Optimal Measurement Matrices for Neighbor Discovery

Arash Saber Tehrani
Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089, USA
email: saberteh@usc.edu

Alexandros G. Dimakis
Department of Electrical Engineering
University of Texas at Austin
Austin, TX 78713, USA
email: dimakis@austin.utexas.edu

Giuseppe Caire
Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089, USA
email: caire@usc.edu

***Abstract*—We study the problem of neighbor discovery in which each node desires to detect nodes within a single hop. Each node is assigned a unique signature known by all other nodes. The problem can be considered as a compressed sensing problem. We propose a explicit—non-random—construction for the signatures. Further, we suggest the basis pursuit to detect the neighbors and offer a guarantee for its performance. Specifically, we show that the average number of errors can be made arbitrary small as the number of nodes in the network grows. Our result does not depend on the density of the network, *i.e.*, how the average number of neighbors scales with respect to the total number of nodes.**

I. Introduction

The expansion of mobile services has given rise to distributed wireless networks. In these networks devices transmit data from one to another across multiple wireless links—hops—in a distributed fashion. Due to power and geometric constraints, each node can only communicate with—transmit and receive data to—a small number of nodes, *i.e.*, its neighbors. The process of determining these neighbors by a node is called neighbor discovery. Neighbor discovery is crucial and is the first task that must be performed in a network, since it makes running other operations such as scheduling or routing of data possible.

Many algorithms have been proposed for neighbor discovery. Most of these algorithms, such as “birthday protocol” [1], directional antenna neighbor discovery [2], and slotted random transmission and reception [3] rely on random access discovery. That is, the nodes choose to transmit or listen randomly in each time slot so that each node gets a chance to hear its neighbors. Of course, such random access schemes allow one successful transmission at a time and thus these methods require long discovery periods to be reliable.

Fast discovery of neighbors is essential especially in mobile networks. Thus it would be ideal to reduce the discovery period as much as possible. An interesting approach introduced in [4] and [5] is for all nodes to transmit their unique signatures simultaneously. Then each node detects the signatures that are present and mark the owners as its neighbors. This can be done using fast multiuser detectors which are well-understood in the context of code-division multiple access (CDMA). These works, however, do not address the problem of implementing coherent detection without training. Evidently, training for channel estimation before neighbor discovery is impossible.

A discovery method which is both fast and non-coherent seems necessary. Work in [6] and [7] offer such methods. Specifically, they introduce a simple non-coherent (energy) detection scheme which is based on on-off signaling. That is, the signature of users are random binary sequences of zeros and ones. The novel idea is that a node gets to listen to the channel when it is not transmitting, *i.e.*, the zeros of its signature. Each node knows the signature dictionary, *i.e.*, signatures of all nodes. When a node is not transmitting, it simply detects whether it receives any energy, and deduce which nodes are present as neighbors based on the signature dictionary. Let us point out that using an energy detector is the most natural receiver for neighbor discovery over non-coherent transmissions.

A key observation is that number of neighbors a node has is much smaller than total number of nodes in the network. Therefore neighbor discovery for each node is similar to sparse signal recovery and can be treated as a compressed sensing problem. Thanks to this observation, [6] and [7] suggest group testing to detect the neighbors. That is, when a node is not transmitting and does not detect any energy, it marks all the nodes whose signature is non-zero at that given slot as non-neighbors (if they were neighbors some signal power would be detected) This is possible since we assumed each node knows the signature dictionary.

In this paper, we consider a model similar to [6], [7], *i.e.*, non-coherent transmissions while nodes use energy detectors as receivers. For building the unique signatures of the nodes, we use the columns of the measurement matrices we recently built for compressed sensing [8]. We apply the progressive edge growth algorithm [9] to make these matrices. Further, instead of using simple group testing for discovery, we detect the neighbors by linear programming–also known as basis pursuit [10], [11].

The rest of this manuscript is organized as follows. First we discuss our contribution. Then in section II we introduce the system model and the constraints governing it. In section III we describe the compressed sensing problem and our method to assign node's signatures. Section IV discusses the group testing and the method used in [6]. Finally in sections V and VI we compare these methods theoretically and experimentally.

A. Our contribution

We make two main contributions in this paper. First, we introduce explicit construction for signatures of the nodes. For our construction, the probability of error in recovering neighbors under basis pursuit can be made arbitrary small as the total number of nodes grows. Note that by adjusting the design parameters for our construction—mainly the length of the signatures—we can obtain the same performance guarantee for different scales of the average number of neighbors with respect to the total number of nodes. We also show that performance of neighbor discovery with our signature dictionary is superior to that of random signatures under group testing. Secondly, to the best of our knowledge, this is the first work which provides performance guarantee for neighbor discovery under basis pursuit.

