Support recovery in compressive sensing for estimation of Direction-Of-Arrival

Zhiyuan Weng and Xin Wang
Department of Electrical and Computer Engineering
Stony Brook University

***Abstract*—In the estimation of direction-of-arrival (DOA) problem, traditional array signal processing techniques normally use linear arrays sampled at Nyquist rate, and the inter-element distance in the linear array is required to be less than or equal to half of the wavelength to avoid angular ambiguity. The emerging Compressive Sensing(CS) theory enables us to use random array to sample the signal at much lower rate and still be able to recover it. To use this theory, the spatial signal should be sparse and it is always the case in practice. In this paper, we propose to apply the compressive sensing theory to reduce the spatial samples, i.e., to reduce the number of antenna elements. Instead of only showing the benefit of using CS theory, we analyze the performance of the angular estimation using the random array, i.e., we analyze the performance when the measurement is Fourier ensemble in terms of support recovery. We provide the sufficient and necessary conditions for the reliable support estimation.**

I. Introduction

Array signal processing has long been a research interest due to its wide application in fields such as sonar, radar and wireless communication. Direction-of-arrival(DOA) estimation is one of the main problems in array signal processing. There are many existing methods for the estimation of DOA using antenna arrays. Minimum variance distortionless response(MVDR) and multiple signal classification(MUSIC) algorithms are two popular ones[1]. These traditional methods often require sampling the received signal at the Nyquist-rate, which might be expensive especially for applications that need a large number of antenna elements such as interferometric array used for radio astronomy.

With the recent rapid advances of Compressive Sensing(CS) theory[2], the signal can be sampled at a much lower rate while it is still able to recover the signal, as long as the signal is sparse in some domain. In the DOA estimation problem, there are usually a limited number of sources, which means the signal is sparse in the angular domain. This naturally brings compressive sensing into the DOA estimation problem. Some researchers have already applied CS to DOA estimation[3], [4]. However, none of existing studies investigated the performance of CS for DOA estimation quantitatively. In this paper, we focus on the theoretical performance of DOA estimation using compressive sensing. Different from [3], [5] which exploit the sparsity of the received signals to reduce the number of temporal-domain samples similar to the majority of the work in the field, we exploit CS to reduce the number of samples in the spatial domain or to improve the angular domain resolution using a given number of antenna elements. Instead of using a uniform linear array (ULA)[4], we consider a random linear array (RLA) where the array is linear and the elements of the array are randomly distributed following the uniform distribution within a distance range.

In order to find the DOA through CS, the bearing space will be discretized into N distinct bearing angles, or called N grids. A vector signal $\boldsymbol{\theta}$ with length N can be formulated, with each item of the signal corresponding to the amplitude of the signal strength in each of the N directions. As mentioned above, with a small number of sources, $\boldsymbol{\theta}$ will be sparse with many elements to be zero. The determination of which elements of an unknown sparse signal are non-zero based on a set of noisy linear observations is called support recovery. Obviously, from the perspective of compressive sensing, the problem of DOA estimation can be considered as a support recovery problem. The measurement matrix is Fourier ensemble, i.e. the rows of the measurement matrix are drawn randomly from the rows of a Discrete Fourier Transform (DFT) matrix. In DOA estimation problem, we are more interested in finding the support thus the corresponding angles than the actual signal value corresponding to the support. There are some existing efforts on the problem of the exact support recovery [6], [7], where the information theoretic limits are given for the exact support recovery. The error of the support recovery, however, is simply represented as 0 if there is no error and 1 otherwise. For the application of antenna arrays for DOA estimation, the ℓ_2-norm support error is of more interest. Inspired by [8] in which the support recovery for generic measurement matrix was analyzed, in this work, we study the support recovery problem when the measurement matrix consists of Fourier ensembles. The Fourier ensembles are more structured than general random matrix. Also the matrix is in the complex field, thus we need to consider complex noise.

The paper is organized as follows. In Section II. we formulate the problem of DOA estimation using the random array. In Section III, we derive the Hammersley-Chapman-Robbins(HCR) bound for complex values. In Section IV, we analyze the unbiased property of the Maximum Likelihood(ML) support estimator and derive the sufficient condition for support recovery. Section V concludes the work.

