Recent Results on Sparse Recovery over Graphs

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

***Abstract*—In this paper, we review our recent results on sparse recovery over graphs, which was motivated by network tomography problems. Our finding has made a new connection between coding theory and graph theory. We also discuss robustness of our proposed measurement construction.**

I. Introduction

Compressive sensing is a new paradigm in signal processing theory, which proposes to sample and recover parsimonious signals efficiently. The basic idea of compressive sensing is that if an object being measured is well-approximated by a lower dimensional object (e.g., sparse vector, low-rank matrix, etc.) in an appropriate space, one can exploit this property to achieve perfect recovery of the object. Compressive sensing [3][9] characterizes this phenomenon for sparse signal vectors, and presents efficient signal recovery schemes, from a small number of measurements. It has been applied to seismology, error correction and medical imaging since the breakthrough works [3][9]. However, its role in networking is still in its early stage [7][8][12][20].

In many of the compressive literature, there are no restrictions on how the compressive sensing matrices are constructed [3][9]. But this is generally not the case in real-world applications. The physical constraints of specific applications make some measurements matrices infeasible, and thus greatly limit the set of feasible measurements. In [22] and [23], we formulated and considered compressive sensing problems where the constructions of the measurement matrices have to satisfy certain graph-theoretic constraints. These graph-constrained compressive sensing problems are mainly motivated by compressive sensing applications in networks.

Network Link Monitoring: In operating communication networks, we are often interested in inferring and monitoring the network performance characteristics, such as delay and packet loss rate, associated with each link. However, making direct measurements and monitoring for each link can be costly and operationally difficult, often requiring the participation from routers or potentially unreliable middle network nodes, while end-to-end communication path characteristics are often readily available. The problem of quickly inferring and monitoring the network link characteristics from indirect end-to-end (aggregate) measurements falls in the area of *network tomography*, which is useful for network traffic engineering [20] and fault diagnosis [2][4][6][10] [11][13][14][15][16][17][18][19][21].

Suppose that we have probes along m source-destination pairs over a network (m is smaller than the number n of network links in the networks). We are interested in identifying congested links with large delays or high packet loss rates from the probe measurements. We note that the delay over each source-destination pair is a sum of the delays over each link on the route between this source-destination pair, giving a natural linear mixing of the link delays on the route. Abstractly, let $\mathbf{x}$ be an $n \times 1$ non-negative vector whose j-th element represents the delay over edge j and let $\mathbf{y}$ be an $m \times 1$ dimensional vector whose i-th element is the end-to-end delay measurement for the i-th source-destination pair. Then

$$\mathbf{y} = A\mathbf{x}, \tag{1}$$

where A is an $m \times n$ matrix, whose element in the i-th row and j-th column is '1' if the i-th source-destination pair routes through the j-th link and '0' otherwise.

$n > m$ means we only have an underdetermined system, but it is reasonable to assume that there are only a small fraction of links that are congested, i.e., their link delays are considerably larger than the delays over other links. This provides the foundation to link our network tomography problems to compressive sensing. However, there are important differences between network tomography problems and general compressive sensing formulations:

- the element $A_{i,j}$ from A is either 0, when the measurement path i does not go through link j, or an integer b, when the measurement path i goes through link j for $b > 0$ times. Generally, the number b is '1', which often makes the matrix a '0' and '1' matrix.
- More importantly, all the nonzero elements in row i of A must correspond to a *connected* communication path (or walk).

It is natural to ask whether the idea of compressive sensing still works using these graph-constrained measurements. In this paper, we review and compare two approaches for constructing graph-constrained measurements: random walk constructions and deterministic constructions. We also provide our recent results on the robustness of graph-constrained compressive sensing constructions.

The paper is organized as follows. In Section II, we give a general mathematical model for graph-constrained compressive sensing. In Section III, we review a random walk approach for constructing graph-constrained measurements. In Section IV, we review a deterministic construction of measurement matrices for compressive sensing, and compare the deterministic construction with the random walk construction. Section V presents results on the robustness of graph-constrained compressive sensing. We conclude in Section VI.

II. Mathematical Formulation

Consider a graph $G = (V, E)$, where V denotes the set of nodes with cardinality $|V| = n$ and E denotes the set of links.