II. System Model

Let us briefly discuss the notation used in the sequel. Bold font capital letters such as $\mathbf{S}$ denote matrices. We denote the (i,j)th element of $\mathbf{S}$ by S_{ij}. Further, bold font small letters stand for vectors. For example $\mathbf{s^i}$ is the i^{th} column of matrix $\mathbf{S}$. Finally s_j is the j^{th} element of vector $\mathbf{s}$.

We consider a network with n nodes. Our goal is for each node to identify its neighbors. For that, we assign to each node a unique signature. The signatures are known by all nodes. As in standard CDMA, this assumption is easily fulfilled. To consider network with different densities, let each node i be a neighbor of another node j with probability p. We call p the "neighbor density". As a result, each node has $p \cdot (n-1)$ neighbors in average.

Let the $m \times n$ matrix $\mathbf{S}$ denote the signature matrix whose columns $\mathbf{s^j}$ for $j = 1, \cdots, n$ represent the signature of the j^{th} node. We consider the on-off signaling. Thus $S_{rj} = 1$ means that node j transmits on the r^{th} slot and $S_{rj} = 0$ means that node j stays quiet and receive. We assume a coarse system frame synchronization such that all nodes transmit their signatures simultaneously.

The receiver is simply an energy detector which reads the level of energy received. Energy at the receiver is offset by the thermal floor such that 0 corresponds to just noise. In practice, each slot consists of a some random-noise sequence and by definition a neighbor is such if its received power is sufficiently above the thermal floor (sufficient raise over thermal). Since statistical fluctuations are essentially negligible given the fact that the energy levels are integrated over many chips per slot, a noiseless model where non-negative energy levels add up (non coherent combining) becomes relevant in this context. Thus, reading any energy by the receiver means the presence of transmission by at least one of the neighbors. Further, we assume that the energy of the signal received by a node from the non-neighbors is too low to be detectable. We leave the case which consider interference from non-neighbors as a future direction. Further, let all the nodes transmit with the same power γ. Thus the neighbor relation is reciprocal. That is, node i can transmit data to node j if and only if node j can transmit data to node i. Further we let h_{ij} denote the channel coefficients.

As discussed, we consider the half-duplex condition, where node i can only receive signal when it does not transmit. That is, node i receives on slots r for which $S_{ri} = 0$. Thus, the signal received by the i^th node is

$$\mathbf{y^i} = \mathbf{S^i}\mathbf{H^i}\gamma\mathbf{x^i}, \tag{1}$$

where $\mathbf{y^i}$ is the signal received by the i^{th} node, $\mathbf{S^i}$ is the same as $\mathbf{S}$ with the difference that the rows r with $S_{ri} = 1$ are omitted, $\mathbf{H^i}$ is a diagonal matrix of power of channel coefficients whose elements are $H^i{}_{jj} = h_{ij}^2$, γ is the transmission power of the nodes, and $\mathbf{x^i} \in \{0,1\}^n$ is the vector of neighbor indices of node i. That is $\mathbf{x^i_j} = 1$ if node j is a neighbor of node i and $\mathbf{x^i_j} = 0$ otherwise. Since a node cannot be its own neighbor, we set $\mathbf{x^i_i} = 0$. Let $\mathbf{v^i} = \mathbf{H^i}\gamma\mathbf{x^i}$. Note that for neighbor density p, $\mathbf{x^i}$ in average has $(n-1)p$ non-zero entries.

For a node that wishes to detect its neighbors, two kinds of errors are possible: *miss* and *false-alarm*. If an actual neighbor is missed in the detection, it is called a miss. On the other hand, if a non-neighbor is falsely declared as a neighbor, it is called a false-alarm.

For $m < n$, (1) is an under-determined system of linear equations. Number of neighbors, however, are much smaller than the number of nodes and thus $\mathbf{x^i}$ is a sparse vector. Thanks to sparsity of $\mathbf{x^i}$, we can recover it by much fewer measurements, *i.e.*, the length of the signatures m can be much smaller than the number of nodes n.

Different methods have been suggested to solve neighbor discovery problem. Here we consider group testing and compressed sensing.