II. Problem Formulation

In this paper, we consider the problem of DOA estimation using a random linear array. For the same array size, a random linear array whose antenna elements are randomly distributed within a given distance could potentially use much fewer number of antennas compared with a conventional uniform linear array. Let the number of antenna elements be M. In other words, M is the number of spatial samples. Suppose the received signal is $\mathbf{y} = [y_1, y_2, \cdots, y_M]^T$, where y_m is the signal received by the mth antenna. The receiving steering vector can be represented as

$$\mathbf{a}(\alpha) = \left[e^{j2\pi f_0 \tau_1(\alpha)}, \ldots, e^{j2\pi f_0 \tau_M(\alpha)}\right]^{\mathrm{T}} \tag{1}$$

where $\tau_i(\alpha)$ is the signal impinging time delay difference between each antenna element to the reference element when the target is at the direction α and f_0 is the frequency of the carrier corresponding to the signal received, and $\alpha \in [0, \pi]$. We set a virtual element located on the origin as the reference and assume its $\tau_{virtual}(\alpha)$ is zero. Suppose there are K targets. Then $\mathbf{y}$ could be written as

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{a}(\alpha_k)\eta_k + \mathbf{z}, \tag{2}$$

where

$$\mathbf{z} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}). \tag{3}$$

η_k denotes the amplitude of the signal from the kth target and α_k denotes the direction of the kth target. We discretize the angular domain $[0, \pi]$ into N grids where $N \gg M$, and assume the targets only lie within the N grids. The amplitude of the signal strength from each direction(i.e., corresponding to one of the grids) forms the signal $\boldsymbol{\theta} \in \mathbb{C}^N$. Since there are K targets, the number of non-zero elements in $\boldsymbol{\theta}$ is K. We say that $\boldsymbol{\theta}$ is K sparse if $K \ll N$. We use $\mathbf{s}$ to denote the indexes of the non-zero elements of $\boldsymbol{\theta}$.

$$\mathbf{s}(\boldsymbol{\theta}) \triangleq [n_1, n_2, \cdots, n_K]. \tag{4}$$

The corresponding non-zero entries of $\boldsymbol{\theta}$ are denoted as

$$\boldsymbol{\theta}_{\mathbf{s}} \triangleq [\theta_{n_1}, \theta_{n_2}, \cdots, \theta_{n_K}], \tag{5}$$

where we assume that $0 \le n_1 < n_2 < \cdots < n_K \le N-1$. Note that the angle domain $[0, \pi]$ is not equally discretized. Instead, we discretize a transformed variable u into N equal grids, where $u = \cos\alpha$ and $u \in [-1, 1]$. By doing this, we intend to keep the coherence value between every two adjacent columns of the measurement matrix (in equation (9)) equal to minimize the maximum coherence of the matrix and therefore reduce the number of measurements needed for the reliable recovery of the signal. Note that there is a trade-off between increasing the angular resolution and reducing the coherence parameter. For a given number of measurements, a higher resolution thus a larger N leads to a higher coherence value and thus reduces the probability of successful recovery. On the other hand, a lower resolution leads to a higher probability of successful recovery. Now equation (2) could be rewritten as

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{a}(\arccos\frac{2n_k - N}{N})\theta_{n_k}. \tag{6}$$

Let $c_n = n - N/2$, where the integer $n \in [0, N-1]$. Thus $c_n \in \{-N/2, -N/2+1, \cdots, N/2-1\}$. Further, we assume the distance between the mth antenna and the reference antenna is $\frac{\lambda}{2}d_m$, where λ is the wavelength of the carrier frequency and d_m is uniformly distributed in the interval $[0, D]$. This means $D\frac{\lambda}{2}$ is the array size. Clearly, $\lambda = \frac{c}{f_0}$ where c is the speed of light. Then equation (6) can be written as:

$$y_m = \sum_{k=1}^{K} \exp\left(j2\pi\frac{d_m c_k}{N}\right)\theta_{n_k}. \tag{7}$$