978-1-4673-0323-1/11/$26.00 ©2011 IEEE

Each node i is associated with a real number x_i, and we say vector $\mathbf{x} = (x_i, i = 1, ..., n)$ is associated with G. Let $T = \{i \mid x_i \neq 0\}$ denote the support of $\mathbf{x}$, and let $\|\mathbf{x}\|_0 = |T|$ denote the number of non-zero entries of $\mathbf{x}$, we say $\mathbf{x}$ is a k-sparse vector if $\|\mathbf{x}\|_0 = k$. We aim to infer the vector $\mathbf{x}$ from indirect additive observations.

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

(A1): A set S of nodes can be measured together in one measurement if and only if G_S is connected. **(A2)**: The measurement is an additive sum of values at the corresponding nodes.

(A1) captures the graph constraints.

(A2) follows from the additive property of many network characteristics, e.g. delays and packet loss rates.

Note that for the mentioned network link monitoring application, we can abstract each network link as a node in the graph G, and two nodes in G are connected by an edge in $G = (V, E)$ if and only if the corresponding network links share the same network terminal. It is essential to notice the distinction between the "network link" and "graph edge" when translating the network link monitoring application to this graph model.

Let $\mathbf{y} \in \mathcal{R}^m$ $(m \ll n)$ denote the vector of m measurements. Let A be an $m \times n$ measurement matrix with $A_{ij} = 1$ $(i = 1, ..., m,\ j = 1, ..., n)$ if and only if node j is included in the ith measurement and $A_{ij} = 0$ otherwise. Then we have $\mathbf{y} = A\mathbf{x}$. We say 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. The advantage of sparse recovery is that with the non-adaptive measurement matrix A, it can identify n-dimensional vectors from m $(m \ll n)$ measurements as long as the vectors are sparse.

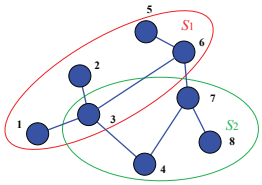

Fig. 1: Network Example

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

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

Given a graph G with n nodes, let $M^G_{k,n}$ denote the minimum number of non-adaptive measurements needed to identify all k-sparse vectors associated with G. Let $M^C_{k,n}$ denote the minimum number of non-adaptive measurements needed in a complete graph with n nodes. In complete graphs, since any subset of nodes can be measured together, any 0-1 matrix is a feasible measurement matrix. Existing results [26], [24], [31] show that with overwhelming probability a random 0-1 matrix with $O(k \log(n/k))$ rowscan identify all k-sparse vectors, and we can recover the sparse vector by ℓ_1-minimization, which returns the vector with the least ℓ_1-norm[1] among those that can produce the obtained measurements. Then we have

$$M^C_{k,n} = O(k \log(n/k)). \tag{2}$$

We will use (2) for the analysis of construction methods. Explicit constructions of measurement matrices for complete graphs also exist, e.g., [24], [27], [28], [31]. In this paper, given a graph, we will discuss how to construct graph-constrained measurement matrices and give general bounds on how many measurements we need.

III. Graph-constrained Measurements from Random Walks

In this section, we consider construction of measurement matrices using random walks on the graph. In our construction, each row of the measurement matrix A corresponds to a single random walk of a predetermined length t on the graph. If the i-th random walk, which corresponds to the i-th row of the measurement matrix, goes through node j, then $A_{i,j} = 1$; otherwise, $A_{i,j} = 0$. For each random walk, we uniformly randomly pick a starting vertex from V and then perform a standard random walk over the graph. The length of the random walk is denoted by t.

It has been demonstrated in [22] that $O(k \log(n))$ measurements are enough for recovering any k-sparse link vector for a sufficiently connected graph with n nodes.

1) Graph Assumptions: Before we proceed, following the works on graph-constrained group testing [11], [5], we introduce the following assumptions on the graphs.

The undirected graph $G = (V, E)$ is called a (D, c) uniform graph if for some constant c, the degree of each vertex $v \in V$ is between D and cD. Suppose that a standard random walk over the graph has a stationary distribution μ over the nodes. The δ-mixing time of G is defined as the smallest t' such that a random walk of length t' starting at any vertex in G ends up having a distribution μ' such that $\|\mu - \mu'\|_\infty \leq \delta$. We define $T(n)$ as the δ-mixing time of G for $\delta = \frac{1}{(2cn)^2}$.

