

# Recovery guarantees for multifrequency chirp waveforms in compressed radar sensing

Nithin Sugavanam and Emre Ertin  
The Ohio State University

**Abstract**—Radar imaging systems transmit modulated wideband waveform to achieve high range resolution resulting in high sampling rates at the receiver proportional to the bandwidth of the transmit waveform. Analog processing techniques can be used on receive to reduce the number of measurements to  $N$ , the number of potential delay bins. If the scene interrogated by the radar is assumed to be sparse consisting of  $K$  point targets, results from compressive sensing suggest that number of measurements can be further reduced to scale with  $K \log N$  for stable recovery of a sparse scene from measurements with additive noise. While unstructured random projectors guarantee successful recovery under sparsity constraints, they cannot be implemented in the radar hardware in practice. Recently, structured random Toeplitz and Circulant matrices that result from using stochastic waveforms in time delay estimation setting have been shown to yield recovery guarantees similar to unstructured sensing matrices. However, the corresponding transmitter and receiver structures have high complexity and large storage requirements. In this paper, we propose an alternative low complexity compressive wideband radar sensor which combines multitone signal chirp waveform on transmit with a receiver that utilizes an analog mixer followed with a uniform sub-Nyquist sampling stage. We derive the recovery guarantees for the resulting structured measurement matrix and sufficient conditions for the number of tones. The only random component of our design is the sparse tone spectrum implementable efficiently in hardware. Our analytical and empirical results show that the performance of our scheme is in par with unstructured random sensing matrices and structured Toeplitz and Circulant matrices with random entries.

**Index Terms**—Compressive sensing, mutual coherence, Structured measurement matrix, Linear Frequency modulated waveform, Radar.

## I. INTRODUCTION

Radar imaging systems acquire information about the scene of interest by transmitting pulsed waveforms and analyze the received backscatter energy to form an estimate of the range and amplitude of the reflectors in the scene. These range profiles from multiple pulses and/or multiple antenna elements can be processed jointly to solve a multitude of inference tasks including detection, tracking and classification [1]. In this paper, we focus on the problem of estimation of range and amplitude of reflectors in the scene using a single modulated wideband pulse  $\phi(t)$  of bandwidth  $B$ . The resolution of the echo imaging system is directly proportional to the bandwidth of the transmitted signal. Assuming the support of the observed delays are known to lie on an interval  $T$  (termed as range swath in radar literature), then the unknown range profile can be discretized into  $N = BT$  delay bins. The signal model at the receiver can be written as  $y(t) = \sum_{n=1}^N x_n \phi(t - n\Delta) + n(t)$ ,

where  $\Delta = 1/B$ ,  $n(t)$  is the receiver noise and  $x_n$  denotes the complex scattering coefficients associated with the  $n$ 'th delay bin. Commonly the received signal is matched filtered with a copy of the transmitted pulse to detect reflectors in range and estimate their complex amplitude of backscattered energy. Direct digital implementation of the matched filtering step requires quadrature sampling of the received signal for the pulse duration with sampling rate matching the bandwidth of the transmit waveform. Alternatively, the matched filtering can be implemented in the analog domain where the number of samples are reduced to  $N$  to cover the delay support. However, the match filter output still requires to be sampled at the Nyquist rate corresponding to the system bandwidth, rendering digital and analog matched filter receivers for *arbitrary* waveforms impractical for bandwidths exceeding gigahertz for current analog to digital converter (ADC) technology.

If a linear frequency modulated waveform (LFM)  $\phi(t) = e^{j\beta t^2}$  is used on transmit, the matched filtering can be approximately implemented through mixing the received signal with a reference LFM waveform and low pass filtering the mixer output. At the mixer output each copy of the waveform delayed by  $\Delta$  appears as a tone whose frequency is given by  $\beta\Delta$ . This pre-processing step is termed stretch processing [1], [2] and can result in substantially reduced sampling rate for the ADC used in the mixer output if the delay support  $T$  is smaller than the pulse length. Specifically, received signal at stretch processor's output can be written as:

$$y(t) = \sum_{n=1}^N x_n e^{j(n\beta\Delta)t}, \quad (1)$$

In essence, stretch processing converts range profile estimation problem into frequency spectrum estimation problem with Nyquist rate samples in time obtained after analog processing. If the scene can be assumed to have fewer targets  $K$  than the number of delay bins  $N$ , well-known results [3], [4] from the Compressive sensing (CS) shows successful reconstruction with sub-Nyquist samples is possible with the number of measurements  $M$  scaling with  $K \log N$ , if appropriate measurement operators can be implemented. Furthermore, there are numerous tractable algorithms with provable performance that are either based on convex relaxation [5]–[7] or greedy methods [8], [9] to solve the reconstruction problem. Motivated by these advances, compressed sensing techniques have been applied to a variety of problems in Radar [10]: range profile estimation [11], single pulse systems for range-doppler estimation in [12], single pulse multiple transmit and receive

system for range-doppler and azimuth estimation and target detection in [13], [14], remote sensing in [15], direction of arrival estimation in [16]. CS based radar sensors based on pure random waveforms [17], Xampling framework [18] and Random Modulator Pre-Integrator (RMPI) framework [19] using the receiver structure from [20] have also been implemented. A common theme in most of the CS literature has been randomization as it leads to measurement matrices that have provable recovery guarantees. Implementation of the randomness into compressive radar systems has proven to be a challenging task in practice; uncorrelated random signals with high peak-to-average power ratio is mismatched to the nonlinear power amplifiers used in radar systems and the system bandwidth (and as a result range resolution) is limited as digital to analog converters (DAC) have to be employed for generation of precise random signals for transmit and receive mixing.

The LFM pulse model in (1) provides an alternative strategy as it converts the range estimation problem into an equivalent sparse frequency spectrum estimation problem. Uniform subsampling in this setting has poor performance [21]. Non-uniform random sub-sampling can be used to obtain measurements with low mutual coherence [21], [22]. However, non-uniform sampling with commercially available ADCs still requires it to be rated at the Nyquist rate to accommodate closely spaced samples. We propose to use low speed *uniform* sub-sampling using a high analog bandwidth ADC in the sparse frequency spectrum estimation setting and push randomness to transmit signal structure to obtain compressive measurements. ADCs whose analog bandwidth exceeds their maximum sampling rate by several factors is readily available and used routinely in pass-band sampling. This compressive radar structure proposed in [23] uses linear combination of LFM waveform at the transmitter with randomly selected center frequencies, while maintaining the simple standard stretch processing receiver structure. The output of the stretch processor receiver is given by

$$y(t) = \sum_{n=1}^N x_n \sum_{k=1}^{N_c} e^{j\phi_{n,k}} e^{j(n\beta\Delta + \omega_k)t}$$

where  $\phi_{n,k}$  is a predetermined known complex phase,  $N_c$  is the number of tones modulating the LFM waveform. We observe that under the proposed compressive sensor design each delayed copy of the transmitted waveform is mapped to multi-tone spectra with known structure. As shown in this paper this known multi-tone frequency structure enables recovery from aliased time samples with provable guarantees. These results complement previous work which has shown good empirical performance in simulation and measurements [24].

The contributions put forth in this sequel is that we present the theoretical analysis for this multi-frequency LFM system. We show that the system with a relatively small number of LFM waveform has performance guarantees similar to a matrix with independent random entries for a sufficiently large number of tones modulating the LFM waveform. We also present a numerical analysis comparing our system with other measurement schemes.