Put it into the matrix form and consider the noise, we obtain

$$\mathbf{y} = \mathbf{x} + \mathbf{z} \tag{8}$$

where $\mathbf{x} = \mathcal{F}_X\boldsymbol{\theta}$, $\mathcal{F}_X$ is called the measurement matrix and

$$(\mathcal{F}_X)_{m,n} = \exp\left(j\frac{2\pi}{N}c_n \cdot d_m\right), \quad 1 \le m \le M,\ 0 \le n \le N-1. \tag{9}$$

In practice, the array size $D\frac{\lambda}{2}$ is given and fixed. So the best choice for N is $N = D$ so that $-\frac{N^2}{2} \le c_n \cdot d_m \le \frac{N(N-1)}{2}$. This also makes $\mathcal{F}_X$ a partial Fourier matrix as desired. If $N > D$, the number of grids would become too large for the array to distinguish between grid cells. Intuitively, we also cannot make N arbitrarily large. Although a higher N would increase the angular resolution, it would also increase the probability of recovery failure and result in lower performance. On the other hand, if we set $N < D$, we can not fully exploit a given size of array to obtain a good resolution. In the following discussion, we assume $N = D$ so we only consider the effect of N. Our objective is to estimate the support of the signal $\mathbf{s}(\boldsymbol{\theta})$, i.e. the positions of the non-zero elements in the signal which correspond to the angles of the targets.

In the estimation theory, the Cramér-Rao bound is used to bound the variance of error of unbiased estimators. However in this work, we want to estimate values in a predetermined set. The possible positions are the finite N points. In such a case, the Hammersley-Chapman-Robbins(HCR) bound[9] provides a stronger lower bound on unbiased estimators.

III. Necessary Condition

In this section, we will provide the necessary condition for the support recovery. That is, we will find a lower bound on the number of measurements needed for the reliable support recovery. We first give the HCR bound for the unbiased estimation of the support, which helps to derive the necessary condition. Suppose $\boldsymbol{\delta}$ is the vector of unknown parameters and we are only interested in estimating the subset of the unknown parameters. Assume $\boldsymbol{\delta} = [\boldsymbol{\delta_1}, \boldsymbol{\delta_2}]$, where only the estimation of $\boldsymbol{\delta_1}$ is of interest but not the estimation of $\boldsymbol{\delta_2}$. Let $\hat{\boldsymbol{\delta_1}}(\mathbf{y})$ denote the unbiased estimator of $\boldsymbol{\delta_1}$. $\mathrm{P}(\mathbf{y};\boldsymbol{\delta})$ denotes the probability density function of $\mathbf{y}$ conditioning on $\boldsymbol{\delta}$. Then the HCR bound says the trace of the covariance matrix of $\hat{\boldsymbol{\delta_1}}(\mathbf{y})$ is bounded by

$$\mathrm{tr}\left[\mathrm{cov}(\hat{\boldsymbol{\delta_1}})\right] \ge \sup_{\boldsymbol{\delta'} \ne \boldsymbol{\delta}} \frac{\|\boldsymbol{\delta_1} - \boldsymbol{\delta'_1}\|^2}{\int_{\mathbb{C}^M} \frac{\mathrm{P}^2(\mathbf{y};\boldsymbol{\delta'})}{\mathrm{P}(\mathbf{y};\boldsymbol{\delta'})} d\mathbf{y} - 1} \tag{10}$$

in which $\boldsymbol{\delta'}$ is in a set specified according to the *a priori* information. In our case, $\boldsymbol{\delta_1}$ corresponds to the support of

$\boldsymbol{\theta}$, which is denoted by $\mathbf{s}$ here. It belongs to the parameter set of integer numbers. $\boldsymbol{\delta_2}$ corresponds to the values on the support. The density function $\mathrm{P}(\mathbf{y};\boldsymbol{\delta})$ is Gaussian. Note that if the support is known, the value on the support can be determined by minimum mean square error method and $\mathbf{x}$ can be obtained by projecting the noisy observation $\mathbf{y}$ into the subspace specified by the support. Consequently, we have the following theorem.