III. Compressed Sensing

The simplest noiseless compressed sensing (CS) problem for exactly sparse signals consists of recovering the sparsest real vector $\mathbf{v}$ of a given length n, from a set of m real-valued measurements $\mathbf{y}$, given by $\mathbf{A} \cdot \mathbf{v} = \mathbf{y}$; namely

$$\begin{array}{lll} \textbf{CS-OPT:} & \text{minimize} & \|\mathbf{v}\|_0 \\ & \text{subject to} & \mathbf{A} \cdot \mathbf{v} = \mathbf{y}. \end{array}$$

As is well-known, ℓ_0 minimization is NP-hard, and one can relax **CS-OPT** by replacing the ℓ_0 norm with ℓ_1. Specifically,

$$\begin{array}{lll} \textbf{CS-LPD:} & \text{minimize} & \|\mathbf{v}\|_1 \\ & \text{subject to} & \mathbf{A} \cdot \mathbf{v} = \mathbf{y}. \end{array}$$

This LP relaxation is also known as basis pursuit. A fundamental question in compressed sensing is under what conditions the solution given by **CS-LPD** equals (or is very close to, in the case of approximately sparse signals) the solution given by **CS-OPT**, *i.e.*, the LP relaxation is tight. There has been a substantial amount of work in this area, see *e.g.* [12]–[15].

It has been shown [10]–[12] that for recovering k-sparse signals—vectors with at most k non-zero entries—at least $m = \Theta(k \log(n/k))$ measurements are required. Recently, we offered an explicit construction for measurement matrices [8] which matches this bound. That is, our measurement matrices are optimal. We make these matrices by progressive edge growth method [9]. Our measurement matrices have the null-space property and recover sparse signals with probability arbitrary close to 1. We direct the interested reader to [8] for more details.

Consider a bipartite graph with two sets of nodes, namely, m measurement nodes and n object nodes. Each measurement node r for $r = 1, \cdots, m$ is connected to all object nodes that it measures, *i.e.*, $\{j : S_{rj} = 1\}$. Further, let all measurement nodes (object nodes) have the same degree d_r (d_ℓ). Let g be the length of the shortest cycle of this graph. The PEG method [9] builds such bipartite graphs with long girth and then maps each measurement node to a row of the measurement matrix whose entries belonging to adjacent objects is set to one. It can be shown that as n increases, the girth scales $O(\log(n)/\log(d_r d_\ell))$. In [8] we construct a optimal measurement matrix with the following parameters

$$\begin{aligned} m &= c \cdot np\log(1/p) \\ d_r &= c_r \cdot 1/p \\ d_\ell &= c_\ell \cdot \log(1/p), \end{aligned} \tag{2}$$

where c, c_r, c_ℓ are constants which must satisfy $c_\ell = c \cdot c_r$, and p is the neighbor density. Thus, for our construction the girth scales $g = O(\log(n)/\log(1/p))$.

The neighbor discovery problem, however, is slightly different from the normal compressed sensing problem due to the half-duplex nature of transmitters. Specifically, the node i can receive its neighbors' signatures only when it is not transmitting. these are slots r for which $S_{ri} = 1$. As a result, compare with the classical compressed sensing problem, here each node misses d_ℓ measurements, so we need to find a way to compensate for them or show that this does not affect the performance of recovery.

We claim, however, that this does not affect the recovery process, since—due to the long girth of our construction—each node loses at most one of its measurements after omitting the measurements E_i. Assume there exists a node j that had at least two of its measurements removed, namely, $S_{r_1 j}$ and $S_{r_2 j}$ due to omitting E_i. The fact that $r_1, r_2 \in E_i$, however, means that $S_{r_1 i} = S_{r_2 i} = 1$. Thus, a cycle of length four consists of nodes i, r_1, j, and r_2 exists in the graph which contradicts the nature of our construction. As a result, our claim is valid and after the elimination of rows with indices in E_i, each object is measured at least $d_r - 1$ times and the proof presented in [8] goes through unchanged.

Let us present a bound from [8] on the average number of errors—missed and false alarms—that the basis pursuit makes when our construction is used.

Lemma 1: Consider a network of n nodes with neighbor density p. Let $\mathbf{S}$ be a signature dictionary constructed according to (2). Then there exists constant $0 < \delta < 1$ for which the average number of errors is upper bounded as

$$\mathcal{E}_{CS} \leq (n-1)\delta^{d_\ell (d_\ell - 1)^{T-1} - d_\ell}, \tag{3}$$

where T is any integer less than $g/4$.
The coefficient $(n-1)$ in (3) is due to taking union bound over $(n-1)$ possible neighbor nodes in the network.