## Notation and Preliminaries

We denote a vector in  $N$ -dimensional complex domain as  $\mathbf{x} \in \mathbb{C}^N$ .  $\|\mathbf{x}\|_0$  is called as  $\ell_0$  norm, which is given as the number of non-zero elements in a vector. Clearly, this is not a valid norm but is used in formulating the fundamental problem in compressed sensing. We denote  $\|\mathbf{x}\|_1 = \sum_i |x_i|$  as the  $\ell_1$  norm. The Euclidean or  $\ell_2$  norm is given as  $\|\mathbf{x}\|_2 = \sqrt{\sum_i |x_i|^2}$ . We denote a matrix as  $\mathbf{A} \in \mathbb{C}^{M \times N}$ ,  $\mathbf{A}^* \in \mathbb{C}^{N \times M}$  as conjugate transpose of  $\mathbf{A}$ , and  $\mathbf{I}$  as the identity matrix of dimensions dependent on the context. The spectral or operator norm of the matrix is given as  $\|\mathbf{A}\|_{op} = \sigma_{\max}(\mathbf{A})$  the largest singular value of the matrix. The Frobenius norm of a matrix is given as  $\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} |A_{i,j}|^2}$ . Another important quantity of interest is the mutual coherence  $\mu(\mathbf{A})$ , which is a measure of the correlation between the columns of matrix  $\mathbf{A}$ . The mutual coherence is given as  $\mu(\mathbf{A}) = \max_{i \neq j} \frac{|\langle \mathbf{A}_i, \mathbf{A}_j \rangle|}{\|\mathbf{A}_i\|_2 \|\mathbf{A}_j\|_2}$ , where  $\mathbf{A}_i$  is a column of matrix  $\mathbf{A}$ . Another fundamental property for the measurement matrices is called as Restricted Isometry Property (RIP). A measurement matrix  $\mathbf{A} \in \mathbb{C}^{M \times N}$  is said to satisfy RIP of order  $K$ , if for any  $K$ -sparse vector  $\mathbf{x} \in \mathbb{C}^N$

$$(1 - \delta) \|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1 + \delta) \|\mathbf{x}\|_2^2, \\ \text{equivalently, } \delta_K = \max_{\substack{\Gamma \\ \text{card}(\Gamma) \leq K}} \|\mathbf{A}_\Gamma^* \mathbf{A}_\Gamma - \mathbf{I}\| \leq \delta, \quad (2)$$

where  $\Gamma$  is an index set that selects the columns of  $\mathbf{A}$ , and  $\text{card}(\Gamma)$  refers to the number of elements in the set,  $\delta \leq 1$  and  $\mathbf{A}_\Gamma$  is the restriction of  $\mathbf{A}$  having columns indexed by  $\Gamma$ . We denote the expectation operator as  $\mathbf{E}(\cdot)$ . The circularly-symmetric complex Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$  is denoted as  $\mathcal{CN}(\mu, \sigma^2)$ .

## A. Relation to other works

Related results in literature on recovery guarantees for compressed radar sensing can be broadly categorized in two categories: Results relating to *signal reconstruction* establish uniform recovery guarantees for successful reconstruction of all  $K$ -sparse signals, whereas results on *support recovery* is concerned with the detection of non-zero locations of a  $K$ -sparse signal that is assumed to be sampled from a generic statistical model, such as uniformly sampling from all possible subsets of size  $K$  [26].

In this paper we show that randomly sampled  $K$ -sparse signals can be recovered with high probability using LASSO for the structured measurement matrix of the proposed compressive radar sensor sensing scheme. Next, we show that the estimates of mutual coherence and column norms we obtain can be used to provide uniform recovery guarantee following a standard argument.

Table I summarizes related results for support recovery for different measurement matrices. The upper bound on the sparsity level that guarantees successful support recovery for our scheme has an additional  $\log(N + M)$  penalty compared to unstructured Gaussian matrix as shown by *Candes and Plan* in [26] and block Toeplitz matrices with entries sampled from Rademacher distribution as shown by *Bajwa* in [28]. We

Recovery Guarantees from noisy measurements with component-wise noise variance $\sigma^2$					
Matrix Type of size $M \times N$	Mutual Coherence	Sparsity for successful recovery	Spectral Norm	Minimum signal strength	Reference
Random matrix with $(NM)$ independent random entries	$2\sqrt{\frac{\log N}{M}}$	$\mathcal{O}\left(\frac{M}{\log N}\right)$	$\sqrt{\frac{N}{M}} + 1$	$\mathcal{O}(\sigma\sqrt{2\log N})$	[25]–[27]
Toeplitz block matrix with $(N + M)$ random entries	$\mathcal{O}\left(\sqrt{\frac{\log N}{M}}\right)$	$\mathcal{O}\left(\frac{M}{\log N}\right)$	$\mathcal{O}\left(\sqrt{\frac{N}{M}}\right)$	$\mathcal{O}(\sigma\sqrt{2\log N})$	[28]
LFM waveform modulated with $N_c \ll N$ randomly selected tones	$\mathcal{O}\left(\sqrt{\frac{\log N}{M}}\right)$	$\mathcal{O}\left(\frac{M}{\log N \log(N+M)}\right)$	$\mathcal{O}\left(2\sqrt{\frac{N \log(N+M)}{M}}\right)$	$\mathcal{O}(\sigma\sqrt{2\log N})$	This paper

TABLE I  
SUPPORT RECOVERY GUARANTEES FOR DIFFERENT SENSING MATRICES.

note that the random matrix with independent entries is not realizable in radar setting but included in the table to provide a baseline.

Uniform recovery guarantees are often formulated in terms of satisfying RIP property with high probability, since if a measurement matrix satisfies RIP of order  $2K$  such that  $\delta_{2K} \leq \delta \approx \sqrt{2} - 1$ , then all  $K$ -sparse vectors are successfully recovered with a reconstruction error of an oracle estimator that knows the support of the sparse vector or the support of  $K$  largest elements [29].

*Baraniuk et al.* in [30] have shown that random matrices with i.i.d entries from either Gaussian or sub-Gaussian probability distribution satisfy the RIP condition. For any  $\delta \in [0, 1]$ ,  $\delta_K \leq \delta$  if  $K \leq \alpha_4 \frac{M}{\log(N/K)}$  where  $\alpha_4$  is dependent on  $\delta$  and the sub-Gaussian norm of the random variables. Although these unstructured random matrices have remarkable recovery guarantees they do not represent any practical measurement scheme, which leads us to consider classical linear time invariant (LTI) systems.