2) $O(k \log(n))$ measurements are sufficient: In compressive sensing, we adopt an $m \times n$ measurement matrix generated by m independent random walks. From [22], we have the following theorem,

Theorem 1. *There is a degree $D_0 = O(c^2 k T^2(n))$ and $t = O(\frac{nD}{c^3 k T(n)})$ such that whenever $D \geq D_0$, by setting the path lengths $t = O(\frac{nD}{c^3 k T(n)})$ the following holds. If $m = O(c^4 T^2(n) k \log(n))$, then with high probability, all the k-sparse signal vectors can be recovered from m measurements generated by m random walks.*

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

m	Compressive sensing	Group Testing
Graph constrained	$O(k\log(n))$(this paper)	$O(k^2\log(\frac{n}{k}))$[5]
General	$O(k\log(\frac{n}{k}))$[3]	$O(k^2\log(\frac{n}{k}))$[30]

TABLE I: Number of measurements needed in different scenarios

Table I summarizes the results for the number of measurements needed in graph constrained problems or general problems without graph constraints.

It can also be shown that with $m = O(c^4T^2(n)k\log(n))$ measurements generated by random walks, ℓ_1 minimization decoding for sparse recovery can be used to recover k-sparse signal vectors.

IV. Graph-constrained Measurements from Deterministic Constructions

A. Measurement Construction Based on r-partition

Different from the random walk measurement, we now give a deterministic measurement construction for graph-constrained compressive sensing based on a concept of r-partition for a general graph G. We first give the definition of r-partition [23].

Definition 1 (r-partition)**.** *Given $G = (V, E)$, disjoint subsets N_i $(i = 1, ..., r)$ of V form an r-**partition** of G if and only if these two conditions both hold: (1) $\cup_{i=1}^{r} N_i = V$, and (2) $\forall i$, $V\backslash N_i$ is a hub for N_i, namely, $V\backslash N_i$ induces a connected subgraph and any node from N_i is directly connected to at least one node from $V\backslash N_i$.*

With the above definition, we have the following theorem about measurement construction using r-partition. The basic idea is to use the hub $V\backslash N_i$ as the bridge, one can then freely get the sum of any subset, say $S' \subseteq N_i$, of nodes in the set N_i. Since the hub $V\backslash N_i$ induces a connected subgraph, one can measure the sum of all the nodes in $V\backslash N_i$. Using the hub $V\backslash N_i$ as the bridge, one can measure the sum of all the nodes in the subset $S' \subseteq N_i$ and $V\backslash N_i$. This is possible because any node in S' is directly connected to the hub $V\backslash N_i$. By a simple subtraction, we know immediately the sum of all the nodes in the subset S'. So we have the following theorem.

Theorem 2. *[23] If G has an r-partition N_i $(i = 1, ..., r)$, then the number of measurements needed to recover k-sparse vectors associated with G is at most $\sum_{i=1}^{r} M^C_{k,|N_i|} + r$, which is $O(rk\log(n/k)) + r$.*

Proof: Note that $M^C_{k,|N_i|} + 1$ measurements are enough to recover k-sparse subvector associated with N_i via its hub $V\backslash N_i$. Note that the one additional measurement is for the measurement offering the sum of all the nodes in $V\backslash N_i$. ■

We next apply this result to the Erdős-Rényi random graph $G(n, p)$, which contains n nodes and there exists an link between any two nodes independently with probability p. Note that if $p \geq (1+\epsilon)\log n/n$ for some constant $\epsilon > 0$, $G(n, p)$ is connected almost surely [25].

Theorem 3. *[23] For Erdős-Rényi random graph $G(n, p)$ with $p = \beta\log n/n$, if $\beta \geq 2 + \epsilon$ for some constant $\epsilon > 0$, then any two disjoint subsets N_1 and N_2 of nodes with $|N_1| = |N_2| = n/2$ form a 2-partition with high probability. Moreover, with high probability the number of measurements needed to recover k-sparse vectors associated with $G(n, p)$ is at most $2M^C_{k,n/2} + 2$, which is $O(2k\log(n/(2k))) + 2$.*

Proof: We reproduce the proof from [23] here. Let N_1 be any subset of V with $|N_1| = n/2$, and let $N_2 = V\backslash N_1$. Then G_{N_1} and G_{N_2} are both Erdős-Rényi random graphs with $n/2$ nodes, and are connected almost surely when $p \geq (2 + \epsilon)\log n/n$.