IV. Group Testing

Neighbor discovery with on/off signatures is analogous to the classical problem of group testing. The group testing problem is to determine defective items among a set of objects by testing different subsets of the objects. The aim is to identify all the defective items with fewest number of measurements.

Algorithm 1 Group testing

```
Input: $\mathbf{y^i}$, $\mathbf{S}$
Initialize: $V \leftarrow \{1, 2, \cdots, n\} \setminus \{i\}$
for $r = 1$ to $m$ do
    if $\mathbf{y^i_r} = 0$ then
        $V \leftarrow V \setminus \{j : S_{r,j} = 1\}$
    end if
end for
Output: mark all nodes in $V$ as neighbors
```

Translating into our language, each row r of $\mathbf{S}$ is considered as a test which measures a subset of nodes $\{j : S_{rj} = 1\}$. Note that our system (1) is noiseless. Thus, if a receiver does not detect any energy, *i.e.*, $\mathbf{y^i_r}$ is zero, then all the nodes for which $S_{rj} = 1$ are non-neighbors. This procedure which is described in Algorithm 1 was suggested in [6]. Algorithm 1 represents a simple suboptimal solution that does not take into account the value of the non-zero components of $\mathbf{y^i}$, but only the location of the zero components. As for the signature dictionary, Zhang *et al.* [6] consider random signatures, *i.e.*, the elements of $\mathbf{S}$ to be i.i.d. Bernoulli random variables with $\mathbb{P}(S_{i,j} = 1) = q$ for all i and j.

Note that, due to the noiseless nature of our system, it is impossible for Algorithm 1 to miss a neighbor. The algorithm, however, may falsely declare some non-neighbors as neighbors. A node j is falsely declared as a neighbor under Algorithm 1 when for all rows r for which node j is transmitting, *i.e.*, $\{r : S_{rj} = 1\}$, at least one other neighbor of node i is also transmitting.

The next lemma from [6] bound the probability of false alarms denoted by $\mathcal{E}_{GT}$ for Algorithm 1.

Lemma 2: Consider a network of n nodes with neighbor density p. If the elements of signature matrix $\mathbf{S}$ are independently Bernoulli variables with parameter q, then the probability of false alarms averaged over all possible realizations of $\mathbf{S}$ and $\mathbf{x^i}$ is upper bounded as

$$\mathcal{E}_{GT} \leq (n-1)e^{-qm(1-(n-1)pq)} \tag{4}$$

The proof can be found in [6].

V. Compressed Sensing versus Group Testing

In this section we compare the performance of group testing over random signatures with compressed sensing over our construction. We first compare the performance of the two based on their average number of errors. Then we analyze the computational complexity of the two algorithm.

A. Average number of errors

Intuitively, we expect compressed sensing to perform better than group testing. Let us present an example to highlight the difference of the two. Consider a network with 4 nodes where nodes 1 and 2 are neighbors of node 0 while node 3 is not. We want to find the neighbors of node 0, which are nodes 1 and 2. Let the following be the signature of the nodes 1, 2, and 3:

$$\mathbf{s_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{s_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{s_3} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Further assume that the channel matrix is $\mathbf{H} = diag(0.7, 0.5, 0)$. Thus, after all nodes transmit, node 0 receives $\mathbf{y^0} = [0.7, 1.2, 0.5]^t$. Apparently, under group testing node 3 is also recognized as a neighbor. This, however, would not occur if we use basis pursuit since the system of linear equations is full rank.

To compare the error performance of the two, we need to analyze the bounds (3) and (4) in the limit, *i.e.*, as n grows. Small length of the signatures is clearly beneficial since it decrease the discovery period. Thus, we consider number of measurements m to be optimal, *i.e.*, $m = cnp\log(1/p)$. Let us remind the reader that it is provably infeasible to recover order of np neighbors with the number of measurements smaller than $\Theta(np\log(1/p))$.