Typically, an active imaging system transmits a signal that interacts with a scene of interest and the acquired measurements are used to estimate characteristics of the scene. The unknown environment is modeled as an LTI system whose transfer function has to be estimated using compressed measurements from the data acquisition step. It is assumed that there exists a sparse or compressible representation of the transfer function in some domain and the goal is to solve the sparse estimation problem with the least possible measurements. This leads to a structured measurement matrix that is either a partial or sub-sampled Toeplitz or circulant matrix. The RIP condition of order  $K$  for partial Toeplitz matrices in the context of channel estimation was established by *Haupt et al.* in [31]. They showed that if the sparsity  $K \leq \alpha_5 \sqrt{\frac{M}{\log N}}$ , then  $\delta_K \leq \delta$ , where  $\alpha_5$  depends on  $\delta$ . This quadratic scaling of number of measurements with respect to sparsity was improved in [32]–[34]. *Romberg* in [32] considered an active imaging system that used waveform with a random symmetric frequency spectrum and acquired compressed measurements using random sub-sampler or random demodulator at the receiver to estimate the sparse scene. The resultant system is a randomly sub-sampled circulant matrix representing the convolution and compression process. It is shown that for a given sparsity level  $K$ , the condition

that  $\delta_{2K} \leq \delta$  is satisfied if the number of measurements  $M \geq \alpha_6 \delta^{-2} \min(K(\log N)^6, (K \log N)^2)$ , where  $\alpha_6 > 0$  is a universal constant independent of the size of problem and  $\delta$ . This was extended by *Rauhut et al.* in [33]. They consider a deterministically sampled random waveform in time domain with samples following Rademacher distribution, which is modeled as a sub-sampled Toeplitz or Circulant matrix with entries sampled from Rademacher distribution. It was shown that for a given sparsity level  $K$ ,  $\delta_K \leq \delta$  with high probability if the number of measurements  $M \geq \alpha_7 \max(\delta^{-1}(K \log N)^{3/2}, \delta^{-2}K(\log N \log K)^2)$ , where  $\alpha_7$  is a universal constant. In the subsequent work by *Krahmer et al.* in [34], the relation between sparsity level and number of measurements is improved and more general random variables are considered such as vectors following sub-Gaussian distribution to generate the Toeplitz or Circulant matrix. It is shown that, for a given sparsity level  $K$  the condition  $\delta_K \leq \delta$  is satisfied if the number of measurements  $M \geq \alpha_8 \delta^{-2} K(\log K \log N)^2$ , where the constant  $\alpha_8$  is a function of only the sub-Gaussian norm of the random variables generating the matrix. We adopt a method similar to [31] and establish the RIP condition of order  $K$  and obtain a similar result stating that  $\delta_K \leq \delta$  if  $M \geq a\delta^{-2}K^2 \log N$ , where  $a > 0$  is independent of  $\delta$ .

The rest of the paper is organized as follows, Section II gives the mathematical model for the multi-frequency chirp model and the statistical model considered for the target. Section III states the main result about the measurement scheme employed for sparse recovery. Section IV contains detailed simulation results of the proposed multi-frequency chirp waveform. Section V contains the detailed proof of the main recovery result. We conclude with some future directions in Section VI.

## II. SIGNAL MODEL AND PROBLEM STATEMENT

### A. Multi-frequency chirp model

We consider a radar sensor with collocated transmitter and receiver antennas employing the compressive illumination framework proposed in [23] and [24] for estimating the range and complex reflectivity of reflectors in the scene. The chirp rate of all the transmitted linear frequency modulated (LFM) waveform is fixed at  $\frac{\beta}{\tau}$ , where  $\beta$  is the bandwidth of each transmitted waveform,  $\tau$  is the pulse duration and  $B = g\beta$  is

the system bandwidth for  $g \geq 1$ . We denote the unambiguous time interval as  $t_u = t_{\max} - t_{\min}$ , where  $t_{\max} = \frac{2R_{\max}}{c}$ ,  $t_{\min} = \frac{2R_{\min}}{c}$ ,  $R_{\max}, R_{\min}$  are the maximum and minimum range in the area of interest, respectively, while  $c$  is the velocity of light in vacuum. The whole space of range is discretized into grids based on the Radar's resolution, therefore we get  $N = Bt_u$  grids. The interval of frequency from  $[0, B]$  is divided into  $N$  grids such that  $f(i) = \frac{i\beta}{N}, i = 0, \dots, N-1$ , which are used as center frequencies for the chirp waveform. From the possible  $N$  waveform, a subset of size  $N_c$  is chosen at random for transmission. We simplify this selection model by considering independent Bernoulli random variables as indicator variables to select LFM waveform such that  $N_c$  waveform are selected on an average. Let  $\gamma_i \in \{0, 1\}$  be the random variable indicating that  $f(i)$  is part of the subset of size  $N_c$ . It can be seen that

$$\gamma_i = \begin{cases} 0 & \text{with probability (w.p.) } 1 - \frac{N_c}{N} \\ 1 & \text{w.p. } \frac{N_c}{N}. \end{cases} \quad (3)$$

The chosen LFM waveform are then scaled by independent random variables such as

- 1) a sequence of independent and identical complex phase with probability density function as  $f_{\Phi}(\phi_i) = \frac{1}{2\pi}, \phi_i \in [0, 2\pi]$ ,

$$c_i = \gamma_i \exp(j\Phi_i), \quad (4)$$

- 2) a sequence of scaling variables following Rademacher distribution given by

$$\xi_i = \begin{cases} -1 & \text{w.p. } 0.5 \\ 1 & \text{w.p. } 0.5 \end{cases} \quad (5)$$

$$c_i = \gamma_i \xi_i.$$

We choose the model in (5), which states that the  $N_c$  chirp waveform are scaled by random signs in our analysis but use the model in (4) in simulation results.

The transmitted signal can be written as

$$s(t) = \frac{1}{\sqrt{MN_c}} \sum_{i=1}^{N_c} c_i \exp\left(j2\pi\left(f_c + \frac{i\beta}{M}\right)t + \frac{\beta}{2\tau}t^2\right),$$

where  $0 \leq t \leq \tau$ . The receiver utilizes stretch processing at the same chirp rate as the transmitter and a fixed reference frequency  $f_d = f_c$  to demodulate the carrier frequency and estimate the round-trip delay. The overall duration of the de-chirping waveform is  $t_u + \tau$ . The sampling rate employed at the receiver is  $F_s = \frac{\beta}{\tau}t_u$ . The total number of samples in the pulse duration  $\tau$  is  $M = \beta t_u$ . The output samples of the stretch processor due to the target at different delay bins  $\Delta_m = \frac{m}{g\beta}$  are

$$y(k) = \frac{1}{\sqrt{MN_c}} \sum_{i=1}^{N-1} \sum_{m=0}^{N-1} c_i \exp\left(-j2\pi \frac{im}{N}\right) \times \exp\left(2\pi j \left(\frac{ip}{M} - \frac{m}{N}\right)k\right) x(m) + w(k),$$

where  $k = 0, \dots, M-1$ ,  $w_k$  is measurement noise process with 0 mean and variance  $\sigma^2$ ,  $p = \frac{\tau}{t_u} \in \mathbb{Z}$  and  $x(m)$  is the

complex scattering coefficient due to a target at the delay bin  $\Delta_m$ . This can be compactly written as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}, \quad (6)$$

where  $\mathbf{y}, \mathbf{w} \in \mathbb{C}^M$ ,  $\mathbf{A} \in \mathbb{C}^{M \times N}$ , and  $\mathbf{x} \in \mathbb{C}^N$ . The sensing matrix  $\mathbf{A}$  can be represented as a series of deterministic matrices corresponding to the response to each of the chirp waveform scaled by zero mean random coefficients as shown

$$\mathbf{A} = \sum_{i=0}^{N-1} c_i \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i. \quad (7)$$