Theorem III.1. *Assume $\hat{\mathbf{s}}(\mathbf{y})$ is an unbiased estimator of the support $\mathbf{s}$, the HCR lower bound on the covariance of $\hat{\mathbf{s}}(\mathbf{y})$ is given by*

$$\mathrm{tr}\left[\mathrm{cov}(\hat{\mathbf{s}})\right] \geq \sup_{\mathbf{s}':\mathbf{s}'\neq\mathbf{s}} \frac{\|\mathbf{s}-\mathbf{s}'\|^2}{e^{2\|\mathbf{x}-\mathbf{x}'\|^2/\sigma^2}-1} \tag{11}$$

where $\mathbf{s}'$ and $\mathbf{x}'$ denote the support value other than $\mathbf{s}$ and the projection of $\mathbf{x}$ into the subspace specified by $\mathbf{s}'$, respectively.

Proof: Due to the complex gaussian noise, the likelihood probability can be written as

$$\mathrm{P}(\mathbf{y}|\mathbf{x}) = \frac{1}{\pi^n\sigma^{2n}}\exp\left(-\frac{\|\mathbf{y}-\mathbf{x}\|^2}{\sigma^2}\right). \tag{12}$$

Then we have

$$\begin{aligned}&\frac{\mathrm{P}^2(\mathbf{y}|\mathbf{s}')}{\mathrm{P}(\mathbf{y}|\mathbf{s})}\\ =&\prod_{i=1}^{M}\frac{1}{\pi\sigma^2}\exp\left(-\frac{(y_i-2x_i'+x_i)^2-2(x_i'-x_i)^2}{\sigma^2}\right)\end{aligned} \tag{13}$$

where M is the length of vector $\mathbf{y}$. Then the denominator of the right hand side of (10) becomes

$$\int_{\mathbb{C}^n}\frac{\mathrm{P}^2(\mathbf{y}|\mathbf{s}')}{\mathrm{P}(\mathbf{y}|\mathbf{s})}dy - 1 = \exp\left(\frac{2\|\mathbf{x}-\mathbf{x}'\|^2}{\sigma^2}\right)-1. \tag{14}$$

Therefore, we obtain the inequality (11). ■

The support recovery is considered to be *reliable* [8] if

$$\lim_{N\to\infty}\|\mathbf{s}(\boldsymbol{\theta})-\mathbf{s}(\hat{\boldsymbol{\theta}})\| = 0. \tag{15}$$

For unbiased estimators, this is equivalent to

$$\lim_{N\to\infty}\mathrm{tr}\left[\mathrm{cov}\left(\mathbf{s}(\hat{\boldsymbol{\theta}})\right)\right] = 0. \tag{16}$$

Next we will derive the necessary condition for the reliable support recovery. Intuitively, the recovery performance will depend on the value of the minimum element in $\boldsymbol{\theta}$, since the larger the value is, the easier it is to distinguish the measured signals from the noise. We assume the minimum element in $\boldsymbol{\theta}$ is θ_{min}. We need to use Bernstein's inequality[10] to prove the next theorem.

Lemma III.1. *(Bernstein's inequality) Let $Z_l, l = 1,\cdots,M$ be a sequence of independent real-valued random variables with mean zero and variance $\mathbb{E}[Z_l^2] \leq v$ for all $l = 1,...,M$. Assume that $|Z_l| \leq B$ almost surely. Then*

$$\mathbf{P}\left(\sum_{l=1}^{M}Z_l \geq x\right) \leq \exp\left(-\frac{1}{2}\frac{x^2}{Mv+Bx/3}\right) \tag{17}$$

Theorem III.2. *Let the measurement matrix be Fourier ensemble $\mathcal{F}_X \in \mathbb{C}^{M\times N}$. Let $\theta_{\min}$ denote the minimum none-zero entry of $\boldsymbol{\theta}$. Then a support recovery is unreliable if*

$$M < \max\left\{K, \frac{\sigma^2\log((N-K)^2K)}{4K^2|\theta_{\min}|^2}\right\} \tag{18}$$