Starting with $\mathcal{E}_{GT}$, after inserting $m = cnp\log(1/p)$ in (4) we get

$$\mathcal{E}_{GT} \leq ne^{-cnpq(1-npq)\log(1/p)}$$

For the exponential part to be decreasing as $n \to \infty$, we require $npq < 1$. The above bound is minimized for $q = \frac{1}{2np}$ whcih gives

$$\mathcal{E}_{GT} \leq ne^{-\frac{c}{4}\log(1/p)} = np^{c/4}$$

On the other hand, as shown in [8], using our construction (2), and choosing the right values for c, c_r, and c_ℓ, $\mathcal{E}_{CS}$ becomes a polynomial function of $1/n$. As shown in Fig. 1 our measurement matrices under basis pursuit performs strictly better than group testing over random zero-one matrices. Further, as the neighbor density p increases, the performance of group testing deteriorate while the performance of our method remains unchanged. The worst case for group testing occurs when expected number of neighbors scale linearly in n for which number of errors also grows linearly in n.

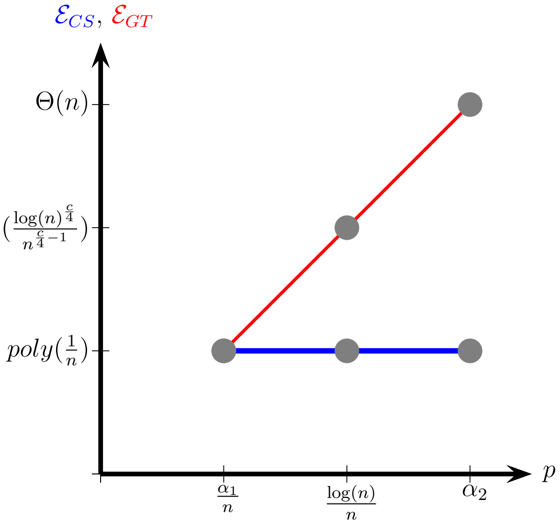

Fig. 1: Comparing the average number of errors of basis pursuit $\mathcal{E}_{CS}$ with group testing $\mathcal{E}_{GT}$ for different neighbor density p. $0 < \alpha_1 < 1$ is any constant.

B. Computational complexity

Here we compare the computational complexity of the two method. As shown in [6], Algorithm 1 has complexity $O(qmn)$. On the other hand, basis pursuit **CS-LPD** is a linear program which has complexity $O(n^{3.5})$ [16]. Both methods have complexity polynomial in n. The complexity of group testing, however, is much smaller than basis pursuit.

Group testing saves on computational power on the cost of having more errors. Thus, it makes sense to apply group testing for neighbor discovery if the nodes in the network have low computational power. On the other hand, if high accuracy in detection of the neighborhood of nodes—due to network tasks such as routing and scheduling—is required, then it is reasonable to use basis pursuit.

VI. Experimental Results

In this section we try to confirm our theoretic results with simulations. We consider a network with 2000 nodes and three cases $np = 20, 100$, and 200. This allows us to analyze the performance of the recovery algorithms as expected number of neighbors grow. In the first experiment, for each value of np we consider a different length of signatures, namely, $m = 200, 400$, and 800. To compare our work woth Zhang *et al.* [6], we build random

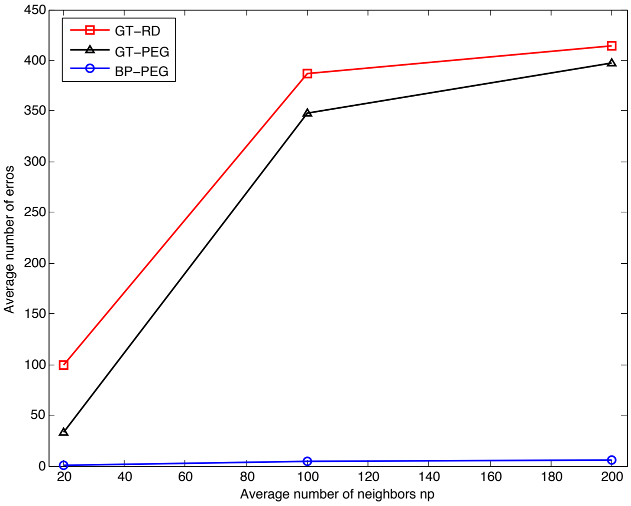

Fig. 2: Comparing the average number of errors of GT-RD, GT-PEG, and BP-PEG over a network of 2000 nodes for different values of $np = 20, 100, 200$ with varying length of signatures $m = 200, 400, 800$.

dictionaries with $q = \frac{1}{40}, q = \frac{1}{100}$, and $q = \frac{1}{200}$ for different values of np and used group testing as the detector. This algorithm is denoted by GT-RD. We also used group testing detector with our construction which is denoted by GT-PEG. Finally, we considered our construction (2) with basis pursuit detector which we called BP-PEG.