The individual components are as follows

$$\bar{\mathbf{A}} = \frac{1}{\sqrt{MN_c}} [\bar{\mathbf{A}}(0) \quad \dots \quad \bar{\mathbf{A}}(N-1)]$$

$$\bar{\mathbf{A}}(r) = \left[ 1 \quad \exp(-2\pi j \frac{r}{N}) \quad \dots \quad \exp(-2\pi j \frac{r(M-1)}{N}) \right]^T,$$

$$\mathbf{D}_i = \text{diag} \left[ 1 \quad \exp(-j2\pi \frac{i}{N}) \quad \dots \quad \exp(-j2\pi \frac{i(N-1)}{N}) \right],$$

$$\mathbf{H}_i = \text{diag} \left[ 1 \quad \exp(j2\pi \frac{ip}{M}) \quad \dots \quad \exp(j2\pi \frac{ip(M-1)}{M}) \right], \quad (8)$$

where  $i = 0, \dots, N-1$  and  $r = 0, \dots, N-1$ ,  $\bar{\mathbf{A}} \in \mathbb{C}^{M \times N}$  are the samples from tones that correspond to each delay bin generated as a result of the de-chirping process in case of a single chirp system,  $\mathbf{H}_i \in \mathbb{C}^{M \times M}$  is the shift in frequency due to the  $i^{\text{th}}$  chirp waveform, and  $\mathbf{D}_i \in \mathbb{C}^{N \times N}$  is the phase term associated with different delay bins due to the  $i^{\text{th}}$  chirp. Each column of sensing matrix  $\mathbf{A}$  can also be represented as

$$\mathbf{A}(m) = \mathbf{E}_m \mathbf{F} \mathbf{G}_m \mathbf{c}, \text{ where}$$

$$\mathbf{E}_m = \text{diag} \left[ 1 \quad \exp(-j2\pi \frac{m}{N}) \quad \dots \quad \exp(-j2\pi \frac{m(M-1)}{N}) \right]$$

$$\mathbf{F} = \frac{1}{\sqrt{MN_c}} [\mathbf{F}(0) \quad \dots \quad \mathbf{F}(N-1)]$$

$$\mathbf{F}(r) = \left[ 1 \quad \exp(2\pi j \frac{rp}{M}) \quad \dots \quad \exp(2\pi j \frac{rp(M-1)}{M}) \right]^T,$$

$$\mathbf{G}_m = \text{diag} \left[ 1 \quad \exp(-j2\pi \frac{m}{N}) \quad \dots \quad \exp(-j2\pi \frac{m(N-1)}{N}) \right] \quad (9)$$

where  $m = 0, \dots, N-1$ ,  $r = 0, \dots, N-1$ ,  $\mathbf{E}_m \in \mathbb{C}^{M \times M}$  represents the tone generated due to target present at  $m^{\text{th}}$  delay bin,  $\mathbf{F} \in \mathbb{C}^{M \times N}$  are the different chirp center frequencies,  $\mathbf{G}_m \in \mathbb{C}^{M \times M}$  is the phase term due to different chirp frequencies for a particular delay bin  $m$ , and  $\mathbf{c} \in \mathbb{C}^N$  is the random vector that selects the chirp waveform and scales them. A closer inspection of matrix  $\mathbf{F}$  reveals that each of the center frequencies used to shift the chirp waveform is being aliased into lower frequency tones as we are sampling at Sub-Nyquist rate. We assumed  $p \in \mathbb{Z}$  in order to simplify the analysis as we get sub-sampled Discrete Fourier Transform (DFT) matrices. We impose an additional condition that  $p$  should be co-prime with  $M$  in order for  $N$  frequency tones to be uniformly mapped onto  $M$  frequency bins, where  $M \leq N$ . A simple example of  $p = M+1$ , which makes  $p$  co-prime with  $M$ , circularly maps the  $N$  possible frequencies into  $M$  bins.

### B. Target model

We consider a statistical model similar to *Strohmer and Friedlander* in [13] for the sparse range profile of targets. We assume that the targets are located at the  $N$  discrete locations corresponding to different delay bins. The support of the  $K$ -sparse range profile is chosen uniformly from all possible subsets of size  $K$ . The complex amplitude of non-zero components is assumed to have an arbitrary magnitude and uniformly distributed phase in  $[0, 2\pi]$ .

### C. Problem statement

Given a sparse scene with targets following the statistical model discussed in previous section, and measurement scheme in (6) with  $M \ll N$  and sparsity level  $K \ll N$ , the goal of compressed sensing [3] is to recover the sparse or compressible vector  $\mathbf{x}$  using minimum number of measurements in  $y$  constructed using random linear projections. The search for sparsest solution can be formulated as an optimization problem given below

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0, \text{ subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \eta, \quad (10)$$

where  $\eta$  is the noise variance. This was shown to be NP-hard and hence, intractable [35], and many approximate solutions have been found. One particular solution is to use convex relaxation technique to modify the objective as  $\ell_1$  norm minimization instead of the non-convex  $\ell_0$  norm given by,

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \eta. \quad (11)$$

This approach has been shown to successfully recover sparse or compressible vectors [6], [29] given that the sub-matrices formed by columns of sensing matrix are well conditioned. Our analysis is based on LASSO [7], which is a related method that solves the optimization problem in (11). It has been shown in [26] that for an appropriate choice of  $\lambda$  and conditions on measurement matrix, the support of the solution of the below mentioned optimization problem coincides with the support of the solution of the intractable problem in (10),

$$\min_{\mathbf{x}} \lambda \|\mathbf{x}\|_1 + \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2. \quad (12)$$

The goal of our analysis is to show that the measurement model given in (7) satisfies conditions on mutual coherence given in [26] and to find a bound on the sparsity level of range profile, which guarantees successful support recovery of almost all sparse signals using LASSO with high probability from noisy measurements. In addition, we also provide an estimate of the number of measurements required for the sensing matrix representing our scheme to satisfy the RIP condition. The next section presents our main results of our analysis.

## III. RECOVERY GUARANTEES

In order to obtain the non-asymptotic recovery guarantee for our system employing multiple chirps, we find an estimate of the tail bounds of mutual coherence and spectral or operator norm of our measurement matrix. Using the estimates, we also provide conditions for RIP condition of order  $K$  to hold.

We make use of the Matrix Bernstein inequality given in lemma 6 to get a tail bound on the operator norm for the measurement matrix given in (7).

*Lemma 1:* Given the measurement matrix model in (7), if  $N_c \geq \frac{4}{9} \log(N + M)$ , then we can bound the tail probability for the operator norm as follows

$$\begin{aligned} P \left( \|\mathbf{A}\|_{op} \geq 2(1 + \epsilon) \sqrt{\frac{N \log(N + M)}{M}} \right) \\ \leq \left( \frac{1}{N + M} \right)^{\alpha_2 - 1}, \end{aligned} \quad (13)$$

$\forall t > 0, \epsilon > 0$ , where

$$\alpha_2 = \frac{2(1 + \epsilon)^2}{1 + \frac{2(1 + \epsilon)}{3} \sqrt{\frac{\log(N + M)}{N_c}}}.$$

In addition, we also obtain an estimate of the expected value of the operator norm of measurement matrix given as

$$\mathbf{E} \left( \|\mathbf{A}\|_{op} \right) \leq \sqrt{\frac{2N}{M} \log(N + M)} + \frac{\log(N + M)}{3} \sqrt{\frac{N}{MN_c}}. \quad (14)$$

The following results about the Euclidean norm of columns and mutual coherence are obtained using concentration inequalities of quadratic forms of sub-Gaussian random vectors given in [36], which is extended to the complex domain in lemma 12.