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

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

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

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

It is now interesting to compare the results of deterministic construction and random-walk construction. [5] considers group testing over Erdős-Rényi random graphs using random walks and shows that $O(k^2\log^3 n)$ measurements are enough to identify up to k non-zero entries in an n-dimensional logical vector provided that $p = \Theta(k\log^2 n/n)$. Here with compressed sensing setup and 2-partition results, we can recover k-sparse vectors in $\mathcal{R}^n$ with $O(2k\log(n/(2k))) + 2$ measurements when $p > (2 + \epsilon)\log n/n$ for some $\epsilon > 0$. Note that this result also improves over the result in [22], which requires $O(k\log^3 n)$ random-walk measurements for compressed sensing on $G(n, p)$. So for certain

V. Sensitivity analysis of the Hub Measurement

The key idea to design measurements on graphs is that we can use a connected subset H of nodes as a hub to freely measure the set S of nodes that are directly connected to the hub. We measure the hub with one measurement. The sum of any subset of nodes in S is obtained by first measuring the sum of these nodes and H, and then deleting the sum of H. This raises the issue that if the measurement of H has an error, and all the measurements we take over S using H a hub are correct, this single error in H is propagated into all the measurements since we need to delete the erroneous measurement of H from every other measurement.

Mathematically, let $\mathbf{x}_S$ denote the sparse vector associated with S, and let $\mathbf{x}_H$ denote the vector associated with H and let $A^{m\times|S|}$ be the measurement matrix that can identify k errors on 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

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

is the measurement matrix for detecting k errors in S using hub H, where W is an m by $|H|$ matrix with all '1's, $\mathbf{0}_{|S|}$ is a $|S|$-dimensional column vector of all '0's, and $\mathbf{1}_{|H|}$ is a $|H|$-dimensional 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

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

or equivalently

$$\mathbf{z} - z_0\mathbf{1}_m = A\mathbf{x}_S,$$

and $\mathbf{x}_S$ can be correctly covered given $\mathbf{z}$, z_0 and A by the Compressed Sensing theory. Now if there is some error e_0 in the last measurement, i.e.,

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

then when recovering $\mathbf{x}_S$, we have

$$A\mathbf{x}_S = \mathbf{z} - \hat{z}_0\mathbf{1}_m = \mathbf{z} - z_0\mathbf{1}_m - e_0\mathbf{1}_m.$$

Then error e_0 in the hub measurement can in fact lead to errors in every other measurement of nodes in S, and finally the error in the recovery of $\mathbf{x}_S$.

To eliminate the impact of the errors in the hub measurements on the recovery accuracy, we model the errors in the hub measurements as entries of an augmented sparse signal to recover. In the above example, let $\mathbf{x}' = [\mathbf{x}^T \quad e_0]^T$, let $F' = [F \quad \mathbf{e}_{m+1}]$, where $\mathbf{e}_i$ is a column vector with '1' on the ith entry and '0' elsewhere. Then the measurements and the augmented signal are related by

$$\begin{bmatrix} \mathbf{z} \\ \hat{z}_0 \end{bmatrix} = F'\mathbf{x}' = F\mathbf{x} + e_0\mathbf{e}_{m+1} = \begin{bmatrix} A\mathbf{x}_S + \mathbf{1}^T\mathbf{x}_H \\ \mathbf{1}^T\mathbf{x}_H + e_0 \end{bmatrix}.$$

When recovering $\mathbf{x}_S$, we delete the sum of the hub from every other measurement, and the obtained equations are

$$A\mathbf{x}_S - e_0\mathbf{1}_m = \mathbf{z} - \hat{z}_0\mathbf{1}_m,$$

with the equivalent matrix form

$$A'\mathbf{x}'_S = \mathbf{z} - \hat{z}_0\mathbf{1}_m,$$

where $A' = [A \quad -\mathbf{1}_m]$, and $\mathbf{x}'_S = [\mathbf{x}_S^T \quad e_0]^T$. We know that with the measurement matrix A, one can recover all k-sparse $|S|$-dimensional vectors $\mathbf{x}_S$, but with the erroneous hub measurement, the question now is can one recover $|S|+1$-sparse vector $\mathbf{x}'_S$ with the measurement matrix A'? If the answer is yes, then with the measurement design method we proposed earlier, by augmenting the measurement matrix and the sparse vector to recover, we can easily recover the sparse vector together with the errors, if any, in the hub measurements. We next show that under certain conditions, the statement is indeed true, and with the same way of measurement construction as we proposed earlier, one can recover sparse vectors with the presence of errors in hub measurements.