Proof: Assuming the correct support is $\mathbf{s} = (0,1,2,...,K-1)$, the maximum ℓ_2-norm error will happen when the incorrect one is $\mathbf{s}' = (N-K,...,N-2,N-1)$. Thus $\|\mathbf{s}-\mathbf{s}'\|^2{=}(N-K)^2K$. On the other hand

$$\begin{aligned}\mathbf{x}-\mathbf{x}' &= \mathcal{F}_X(\boldsymbol{\theta}-\boldsymbol{\theta}') &(19)\\ &= \sum_{l=0}^{K-1}\theta_{n_l}(\phi_l-\phi_{N-K+l}) &(20)\\ &> K\theta_{\min}(\phi_l-\phi_{N-K+l}) &(21)\end{aligned}$$

where ϕ_l denotes the lth column of the measurement matrix $\mathcal{F}_X$. Taking the norm and multiplying both sides by $2/\sigma^2$, we have

$$\frac{2\|\mathbf{x}-\mathbf{x}'\|^2}{\sigma^2} > \frac{2K^2|\theta_{\min}|^2}{\sigma^2}\|\phi_k-\phi_l\|^2 \tag{22}$$

Note that it is impossible to achieve the reliable support recovery if the right hand side of (11), $\|\mathbf{s}-\mathbf{s}'\|^2/(e^{2\|\mathbf{x}-\mathbf{x}'\|^2/\sigma^2}-1)$ stays away from zero as $N\to\infty$. This implies that

$$\lim_{N\to\infty}\mathrm{P}\left(\frac{2\|\mathbf{x}-\mathbf{x}'\|^2}{\sigma^2} < \log((N-K)^2K)\right) = 1. \tag{23}$$

By (22) we have

$$\lim_{N\to\infty}\mathrm{P}\left(\|\phi_k-\phi_l\|^2 < \frac{\sigma^2\log((N-K)^2K)}{2K^2|\theta_{\min}|^2}\right) = 1. \tag{24}$$

Therefore the complement of the event

$$\lim_{N\to\infty}\mathrm{P}\left(\|\phi_k-\phi_l\|^2 > \frac{\sigma^2\log((N-K)^2K)}{2K^2|\theta_{\min}|^2}\right) = 0. \tag{25}$$

To see the distribution of $\|\phi_k-\phi_l\|^2$, first we have

$$\begin{aligned}&\|\phi_k-\phi_l\|^2\\ =&\sum_{i=1}^{M}\left|e^{j\frac{2\pi c_k d_i}{N}}-e^{j\frac{2\pi c_l d_i}{N}}\right|^2\\ =&\sum_{i=1}^{M}\left(2-e^{j\frac{2\pi(c_k-c_l)d_i}{N}}-e^{-j\frac{2\pi(c_k-c_l)d_i}{N}}\right)\\ =&\sum_{i=1}^{M}\left(2-2\cos\left(\frac{2\pi d_i(c_k-c_l)}{N}\right)\right).\end{aligned} \tag{26}$$

Suppose $u_i = \cos\left(\frac{2\pi d_i(c_k-c_l)}{N}\right)$ where d_i is uniformly distributed in $[0,D]$. Suppose $D' = D(c_k-c_l)$. To obtain the probability density function of u_i, we need to discuss two cases separately. The first case is when $kN \leq D' < (k+\frac{1}{2})N$ and the second case is when $(k+\frac{1}{2})N \leq D' < (k+1)N$.

It is not difficult to show that when D' thus k are large enough, the remainder $D'\%N$ plays a less role and can be ignored, and both cases can be approximated to

$$\mathrm{Var}[u_i] \approx \frac{1}{2} \tag{27}$$

$$\mathbb{E}[u_i] \approx 0 \tag{28}$$

When $D' = N$ or more generally $D' = kN$ ($D'\%N = 0$), the two cases would also have the exactly the same results. Intuitively, when $D = N$, the impact of random placement of antennas is equivalent to the case of random selection of rows from a DFT matrix. If $D' = kN$, this is equivalent to selecting rows from a matrix which is formed by concatenating k identical DFT matrices vertically. In the case when $D'\%N \neq 0$, as D' and k go to infinity, the effect will be similar to the case $D' = kN$, and therefore can be approximated by (27) and (28).