*Lemma 2:* The concentration inequality for the minimum of Euclidean norm of any column  $m$  of  $\mathbf{A}$  is given as follows

$$\begin{aligned} P \left( \min_m \|\mathbf{A}(m)\|_2^2 \leq 1 - \epsilon \right) \leq \\ 4N \exp \left( -Md \left( \epsilon \frac{q^*}{\left(\frac{N_c}{N}\right)^{\frac{2}{q^*} - 1}} \right)^2 \right), \end{aligned} \quad (15)$$

where  $d > 0$  is a universal constant,  $\epsilon \in (0, 1)$ , and

$$q^* = \max \left( 1, 2 \log \left( \frac{N}{N_c} \right) \right).$$

*Lemma 3:* If  $M \geq (\log N)^3$  then there exists constant  $\alpha_3 > 0$  such that the mutual coherence of our sensing matrix is bounded by

$$\begin{aligned} P \left( \mu(\mathbf{A}) \geq \frac{\alpha_3 + \epsilon}{1 - \epsilon_1} \sqrt{\frac{\log N}{M}} \right) \leq \\ \begin{cases} \frac{2}{N^{u_1 - 2}} + 4N \exp(-Md\epsilon^2), & \text{if } \log N > \frac{q^*}{\left(\frac{N_c}{N}\right)^{(2/q^* - 1)}} \alpha_3 \\ \frac{2}{N^{u_2 - 2}} + 4N \exp(-Md\epsilon^2), & \text{otherwise} \end{cases} \end{aligned} \quad (16)$$

where  $d > 0$  is a universal constant,  $\epsilon > 0, \epsilon_1 \in (0, 1)$  are arbitrary constants, and

$$\begin{aligned} u_1 &= d \left( \frac{q^*(\alpha_3 + \epsilon)}{\frac{N_c}{N} \left(\frac{2}{q^*} - 1\right)} \right)^2, \\ u_2 &= \frac{q^*(\alpha_3 + \epsilon)}{\frac{N_c}{N} \left(\frac{2}{q^*} - 1\right)} d \log N, \end{aligned}$$

$$q^* = \max \left( 1, 2 \log \left( \frac{N}{N_c} \right) \right),$$

$$\bar{\epsilon} = \left( \epsilon_1 \frac{q^*}{\left( \frac{N_c}{N} \right)^{\frac{2}{q^*} - 1}} \right).$$

*Theorem 1:* For a measurement model  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$ , where  $\mathbf{A}$  is defined in (7) such that  $\mathbf{x}$  is drawn from a  $K$ -sparse model in complex domain and  $\mathbf{w} \sim \mathcal{CN}(0, \sigma^2 I)$ , the following conditions guarantee successful support recovery from solving (12) with regularizer  $\lambda = 2\sigma\sqrt{\log N}$ ,

$$K \leq K_{\max} = \frac{(1 - \epsilon_1) \alpha_1 M}{(1 + \epsilon)^2 \log(N) \log(N + M)}, \quad (17)$$

$$M \geq (\log N)^3, \log N \geq \frac{q^*}{\left( \frac{N_c}{N} \right)^{\frac{2}{q^*} - 1}} \alpha_3 \quad (18)$$

$$N_c \geq \max \left( \frac{4}{9} \log(N + M), \nu N \right), \quad (19)$$

$$\min_{k \in I} |x_k| > \frac{8}{\sqrt{1 - \epsilon_1}} \sigma \sqrt{2 \log(N)}, \quad (20)$$

with probability  $\bar{p}_4(1 - \bar{p}_1 - \bar{p}_2 - \bar{p}_3)$  for some  $\alpha_3 > 0, \alpha_1 > 0, \epsilon > 0, \epsilon_1 \in (0, 1), \nu \ll 1, d > 0$  is a universal constant independent of  $N, M$ , where

$$\bar{p}_1 = \frac{2}{N^{u_1 - 2}} + 4N \exp(-dM\bar{\epsilon}^2)$$

$$\bar{p}_2 = \left( \frac{1}{N + M} \right)^{\alpha_2 - 1} + 4N \exp(-dM\bar{\epsilon}^2)$$

$$\bar{p}_3 = 4N \exp(-dM\bar{\epsilon}^2),$$

$$\bar{p}_4 = 1 - 2N^{-1}(2\pi \log(N) + KN^{-1}) - \mathcal{O}(N^{-2 \log 2}),$$

$$\alpha_2 = \frac{2(1 + \epsilon)^2}{1 + \frac{2(1 + \epsilon)}{3} \sqrt{\frac{\log(N + M)}{N_c}}}$$

$$u_1 = d \left( \frac{q^*(\alpha_3 + \epsilon)}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)} \right)^2$$

$$q^* = \max \left( 1, 2 \log \left( \frac{N}{N_c} \right) \right),$$

$$\bar{\epsilon} = \left( \epsilon_1 \frac{q^*}{\left( \frac{N_c}{N} \right)^{\frac{2}{q^*} - 1}} \right).$$

The proof in section V involves direct application of lemma 5 and uses the estimates of the spectral norm and the mutual coherence of the measurement matrix.

*Theorem 2:* For the measurement matrix  $\mathbf{A}$  given in (7) and any  $\delta, \epsilon \in [0, 1]$  such that  $\delta + \epsilon < 1$ , the RIP condition given as  $\delta_K \leq \delta + \epsilon$  is satisfied with probability  $1 - p_5 - p_6$  if the number of measurements  $M \geq a\delta^{-2}K^2 \log N$ , where

$$p_5 = \frac{1}{N^{(u_3 - 2)}},$$

$$p_6 = 4N \exp \left( -d \left( \epsilon \frac{q^*}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)} \right)^2 M \right),$$

$$u_3 = a \left( \frac{q^*}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)} \right)^2,$$

$a > 0$  is a constant independent of  $N, M$ .

We adopt a similar approach as *Haupt et. al.* in [31] and utilize the estimates of inner-product of columns of sensing matrix and norms to obtain a simple bound on the number of measurements required to guarantee RIP of order  $K$ .

### Discussion

The support recovery guarantee stated in Theorem 1 is satisfied for almost all  $K$ -sparse vectors sampled from the generic sparse signal model discussed earlier, *i.e.* given a measurement matrix one could find a  $K$ -sparse vector (with arbitrarily small probability varying  $N, M$  and  $N_c$ .) for which the recovery fails. This differs from the worst-case guarantees as well as reconstruction error bounds that depend on Restricted Isometry Property (RIP) given in Theorem 2. The exponent in probability tail bounds for quantities such as mutual coherence  $\mu(\mathbf{A})$ , spectral norm  $\|\mathbf{A}\|_{op}$  of the measurement matrix are controlled by number of chirp waveform employed  $N_c$ . Specifically, it can be seen that the upper bound to the tail probability of the above estimates reduce as  $N_c$  increases until  $N_c \leq \frac{N}{e}$ . We also show empirically in section IV that the expected value of mutual coherence  $\mu(\mathbf{A})$  reduces as the number of chirp waveform increases. Specifically, it converges in mean to the mutual coherence of an unstructured random matrix  $\mathbf{G}$  with independent Gaussian entries and thereby converges in probability as well. Typically, smaller values of  $\mu(\mathbf{A})$  are desirable for robust recovery as shown by *Candes and Plan* in [26] as it ensures that the Grammian matrices of the submatrices formed using a subset of columns of sensing matrix  $\mathbf{A}$  are well conditioned as shown by *Tropp* in [37]. Since, the minimum value of the signal has to be above the noise floor (20) for successful recovery, we get a condition on the signal to noise ratio SNR for a particular target located at a fixed range bin  $r$  below which the recovery guarantee does not hold, which is given by

$$SNR_r = \frac{|x_r|^2}{\sigma^2} \geq 128\kappa \log N.$$

The authors in [13], [26] also show that the threshold on  $SNR_r$  scales with  $C \log N$  with constant  $C$  determining the probability of successful recovery. In section IV we study the effect of SNR on the reconstruction error using simulations.

