I}) < 2k$ is the same for every given j, and let it denoted by π_d^I. Let P_d^I denote the probability that $\mathrm{rank}(A_I) < 2k$, then

$$P_d^I \leq (\pi_d^I)^{\lfloor \frac{m}{2k} \rfloor}. \tag{33}$$

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

$$\pi_d^I \leq 1 - P(A_{S_j I} \text{ is a row permutation of } B). \tag{34}$$

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 $P(X_k = 1) \geq 1/2$. Since B has $(2k)!$ different row permutations,

$$P(A_{S_j I} \text{ is a row permutation of } B) \geq (2k)!/2^{4k^2+2k-1}. \tag{35}$$

Combining (33), (34) and (35), we have

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

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

APPENDIX B
PROOF OF PROPOSITION 1

We view nodes $2i - 1$ and $2i$ as a group for every i ($1 \leq i \leq \lfloor \frac{n}{2} \rfloor$), denoted by B_i. Consider the special case that for some t, both nodes in B_t are '1's, and all other nodes are '0's. Then every measurement that passes either node or both nodes in B_t is always '1'. Consider a reduced graph with B_i, $\forall i$ as nodes, and edge (B_i, B_j) ($i \neq j$) 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 nonzero element in the reduced network. Then only to locate two consecutive non-zero elements associated with $\mathcal{G}^4$, we need at least $\lfloor n/4 \rfloor$ measurements, and the claim follows.

APPENDIX C
PROOF OF THEOREM 7

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

We next show that the NP-complete r-coloring ($r \geq 3$) problem is polynomial time reducible to 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 edges (w, u) and (w, v). Let W denote the set of nodes added.

Add a edge between every pair of nodes in V not already joined by a edge. Let H denote the augmented graph and let $V' = V \cup W$. We claim that if there exists an r-partition of H, then we can obtain an r-coloring of G, and vice versa.

Let S_i $(i = 1, \ldots, r)$ be an r-partition of H. Suppose there exists edge $(u, v) \in E$ s.t. 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 $H_{V' \backslash S_i}$ is disconnected since w does not connect to any node in $V' \backslash S_i$. 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 $C_i \subset V$ $(i = 1, \ldots, r)$ denote an r-coloring of G. We claim that $N_i = C_i$ $(i = 1, \ldots, r-1)$, and $N_r = C_r \cup W$ form an r-partition of H. First note for every $u \in V$, at least one of its neighbors is not in the same set as u. For every $w \in W$, w is directly connected to u and v for some $(u, v) \in E$, and 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 $H_{V' \backslash N_i}$ is connected for all i. $H_{V' \backslash N_r}$ is a complete subgraph, and thus connected. For every $i < r$, let $S_i := V \backslash C_i$, then $V' \backslash N_i = S_i \cup W$. H_{S_i} is a complete subgraph, and thus connected. For every $w \in W$, since its two neighbors cannot be both in C_i, then at least one neighbor belongs to S_i, thus $H_{V' \backslash N_r} = H_{S_i \cup W}$ is connected. N_i $(i = 1, \ldots, r)$ thus forms an r-partition.

REFERENCES

[1] S. S. Ahuja, S. Ramasubramanian, and M. M. Krunz, "Single-link failure detection in all-optical networks using monitoring cycles and paths," *IEEE/ACM Trans. Netw.*, vol. 17, no. 4, pp. 1080–1093, Aug. 2009.

[2] L. Applebaum, S. D. 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, 2009.

[3] 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, Jun. 2011.

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

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

[6] R. Berinde, A. C. Gilbert, P. Indyk, H. Karloff, and M. J. Strauss. (2008). "Combining geometry and combinatorics: A unified approach to sparse signal recovery." [Online]. Available: http://arxiv.org/abs/0804.4666

[7] T. Blumensath, "Compressed sensing with nonlinear observations and related nonlinear optimization problems," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3466–3474, Jun. 2013.

[8] B. Bollobás, *Random Graphs*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[9] B. Bollobás and O. Riordan, "The diameter of a scale-free random graph," *Combinatorica*, vol. 24, no. 1, pp. 5–34, 2004.

[10] Y. Breitbart, C.-Y. Chan, M. Garofalakis, R. Rastogi, and A. Silberschatz, "Efficiently monitoring bandwidth and latency in IP networks," in *Proc. IEEE INFOCOM*, vol. 2. Apr. 2001, pp. 933–942.

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

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

[13] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[14] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson, "Randomness conductors and constant-degree lossless expanders," in *Proc. 34th Annu. ACM STOC*, 2002, pp. 659–668.

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

[16] M. Cheraghchi, A. Karbasi, S. Mohajer, and V. Saligrama, "Graph-constrained group testing," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 248–262, Jan. 2012.

[17] D. B. Chua, E. D. Kolaczyk, and M. Crovella, "Network kriging," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 12, pp. 2263–2272, Dec. 2006.

[18] F. Chung and L. Lu, "The diameter of sparse random graphs," *Adv. Appl. Math.*, vol. 26, no. 4, pp. 257–279, 2001.

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

[20] M. Coates, Y. Pointurier, and M. Rabbat, "Compressed network monitoring for IP and all-optical networks," in *Proc. 7th ACM SIGCOMM IMC*, 2007, pp. 241–252.