As shown in Fig. 2, our construction with basis pursuit BP-PEG performs much better than other two algorithms which also confirms the theoretical result. Although, to analyze group testing fairly, we must state that it requires large length of signatures to perform well.

In the second experiment, we considered the same network but this time we chose the signature lengths to be $m = 1000$. As shown in Fig. 3, although group testing over random signature dictionary performs poorly, it performs reasonably well on our construction. Performance of basis pursuit, however, is still superior.

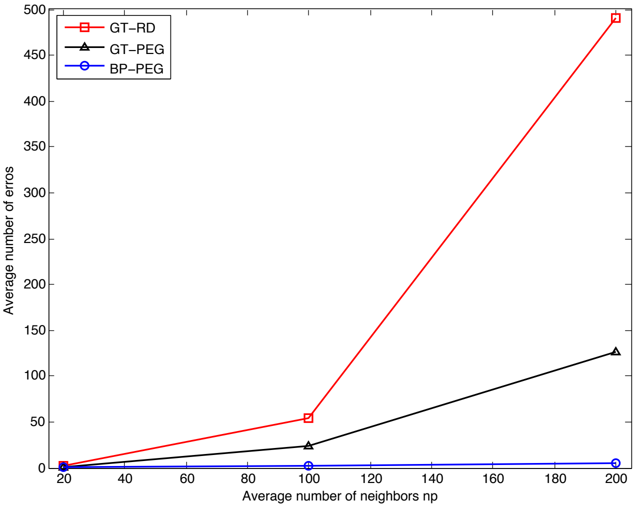

Fig. 3: Comparing the average number of errors of GT-RD, GT-PEG, and BP-PEG over a network of 2000 nodes for different values of $np = 20, 100, 200$ with fixed length of signatures $m = 1000$.

References

[1] M. McGlynn and S. Borbash, “Birthday protocols for low energy deployment and flexible neighbor discovery in ad hoc wireless networks,” in *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing*. ACM, 2001, pp. 137–145.

[2] S. Vasudevan, J. Kurose, and D. Towsley, “On neighbor discovery in wireless networks with directional antennas,” in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 4. IEEE, 2005, pp. 2502–2512.

[3] S. Borbash, A. Ephremides, and M. McGlynn, “An asynchronous neighbor discovery algorithm for wireless sensor networks,” *Ad Hoc Networks*, vol. 5, no. 7, pp. 998–1016, 2007.

[4] D. Angelosante, E. Biglieri, and M. Lops, “Neighbor discovery in wireless networks: a multiuser-detection approach,” *Physical Communication*, vol. 3, no. 1, pp. 28–36, 2010.

[5] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, “Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks,” in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 514–521.

[6] J. Luo and D. Guo, “Neighbor discovery in wireless ad hoc networks based on group testing,” in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*. IEEE, 2008, pp. 791–797.

[7] L. Zhang, J. Luo, and D. Guo, “Neighbor discovery for wireless networks via compressed sensing,” *Performance Evaluation*, 2012.

[8] A. Saber Tehrani, A. Dimakis, and G. Caire. Optimal deterministic compressed sensing matrices. [Online]. Available: http://www-scf.usc.edu/~saberteh/Docs/SublinearRecovery

[9] X. Hu, E. Eleftheriou, and D. Arnold, “Regular and irregular progressive edge-growth tanner graphs,” *Information Theory, IEEE Transactions on*, vol. 51, no. 1, pp. 386–398, 2005.

[10] E. Candes and T. Tao, “Decoding by linear programming,” *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, 2005.

[11] D. Donoho, “Compressed sensing,” *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[12] D. Donoho and J. Tanner, “Neighborliness of randomly projected simplices in high dimensions,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, p. 9452, 2005.

[13] D. Donoho and X. Huo, “Uncertainty principles and ideal atomic decomposition,” *Information Theory, IEEE Transactions on*, vol. 47, no. 7, pp. 2845–2862, 2001.

[14] M. Stojnic, W. Xu, and B. Hassibi, “Compressed sensing-probabilistic analysis of a null-space characterization,” in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, 2008, pp. 3377–3380.

[15] A. Gilbert and P. Indyk, “Sparse recovery using sparse matrices,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 937–947, 2010.

[16] N. Karmarkar, “A new polynomial-time algorithm for linear programming,” in *Proceedings of the sixteenth annual ACM symposium on Theory of computing*. ACM, 1984, pp. 302–311.