### IV. SIMULATION EXAMPLES

For our simulation studies, we consider a system with a bandwidth  $B = 1GHz$  from which we choose center frequency of each chirp waveform randomly with each chirp sweeping a fraction of the system bandwidth  $\beta = \frac{1}{9}B$ . We note that a wideband multi-tone signal with bandwidth  $B$  modulating a LFM waveform of bandwidth  $\beta$  results in a system bandwidth of  $B + \beta$  Hz, but we seek to resolve targets with range resolution that corresponds to  $B$  Hz common to all modulated chirps. The fractional bandwidth defined as the ratio of  $\frac{\beta}{B}$  represents the under-sampling ratio as the stretch processor output is uniformly sub-sampled at that rate. We assume that the minimum and maximum range of the area of interest are  $R_{\min} = 0m$  and  $R_{\max} = 150m$ , respectively.

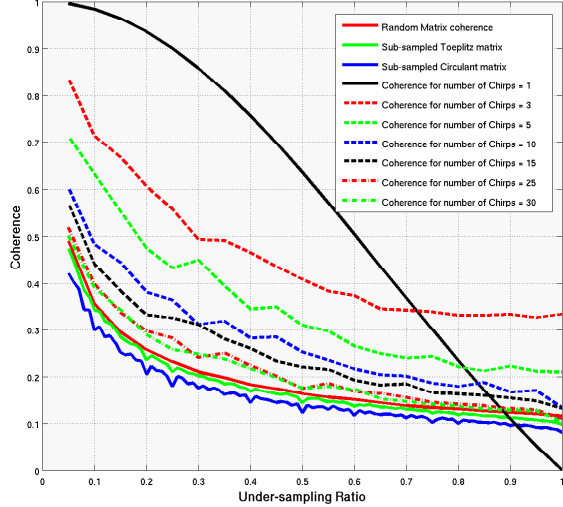


Fig. 1. Mutual Coherence of structured sensing matrices as a function of under sampling ratio. Random Gaussian matrix coherence is provided as a baseline.

The pulse duration  $\tau$  is chosen such that the ratio  $p = \frac{\tau}{t_u} \in \mathbb{Z}$  is co-prime with the number of samples. We make use of the model in (4) to select a subset of  $N_c$  chirp waveforms and scale with a random phase term to obtain the simulation results. The other measurement matrices that we compare the performance of our scheme with are

- 1) matrix  $\mathbf{G}$  with i.i.d. complex Gaussian random entries sampled from distribution  $\mathcal{CN}(0, \frac{1}{M})$ ,
- 2) matrix  $\mathbf{T} = \frac{1}{\sqrt{M}} \mathbf{P}_\Omega \mathbf{T}_1$ , which is a partial Toeplitz matrix, where  $\mathbf{P}_\Omega$  is a uniform sub-sampling operator,

$$\mathbf{T}_1 = \begin{bmatrix} t_N & t_{N-1} & \cdots & t_1 \\ t_{N+1} & t_N & \cdots & t_2 \\ & \vdots & \vdots & \\ t_{2N-1} & t_{2N-2} & \cdots & t_N \end{bmatrix}$$

$t_i \sim \mathcal{CN}(0, 1), i = 1, \dots, 2N - 1$ . and

We generate 100 realizations of matrices  $\mathbf{G}$ ,  $\mathbf{A}$  and  $\mathbf{T}$ , in order to observe the effect of under-sampling on mutual coherence. It can be seen empirically from figure 1 that the mutual coherence of the sensing matrix  $\mathbf{A}$  representing our system converges in mean to the mutual coherence of a sensing matrix  $\mathbf{G}$  as the number of chirps increase. We observe that even for small values of  $\nu \triangleq \frac{N_c}{N} = 0.03$ , the mean of the mutual coherence of our measurement matrix  $\mathbf{A}$  is quite close to the mean of the mutual coherence of matrix  $\mathbf{G}$ . In addition, we also evaluate the coherence of Toeplitz matrix  $\mathbf{T}$ , which is representative of active-imaging schemes using stochastic waveform that are modeled as linear time invariant systems.

Next, we consider the recovery performance of our measurement system used in conjunction with Basis pursuit de-noise using SPGL1 solver developed by *Van den Berg and Friedlander* in [38], [39] to estimate the unknown target locations and their amplitudes in the area of interest. For each

realization of the measurement matrix we generate multiple samples of target range profile with specified sparsity level and scattering coefficient is sampled at specified locations are sampled from a complex Gaussian distribution. The overall target range profile is normalized to get a fixed SNR. We consider a function of mean square error as a performance measure, specifically a thresholding function for a fixed SNR of 25dB. We consider a threshold of 1% on the mean squared error and vary both the number of targets in the scene, and the bandwidth of the chirp waveform, which in turn influence the sampling rate at receiver. Again it is clear from figure 2 that the performance becomes similar to that of the random Gaussian sensing matrix  $\mathbf{G}$  as the number of chirps increase. We also observe that the recovery performance of Toeplitz matrix  $\mathbf{T}$  is similar to our system at lower values of fractional bandwidth  $\frac{\beta}{B} \leq 0.7$  but performs better when  $\frac{\beta}{B} \geq 0.7$ .

Next, we see the influence of noise on sparse target recovery using Basis pursuit de-noise employing our measurement scheme. We fix fractional bandwidth  $\frac{\beta}{B} = 0.4$  and vary the noise variance as well as the number of targets in the scene. In Figure 3 the intensity of the image represents the mean square error in dB scale. Figure 3 shows that as the number of chirps increase, the performance achieved by our scheme in terms of the mean square error approaches the mean square error achieved by the random Gaussian matrix  $\mathbf{G}$ . The reconstruction error for Toeplitz matrix  $\mathbf{T}$  is marginally better at lower values of  $SNR \approx 10dB$  compared to our measurement scheme.