[21] G. Cormode and S. Muthukrishnan, "Combinatorial algorithms for compressed sensing," in *Structural Information and Communication Complexity* (Lecture Notes in Computer Science). Berlin, Germany: Springer-Verlag, 2006, pp. 280–294.

[22] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complex.*, vol. 23, nos. 4–6, pp. 918–925, 2007.

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

[24] D. L. Donoho and J. Tanner, "Sparse nonnegative solution of underdetermined linear equations by linear programming," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 27, pp. 9446–9451, 2005.

[25] R. Dorfman, "The detection of defective members of large populations," *Ann. Math. Statist.*, vol. 14, no. 4, pp. 436–440, 1943.

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

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

[28] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2558–2567, Sep. 2002.

[29] P. Erdõs and A. Rényi, "On the evolution of random graphs," *Pub. Math. Inst. Hungaria Acad. Sci.*, vol. 5, pp. 17–61, 1960.

[30] S. Foucart and M.-J. Lai, "Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$," *Appl. Comput. Harmonic Anal.*, vol. 26, no. 3, pp. 395–407, 2009.

[31] M. Gonen and Y. Shavitt, "A $O(\log n)$-approximation for the set cover problem with set ownership," *Inf. Process. Lett.*, vol. 109, no. 3, pp. 183–186, 2009.

[32] A. Gopalan and S. Ramasubramanian, "On identifying additive link metrics using linearly independent cycles and paths," *IEEE/ACM Trans. Netw.*, vol. 20, no. 3, pp. 906–916, Jun. 2012.

[33] F. Harary and R. Z. Norman, "Some properties of line digraphs," *Rendiconti Circolo Matematico Palermo*, vol. 9, no. 2, pp. 161–168, 1960.

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

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

[36] D. B. Johnson, "Efficient algorithms for shortest paths in sparse networks," *J. ACM*, vol. 24, no. 1, pp. 1–13, 1977.

[37] G. Kutyniok, "Theory and applications of compressed sensing," *GAMM-Mitteilungen*, vol. 36, no. 1, pp. 79–101, 2013.

[38] H. X. Nguyen and P. Thiran, "Using end-to-end data to infer lossy links in sensor networks," in *Proc. 25th IEEE INFOCOM*, 2006, pp. 1–12.

[39] 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, no. 1, pp. 275–285, Feb. 2011.

[40] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 372–386, Feb. 2012.

[41] M. Wang, W. Xu, E. Mallada, and A. Tang, "Sparse recovery with graph constraints: Fundamental limits and measurement construction," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 1871–1879.

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

[43] 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*, Mar. 2010, pp. 1–5.

[44] W. Xu, E. Mallada, and A. Tang, "Compressive sensing over graphs," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 2087–2095.

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

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

Meng Wang (S'06–M'12) received B.E. (Hon.) and M.S. (Hon.) from Tsinghua University, Beijing, China, and Ph.D. from Cornell University, Ithaca, NY, in 2005, 2007, and 2012, all in Electrical Engineering.

She is currently an Assistant Professor in the Department of Electrical, Computer, and Systems Engineering at Rensselaer Polytechnic Institute. Her research interests include communication networks, signal processing, and nonconvex optimization and its applications.

Dr. Wang is a recipient of Jacobs Fellowship of Cornell University in 2008 and 2010.

Weiyu Xu received his B.E. in Information Engineering from Beijing University of Posts and Telecommunications in 2002, and a M.S. degree in Electronic Engineering from Tsinghua University in 2005. He received a M.S. and a Ph.D. degree in Electrical Engineering in 2006 and 2009 from California Institute of Technology, with a minor in Applied and Computational Mathematics.

He is currently an Assistant Professor at the Department of Electrical and Computer Engineering at the University of Iowa. His research interests are in signal processing, compressive sensing, communication networks, information and coding theory.

Dr. Xu is a recipient of the Information Science and Technology Fellowship at Caltech, and the recipient of Charles and Ellen Wilts doctoral research award in 2010.

Enrique Mallada (S'09) received his Ingeniero en Telecomunicaciones degree from Universidad ORT, Uruguay, in 2005. From 2006 to 2008, he was teaching and research assistant in the Department of Telecommunications at Universidad ORT. He obtained his Ph.D. degree in electrical and computer engineering at Cornell University, Ithaca, NY in 2013.

He is currently a postdoctoral researcher at California Institute of Technology. His research interests include communication networks, control, nonlinear dynamics and optimization.

Ao Tang (S'01–M'07–SM'11) received the B.E. (Honors.) in electronics engineering from Tsinghua University, Beijing, China, and the M.S. and Ph.D. degrees in electrical engineering with a minor in applied and computational mathematics from the California Institute of Technology, Pasadena, CA, in 1999, 2002, and 2006, respectively.

He is currently an Associate Professor in the School of Electrical and Computer Engineering at Cornell University, where he conducts research on the control and optimization of engineering networks including communication networks, power networks and on-chip networks.

Dr. Tang was a recipient of the Cornell Engineering School Michael Tien' 72 Excellence in Teaching Award in 2011, the Young Investigator Award from the Airforce Office of Scientific Research (AFOSR) in 2012, and the Presidential Early Career Award for Scientists and Engineers (PECASE) in 2012.