Since the recovery performance varies for different recovery methods, we focus on the ℓ_1-minimization method the widely used in Compressed Sensing. Given measurement matrix A and the measurements $\mathbf{y} = A\mathbf{x}$, ℓ_1-minimization returns vector $\mathbf{x}^*$ with the least ℓ_1-norm among all the vectors $\mathbf{x}'$ such that $A\mathbf{x}' = \mathbf{y}$ and uses $\mathbf{x}^*$ as an estimate the unknown vector $\mathbf{x}$. The following lemma provides the equivalent null space condition of successful sparse recovery via ℓ_1-minimization when the hub error exists.

Lemma 1. *Given the augmented matrix $A' = [A \quad -\mathbf{1}_m]$, ℓ_1-minimization successfully recovers k-sparse vectors $\mathbf{x}_S \in \mathcal{R}^n$ in the presence of some unknown error e_0 in the hub measurement if and only if for every non-zero vector $\mathbf{w}$ such that $A'\mathbf{w} = 0$, and for every set $T \subseteq \{1, ..., n\}$ with $|T| \leq k$, it holds that*

$$\|\mathbf{w}_T\|_1 + |w_{n+1}| \leq \|\mathbf{w}_{T^c}\|_1,$$

where $T^c = \{1, ..., n\} \backslash T$.

The recovery performance also varies for different measurement construction methods, here we consider the random measurement construction for complete graphs in which that every node is included in a measurement independently with probability 0.5, and every measurement is independent of each other. Mathematically, $P(A_{ij} = 1) = 0.5$ and $P(A_{ij} = 0) = 0.5$ independently for every i and j. Let the number of the randomly chosen measurements be $m = O(k\log(n/k))$, and we will choose the scaling constant later. We also add one row to A with all '1's, and it only increases the number of measurements by one. For such a measurement matrix A for complete graphs, we have the following result.

Theorem 4. *Given n nodes that can be measured freely via one hub, if the measurement matrix $A^{(m+1)\times n}$ with $m = O(k\log(n/k))$ for a complete graph has one row of all '1's, and every other entry independently takes value '1' or '0' with equal probability, then with probability at least $1 - O(n^{-\alpha})$ for some constant $\alpha > 0$, ℓ_1-minimization can successfully recover all k-sparse vectors in $\mathcal{R}^n$ even if the hub measurement is erroneous.*

The proof of Theorem 4 lies heavily on Lemma 2, so we first state it as follows.

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

$$(1-\delta)\|\mathbf{x}\|_2^2 \leq \|\Phi_S\mathbf{x}\|_2^2 \leq (1+\delta)\|\mathbf{x}\|_2^2. \tag{3}$$

Proof: Consider an m by n matrix Φ' with each entry taking value $\pm 1/\sqrt{m}$ with equal probability independently.

For every realization of matrix Φ', construct a matrix $\hat{\Phi}$ as follows. For every $i \in \{1, ..., m\}$ such that $\Phi'_{in} = 1/\sqrt{m}$, let $\hat{\Phi}_{ij} = -\Phi'_{ij}$ for all $j = 1, ..., n$. 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 $S \subseteq \{1, ..., n\}$,

$$\Phi'^{T}_{S} \Phi'_{S} = \hat{\Phi}^{T}_{S} \hat{\Phi}_{S}. \quad (4)$$

The Restricted Isometry Property in Compressed Sensing [26] states that for any $\delta > 0$, if $m \geq Ck \log n$ for some constant C and n is large enough, then with probability at least $1 - O(n^{-\alpha})$ for some constant $\alpha > 0$ such that for every set $S \subseteq \{1, ..., n\}$ with $|T| \leq 2k + 1$ and for every vector $\mathbf{x} \in \mathcal{R}^{2k+1}$,

$$(1 - \delta)\|\mathbf{x}\|_2^2 \leq \|\Phi'_S \mathbf{x}\|_2^2 \leq (1 + \delta)\|\mathbf{x}\|_2^2 \quad (5)$$

holds simultaneously.

Since (4) holds for all S, and $\|\Phi'_S \mathbf{x}\|_2^2 = \mathbf{x}^T \Phi'^{T}_{S} \Phi'_S \mathbf{x}$, then the above statement still holds if we replace Φ'_S with $\hat{\Phi}_S$ in (5). Since $\hat{\Phi}$ and Φ follow the same probability distribution, the lemma follows. ■

As the number of nodes in a network goes to infinity, if the number of groups of nodes that can be measured together via some hub remains constant, and the number of nodes in each group goes to infinity, (one example of such network is $\mathcal{G}^4$,) then by applying Theorem 4 with a simple union bound, we know that with high probability ℓ_1-minimization can successfully recover all the sparse vectors even if the hub measurements are erroneous, provided that while designing the measurement matrix for a general graph based on hubs, we randomly generate matrices with each entry taking value '0' and '1' with equal probability independently as the measurement matrices for complete graphs.