## V. PROOFS

*Proof of lemma 1:* The result can be obtained by direct application of lemma 6, using the value of  $L = \sqrt{\frac{N}{MN_c}}$  and  $\nu(\mathbf{A}) = \frac{N}{M}$  obtained from lemma 8 and lemma 9, respectively. The upper bound on the tail probability is given as

$$P(\|\mathbf{A}\|_{op} \geq t) \leq (N + M) \exp\left(\frac{-t^2/2}{\sqrt{\frac{N}{MN_c} \frac{t}{3} + \frac{N}{M}}}\right).$$

By plugging in  $t = 2(1+\epsilon)\sqrt{\frac{N \log(N+M)}{M}}$ , we get the result in (13). For the tail probability to decay, we require that  $\alpha_2 > 1$ . This gives us the condition that  $1/3\sqrt{\frac{\log(N+M)}{N_c}} < 0.5$ , which implies  $N_c \geq \frac{4}{9} \log(N + M)$ . Similarly, using the estimates  $L$ , and  $\nu(\mathbf{A})$  we can bound the expected value of the operator norm of  $\mathbf{A}$  as given in lemma 6. ■

*Proof of lemma 2:* The norm of a column  $m$  of the sensing matrix  $\mathbf{A}$  can be written as follows

$$\|\mathbf{A}(m)\|_2^2 = \mathbf{c}^* \mathbf{B} \mathbf{c},$$

where  $\mathbf{B} = \mathbf{G}_m^* \mathbf{F}^* \mathbf{E}_m^* \mathbf{E}_m \mathbf{F} \mathbf{G}_m$ , and  $\mathbf{c} \in \mathbb{R}^N$  is a sequence of random variables that selects and scales a subset of the chirp waveforms. It can be verified that the diagonal elements of matrix  $\mathbf{B}$  are as follows

$$B_{i,i} = \frac{1}{N_c}, i = 1, \dots, N.$$

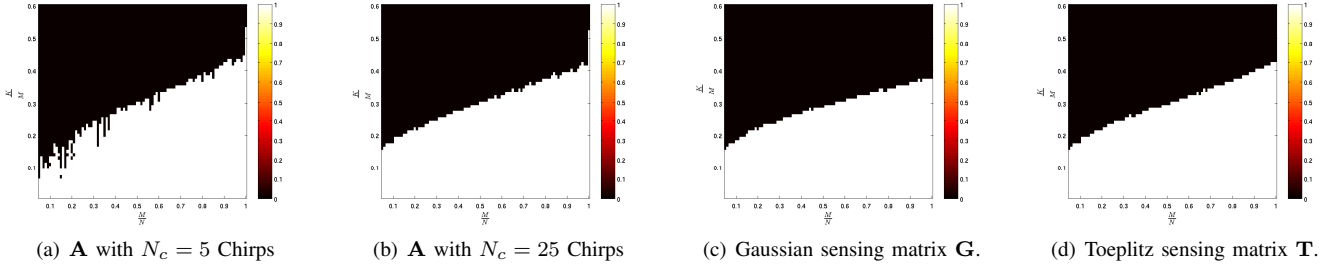


Fig. 2. Phase transition curve for the multi-frequency measurement scheme as number of chirps  $N_c$  increases and other related sensing schemes. Image intensity is either 1 if reconstruction error is below 1% and 0 otherwise. We compare the performance of our multi-frequency scheme with other random structured and unstructured matrices as the sparsity level and fractional bandwidth is varied at Signal to Noise ratio of 25 dB.

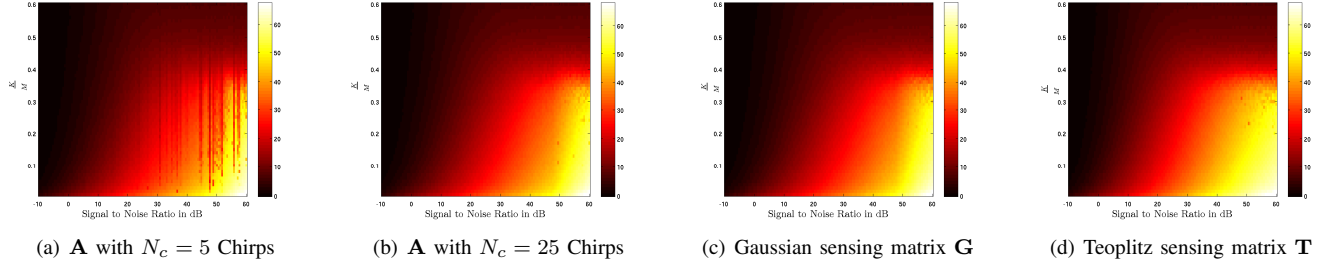


Fig. 3. The intensity values of the images represent the reconstruction error in the log scale. Each image from left to right, represents the recovery error for our measurement scheme as number of chirps  $N_c$  is increased and other measurement schemes as a function of SNR and sparsity level at a fixed fractional bandwidth  $\frac{\beta}{B} = 0.4$ .

Since the random variables  $c_i$  are independent and  $\mathbf{E}(c_i) = 0$ , the off-diagonal terms vanish and we get

$$\begin{aligned} \mathbf{E} \left( \|\mathbf{A}(m)\|_2^2 \right) &= \sum_{i=1}^N \mathbf{E} \left( |c_i|^2 \right) B_{i,i}, \\ &= \frac{1}{N} \sum_{i=1}^N 1, \\ \mathbf{E} \left( \|\mathbf{A}(m)\|_2^2 \right) &= 1. \end{aligned}$$

Using results from lemma 10 and lemma 12 along with the result on sub-Gaussian norm from lemma 11, we have

$$\begin{aligned} P \left( \left| \|\mathbf{A}(m)\|_2^2 - 1 \right| > t \right) &\leq \\ 4 \exp \left( - \frac{d}{\frac{N_c}{N} \left( \frac{2}{q^*} \right) \frac{1}{N_c q^*} \lceil \frac{N}{M} \rceil} \min \left( \frac{t^2}{\frac{N_c}{N} \left( \frac{2}{q^*} \right) \frac{M}{N_c q^*} \lceil \frac{N}{M} \rceil}, t \right) \right), \end{aligned} \quad (21)$$

$$\begin{aligned} P \left( \|\mathbf{A}(m)\|_2^2 < 1 - t \right) &\leq \\ 4 \exp \left( - \frac{d}{\frac{N_c}{N} \left( \frac{2}{q^*} \right) \frac{1}{N_c q^*} \lceil \frac{N}{M} \rceil} \min \left( \frac{t^2}{\frac{N_c}{N} \left( \frac{2}{q^*} \right) \frac{M}{N_c q^*} \lceil \frac{N}{M} \rceil}, t \right) \right) \end{aligned}$$

The concentration inequality for the minimum value of norm of any column can be written as for any  $t \in [0, 1]$ ,

$$\begin{aligned} P \left( \min_m \|\mathbf{A}(m)\|_2^2 \geq 1 - t \right) &\geq 1 - NP \left( \|\mathbf{A}(m)\|_2^2 \leq 1 - t \right), \\ P \left( \min_m \|\mathbf{A}(m)\|_2^2 \leq 1 - t \right) &\leq NP \left( \|\mathbf{A}(m)\|_2^2 \leq 1 - t \right). \end{aligned}$$

Let  $t = \epsilon$  for any  $\epsilon \leq \frac{\left( \frac{N_c}{N} \right)^{\frac{2}{q^*} - 1}}{q^*} \in (0, 1]$ . Using the approximation  $\lceil \frac{N}{M} \rceil \approx \frac{N}{M}$ , we get the required result.  $\blacksquare$