Combine (26), (27) and (28) using (17), we have

$$\begin{aligned} &\mathrm{P}\left(\sum_{i=1}^{M} u_i \geq \frac{\sigma^2 \log((N-K)^2 K)}{4K^2|\theta_{\min}|^2} - M\right) \\ &\leq \exp\left(-\frac{1}{2}\frac{\left(\frac{\sigma^2 \log((N-K)^2 K)}{4K^2|\theta_{\min}|^2} - M\right)^2}{\frac{M}{2} + \frac{1}{3}\left(\frac{\sigma^2 \log((N-K)^2 K)}{4K^2|\theta_{\min}|^2} - M\right)}\right) \end{aligned} \tag{29}$$

where B in (17) is set to 1 since $u_i \leq 1$, and v is 1/2 as derived above. Therefore according to (25), to make the right hand side of (29) go to zero, the exponential component must go to negative infinity. So

$$\lim_{N\to\infty}\left(\frac{1}{2}\frac{\left(\frac{\sigma^2 \log((N-K)^2 K)}{4K^2|\theta_{\min}|^2} - M\right)^2}{\frac{M}{2} + \frac{1}{3}\left(\frac{\sigma^2 \log((N-K)^2 K)}{4K^2|\theta_{\min}|^2} - M\right)}\right) = \infty. \tag{30}$$

Therefore, if

$$M < \frac{\sigma^2 \log((N-K)^2 K)}{4K^2|\theta_{\min}|^2}, \tag{31}$$

reliable recovery is not possible. This completes the proof. ∎

Obviously, it is a necessary condition for the reliable support recovery. In another word, if the support can be reliably recovered, the following condition will be satisfied:

$$M \geq \frac{\sigma^2 \log((N-K)^2 K)}{4K^2|\theta_{\min}|^2}. \tag{32}$$

This provides the lower bound on the number of measurements M necessary for the successful support recovery. From the derivation above, when there is a need to increase N for a higher angular resolution or the signals are sparse with a smaller K or when there are weak signal sources, the spatial number of samples or equivalently the number of antenna elements need to be large. Note, the objective here is to estimate the support instead of the values on the support. As an extreme example, if $K = N$, any estimation of K will be correct since $n_1, n_2, \cdots, n_K$ in (4) is arranged in ascending order. In a practical situation, $K \ll N$, thus K has less impact on M than N.

IV. Sufficient Condition

Before analyzing the sufficient condition for the reliable support recovery, we first consider the performance of the Maximum Likelihood (ML) estimator for ℓ_2-norm support recovery, and find the condition for the estimator to be unbiased. The ML estimator should be equivalent to the least squares solution in the Gaussian noise setting. That is

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\min_{s:|s|=K} \|\mathbf{y} - P_{\mathbf{s}}\mathbf{y}\|_2 \tag{33}$$

where $P_{\mathbf{s}}$ indicates the projection matrix which projects the signal to the subspace spanned by the columns whose positions are $\mathbf{s}$. We use $\mathbf{s}'$ to denote the support other than $\mathbf{s}$ and $\mathbf{F_s}$ to denote the subspace spanned by $\mathbf{s}$. We say the estimation is unbiased if $\mathrm{E}(\hat{\mathbf{s}}) = \mathbf{s}$. As long as the ML estimator becomes unbiased, the HCR bound will be achievable. The following results are derived along the same lines as [8]. We provide them here for the completeness of the work.

Lemma IV.1. *Let* $\mathbf{y} = \mathbf{x} + \mathbf{z}$*, where* $\mathbf{x} = \mathcal{F}_X\boldsymbol{\theta} \in \mathbf{F_s}$, $\mathbf{z} \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$ *and* $\mathbf{s}'$ *be a support vector that is different from* $\mathbf{s}$*. Then the probability that ML estimator selects the incorrect support, which is denoted by* $\mathrm{P}_{\mathrm{ML}}(\mathbf{s}')$ *satisfies*

$$\mathop{\mathrm{P}}_{\mathrm{ML}}(\mathbf{s}') < \mathrm{P}\left(\|\mathbf{z}\| \geq \frac{\|\mathbf{x} - P_{\mathbf{s}'}\mathbf{x}\|}{2}\right) \tag{34}$$

Proof: ML chooses $\mathbf{s}'$ over s if and only if

$$\min_{\mathbf{t}'\in\mathbf{F}_{\mathbf{s}}'} \|\mathbf{y} - \mathbf{t}'\| = \min_{\mathbf{t}\in\mathbf{F_s}} \|\mathbf{y} - \mathbf{t}\| \tag{35}$$