VI. Conclusion

In this paper, we considered the applications of compressive sensing in networks. We formulated the graph-constrained compressive sensing problem and reviewed some recent research progress in designing graph-constrained measurement matrices for compressive sensing. We further compared both random and deterministic graph-constrained measurement matrices. The robustness of the deterministic construction was also studied.

VII. Acknowledgement

The research is supported by NSF under CCF-0835706.

References

[1] S. S. Ahuja, S. Ramasubramanian, and M. Krunz, "SRLG failure localization in all-optical networks using monitoring cycles and paths," Proceeding of INFOCOM 2008

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

[3] E. Candès, "Compressive sampling," *Proc. International Congress of Mathematics*, 3, pp. 1433-1452, Madrid, Spain, 2006.

[4] Y. Chen, D. Bindel, H. Song, and R. H. Katz, "Algebra-based scalable overlay network monitoring: algorithms, evaluation, and applications," in IEEE/ACM Trasanctions on Networking, 2007

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

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

[7] M. Coates, Y. Pointurier, and M. Rabbat, "Compressed network monitoring for IP and all-optical networks," ACM Internet Measurement Conference," San Diego, California, October, 2007. 12 pages, 28

[8] M. Coates, Y. Pointurier, and M. Rabbat, "Compressed network monitoring," IEEE Statistical Signal Processing Workshop, Madison, Wisconsin, August, 2007.

[9] D. Donoho, "High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension ", *Discrete and Computational Geometry* , 102(27), pp. 617-652, 2006, Springer.

[10] N.Duffield, "Network tomography of binary network performance characteristics," IEEE Transcations on Information Theory, 52(12):5373-5388,2006

[11] N. Harvey, M. Patrascu, Y. Wen, S. Yekhanin and V. Chan, "Non-adpative fault diagnosis for all-optical networks via combinatorial group testing on graphs," Proc. of IEEE INFOCOM 2007.

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

[13] N. Feamster and H. Balakrishnan."Detecting BGP configuration faults with static analysis," In Proceedings of the 2nd Symposium on Networked Systems Design and Implementation (NSDI), May 2005.

[14] J. Kleinberg."Detecting a network failure". In Proceedings of the 41st Annual Symposium on Foundations of Computer Science (FOCS), 2000.

[15] M. Laifenfeld and A. Trachtenberg, "Identifying codes and covering problems," IEEE Transaction on Information Theory,vol.54, No. 9, 2008

[16] H. Nguyen and P. Thiran,"Using end-to-end data to infer lossy links in sensor networks," in Proceedings of IEEE INFOCOM, 2006.

[17] H. Nguyen and P. Thiran,"The Boolean solution to the congested IP link location problem: theory and practice," Proceeedings of IEEE INFOCOM 2007.

[18] L. Qiu, P. Bahl, A. Rao and L. Zhou,"Troubleshooting multihop wireless networks," In Proceedings of ACM SIGMETRICS, June 2005.

[19] Y. Wen, V. Chan and L. Zheng, "Efficient Fault Diagnosis Algorithms for All-Optical WDM Networks with Probabilistic Link Failures (invited paper)," IEEE/OSA Journal of Lightwave Technology, Vol. 23, No. 10, October 2005, pp.3358-3371.

[20] Y. Zhang, M. Roughan, W. Willinger and L. Qiu, "Spatio-temporal compressive sensing and Internet traffic matrices," Proceedings of IEEE SIGCOMM, 2009.

[21] Y. Zhao, Y. Chen, and D. Bindel, "Towards unbiased end-to-end network diagnosis," IEEE/ACM Transactions on Networking, Decemeber 2009.

[22] W. Xu, E. Mallada and A. Tang, "Compressive sensing over graphs," Proceedings of IEEE INFOCOM, 2011.

[23] M. Wang, W. Xu, E. Mallada and A. Tang, "Sparse Recovery with Graph Constraints: Fundamental Limits and Measurement Construction," to appear in Proceedings of IEEE INFOCOM 2012.

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

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

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

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

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

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

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

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