Assume that $\|\mathbf{z}\| < \|\mathbf{x} - P_{\mathbf{s}'}\mathbf{x}\|/2$. Then for any $\mathbf{t}' \in \mathbf{F_s}$, we have

$$\begin{aligned} \|\mathbf{y} - \mathbf{t}'\|^2 &= \|\mathbf{x} - \mathbf{t}' + \mathbf{z}\|^2 && (36)\\ &\geq \|\mathbf{z}\|^2 + \|\mathbf{x} - \mathbf{t}'\|^2 - 2\|\mathbf{x} - \mathbf{t}'\|\|\mathbf{z}\| && (37)\\ &> \|\mathbf{z}\|^2 && (38)\\ &= \|\mathbf{y} - \mathbf{x}\|^2 && (39)\\ &\geq \min_{\mathbf{t}\in\mathbf{F_s}} \|\mathbf{y} - \mathbf{t}\|^2. && (40) \end{aligned}$$

∎

Lemma IV.2. *Suppose* $r = \dfrac{\|\mathbf{x} - P_{\mathbf{s}'}\mathbf{x}\|^2}{2\sigma^2}$*, then*

$$\mathop{\mathrm{P}}_{\mathrm{ML}}(\mathbf{s}') < e^{-r/2}\sum_{t=0}^{M-1}\frac{(r/2)^t}{t!} \tag{41}$$

Proof: The random variable $\dfrac{\|\mathbf{z}\|^2}{\frac{1}{2}\sigma^2}$ is distributed according to the chi-square distribution with $2M$ degrees of freedom. By using the cdf of the Chi-square distribution, we obtain

$$\mathop{\mathrm{P}}_{\mathrm{ML}}(\mathbf{s}') < 1 - \frac{\gamma(M, r/2)}{\Gamma(M)} \tag{42}$$

where $\Gamma(M)$ is the Gamma function and $\gamma(M,x)$ is the lower incomplete Gamma function. Then we have

$$\frac{\gamma(M,r/2)}{\Gamma(M)}=\exp(-r/2)\sum_{t=M}^{\infty}\frac{(r/2)^t}{t!} \tag{43}$$

By using Taylor expansion $e^{r/2}=\sum_{t=0}^{\infty}\frac{(r/2)^t}{t!}$, we obtain

$$\frac{\gamma(M,r/2)}{\Gamma(M)}=1-\exp(-r/2)\sum_{t=0}^{M-1}\frac{(r/2)^t}{t!} \tag{44}$$

which completes the proof. ■

Lemma IV.3. *Let $r=\alpha M$ for some constant $\alpha>1$. Then we have*

$$\mathrm{P}(\mathbf{s}')<\frac{r}{\alpha}c(\alpha)^{-r} \tag{45}$$

where

$$c(\alpha)=\frac{\exp\left(\dfrac{\alpha-2}{2\alpha}\right)}{\left(\dfrac{\alpha}{2}\right)^{1/\alpha}} \tag{46}$$

Proof:

$$\begin{aligned}
\underset{\mathrm{ML}}{\mathrm{P}}(\mathbf{s}') <& e^{-r/2}\sum_{t=0}^{M-1}\frac{(r/2)^t}{t!} && (47)\\
<& e^{-r/2}M\frac{(r/2)^M}{M!} && (48)\\
<& e^{-r/2}M\frac{(r/2)^M}{(M/e)^M} \quad \text{due to} \quad M!>(M/e)^M && (49)\\
=& \frac{r}{\alpha}\left(\frac{\exp\left(\dfrac{\alpha-2}{2\alpha}\right)}{\left(\dfrac{\alpha}{2}\right)^{1/\alpha}}\right)^{-r} && (50)
\end{aligned}$$

■

Let $d_{\min}\triangleq\min_{\mathbf{s}':\mathbf{s}'\neq s}\|\mathbf{x}-P_{\mathbf{s}'}\mathbf{x}\|$, $\beta=d_{\min}^2/2M\sigma^2$ and $r_{\min}=\beta M$. Then we have the following two theorems.

Theorem IV.1. *For $\beta>1$, the performance of the ML estimator is upper bounded as*

$$\mathrm{tr}\left(\mathrm{cov}(\hat{\mathbf{s}})\right)<\frac{KMN^2}{2}c(\beta)^{-r_{\min}} \tag{51}$$

where $c(\cdot)$ is defined by (46).

Theorem IV.2. *Under the conditions $M\geq(1+\varepsilon)\log(N)/\beta\log c(\beta)$, for some fixed $\varepsilon>0$ and β bounded away from 1, the ML estimator is asymptotically unbiased as $N\to\infty$.*

The proof of theorem IV.1 and IV.2 are exactly the same as in [8]. Theorem IV.1 provides an upper bound on the performance of the ML estimator and theorem IV.2 proves the unbiased property of the ML estimator, which indicates the error of the ML estimator is lower bounded by the HCR bound. The following theorem shows the sufficient number of measurements for reliable support recovery with the partial Fourier measurement matrix. We use $\Theta(\cdot)$ to indicate a function is bounded both above and below asymptotically. Formally, if $f(n)\in\Theta(g(n))$, then $|g(n)|\cdot k_1\leq|f(n)|\leq|g(n)|\cdot k_2$ for some constant k_1 and k_2 as $n\to\infty$. We use $o(\cdot)$ to indicate a function is dominated by another function asymptotically. That is if $f(n)\in o(g(n))$, then $|f(n)|\leq|g(n)|\cdot\epsilon$ for any ϵ as $n\to\infty$.

Theorem IV.3. *If the minimum value $\theta_{\min}=\Theta(1)$, then $M=\Theta(K\log(N-K))$ will be the sufficient number of measurements to ensure reliable ℓ_2-norm support recovery.*

The proof is omitted due to the limited space, and can be found in [11]. Sufficient conditions for different trends of K are listed in Table IV.

	Necessary conditions	Sufficient conditions
$K=\Theta(N)$ $\theta_{\min}=\Theta(\frac{1}{K})$	$\Theta(\log N)$	*
$K=\Theta(N)$ $\theta_{\min}=\Theta(1)$	*	$\Theta(N)$
$K=o(N)$ $\theta_{\min}=\Theta(\frac{1}{K})$	$\Theta(\log((N-K)^2K))$	*
$K=o(N)$ $\theta_{\min}=\Theta(1)$	$\Theta(\frac{\log((N-K)^2K)}{K^2})$	$\Theta(K\log(N-K))$

V. Conclusion

In this paper, we apply compressive sensing theory to reduce the number of antennas by using random linear array while achieving the similar performance of angular estimation as the whole uniform linear array. We analyze the necessary and sufficient conditions that the number of measurements should satisfy thus the number of antennas needed in order to achieve reliable support recovery for reliable estimation of angles.

References

[1] H.L. Van Trees, *Optimum array processing*, Wiley Online Library, 2002.

[2] E.J. Candès and M.B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[3] A.C. Gurbuz, J.H. McClellan, and V. Cevher, "A compressive beamforming method," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, 2008, pp. 2617–2620.

[4] I. Bilik, "Spatial compressive sensing approach for field directionality estimation.," in *Radar Conference, 2009 IEEE*. IEEE, 2009, pp. 1–5.

[5] V. Cevher, A.C. Gurbuz, J.H. McClellan, and R. Chellappa, "Compressive wireless arrays for bearing estimation," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, 2008, pp. 2497–2500.

[6] A.K. Fletcher, S. Rangan, and V.K. Goyal, "Necessary and sufficient conditions for sparsity pattern recovery," *Information Theory, IEEE Transactions on*, vol. 55, no. 12, pp. 5758–5772, 2009.

[7] M.J. Wainwright, "Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting," *Information Theory, IEEE Transactions on*, vol. 55, no. 12, pp. 5728–5741, 2009.

[8] A. Karbasi, A. Hormati, S. Mohajer, and M. Vetterli, "Support recovery in compressed sensing: An estimation theoretic approach," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 679–683.

[9] JM Hammersley, "On estimating restricted parameters," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 12, no. 2, pp. 192–240, 1950.

[10] "http://en.wikipedia.org/wiki/bernstein_ inequalities_ (probability_theory)," .

[11] Z. Weng and X. Wang, "Support recovery in compressive sensing for estimation of direction-of-arrival," Technical Report CEAS-836, Stony Brook University, Stony Brook, New York, 2011.