*Proof of lemma 3:* We can express the inner-product between any two columns  $m_1$  and  $m_2$  of sensing matrix as

$$\langle \mathbf{A}(m_1), \mathbf{A}(m_2) \rangle = \mathbf{c}^T \bar{\mathbf{B}} \mathbf{c},$$

where  $\bar{\mathbf{B}} = \mathbf{G}_{m_1}^* \mathbf{F}^* \mathbf{E}_{m_1}^* \mathbf{E}_{m_2} \mathbf{F} \mathbf{G}_{m_2}$ . The diagonal terms of the matrix  $\bar{\mathbf{B}}$  is given as

$$\begin{aligned} \bar{B}_{i,i} &= D_M \left( \frac{m_1 - m_2}{N} \right) \frac{1}{N_c} \exp \left( \frac{2\pi j (m_1 - m_2)(i - 1)}{N} \right), \\ D_M \left( \frac{m_1 - m_2}{N} \right) &= \frac{1}{M} \sum_{t=0}^{M-1} \exp \left( 2\pi j \left( \frac{m_1 - m_2}{N} \right) t \right) \end{aligned}$$

where  $i = 1, \dots, N$ . By using the fact that  $c_i$  are zero mean independent random variables, we obtain the following expression for  $\mathbf{E}(\mathbf{c}^T \bar{\mathbf{B}} \mathbf{c})$

$$\begin{aligned} \mathbf{E}(\mathbf{c}^T \bar{\mathbf{B}} \mathbf{c}) &= \sum_i \mathbf{E}(c_i^2) \bar{B}_{i,i} \\ &= D_M \left( \frac{m_1 - m_2}{N} \right) \sum_{i=1}^N \exp \left( \frac{2\pi j (m_1 - m_2)(i - 1)}{N} \right) \\ &= 0. \end{aligned}$$

Using the results from lemma 10 and lemma 12 along with the sub-Gaussian norm result from lemma 11, and making the approximation  $\lceil \frac{N}{M} \rceil \approx \frac{N}{M}$ , we see that there  $\exists \alpha_3 > 0$  such that

$$P \left( \left| \langle \mathbf{A}(m_1), \mathbf{A}(m_2) \rangle \right| > (\alpha_3 + \epsilon) \sqrt{\frac{\log N}{M}} \right) \leq$$



$$4 \exp \left( \frac{-q^* d(\alpha_3 + \epsilon) \log N}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)} h(N) \right),$$

where  $h(N) = \min \left( \frac{q^*(\alpha_3 + \epsilon)}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)}, \sqrt{\frac{M}{\log N}} \right)$ . Using the fact that  $M \geq (\log N)^3$  and  $\left\lceil \frac{N}{M} \right\rceil \approx \frac{N}{M}$ , we get

$$\begin{aligned} P \left( |\langle \mathbf{A}(m_1), \mathbf{A}(m_2) \rangle| > (\alpha_3 + \epsilon) \sqrt{\frac{\log N}{M}} \right) \\ \leq \begin{cases} \frac{4}{N^{u_1}} & \text{if } \log N > \frac{q^*}{\left( \frac{N_c}{N} \right)^{2/q^* - 1}} \alpha_3 \\ \frac{4}{N^{u_2}} & \text{otherwise,} \end{cases} \end{aligned} \quad (22)$$

where

$$\begin{aligned} u_1 &= d \left( \frac{q^*(\alpha_3 + \epsilon)}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)} \right)^2, \\ u_2 &= \frac{q^*(\alpha_3 + \epsilon)}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)} d \log N. \end{aligned}$$

The concentration inequality for the coherence of matrix  $\mathbf{A}$  can be obtained by using the following inequality

$$\begin{aligned} \mu(\mathbf{A}) &= \max_{m_1, m_2} \frac{|\langle \mathbf{A}(m_1), \mathbf{A}(m_2) \rangle|}{\|\mathbf{A}(m_1)\| \|\mathbf{A}(m_2)\|} \\ &\leq \max_{m_1, m_2} |\langle \mathbf{A}(m_1), \mathbf{A}(m_2) \rangle| \max_m \frac{1}{\|\mathbf{A}(m)\|^2} \end{aligned}$$

$$\begin{aligned} P \left( \mu(\mathbf{A}) \geq \frac{\alpha_3 + \epsilon}{1 - \epsilon_1} \sqrt{\frac{\log(N)}{M}} \right) \\ \leq P \left( \max_{m_1, m_2} |\langle \mathbf{A}(m_1), \mathbf{A}(m_2) \rangle| \geq (\alpha_3 + \epsilon) \sqrt{\frac{\log(N)}{M}} \right) \\ + P \left( \min_m \|\mathbf{A}(m)\|_2^2 \leq 1 - \epsilon_1 \right) \\ \leq \frac{N^2}{2} P \left( |\langle \mathbf{A}(m_1), \mathbf{A}(m_2) \rangle| \geq (\alpha_3 + \epsilon) \sqrt{\frac{\log(N)}{M}} \right) \\ + P \left( \min_m \|\mathbf{A}(m)\|_2^2 \leq 1 - \epsilon_1 \right) \end{aligned}$$

Using (15) and (22) in the above expression, we get the result in (16).  $\blacksquare$

*Proof of Theorem 1:* Using  $M \geq \log(N)^3$  and  $\log N > \frac{q^*}{\left( \frac{N_c}{N} \right)^{2/q^* - 1}} \alpha_3$  in (16), the coherence condition given in [26] is satisfied with high probability as shown below

$$\begin{aligned} \mu(\mathbf{A}) &= \mathcal{O} \left( \frac{1}{\log N} \right) \\ \text{w.p. } p_1 &\geq 1 - \frac{2}{N^{u_1 - 2}} + 4N \exp(-dM\bar{\epsilon}^2), \end{aligned} \quad (23)$$

where  $\alpha_3 > 0$  is a constant independent of  $N$  and  $M$ ,  $\epsilon > 0$ ,  $\epsilon_1 \in (0, 1)$ ,

$$u_1 = d \left( \frac{q^*(\alpha_3 + \epsilon)}{\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)} \right)^2,$$

and  $q^* = \max \left( 1, 2 \log \left( \frac{N}{N_c} \right) \right)$ . This establishes the condition in (18). We also note that the exponent in the probability tail bound in (18) depends on  $\frac{N_c}{N}$  as the function  $\frac{N_c}{N} \left( \frac{2}{q^*} - 1 \right)$  is an increasing in  $N_c \in [0, 1/\exp(1)]$  but decreases in  $N_c \in [1/\exp(1), 1]$ . In addition to this, we also believe that  $\alpha_3$  decreases as  $N_c$  increases linearly with  $N$ , which is also verified numerically in section IV. This leads to part of the condition in (19) given as  $\frac{N_c}{N} \geq \nu \ll 1$ .

The measurement matrix in our analysis does not have unit norm columns and in order to apply lemma 5, we normalize the columns. We follow the approach similar to [13]. Let  $\mathbf{D} \in \mathbb{R}^{N \times N}$  diagonal matrix with diagonal entries corresponding to the norm of the column of  $\mathbf{A}$  given by

$$D_{i,i} = \|\mathbf{A}(i)\|_2.$$

The measurement model can be modified as

$$\mathbf{y} = \hat{\mathbf{A}}\mathbf{z} + \mathbf{w},$$

where  $\hat{\mathbf{A}} = \mathbf{A}\mathbf{D}^{-1}$  and  $\mathbf{z} = \mathbf{D}\mathbf{x}$ . Next, we obtain the probability tail bound for the operator norm of the measurement matrix  $\hat{\mathbf{A}}$  with  $\ell_2$  normalized columns. Using lemma 7, we have  $\forall \epsilon > 0, \epsilon_1 \in (0, 1/2]$ , independent of  $N$  and  $M$ ,

$$\begin{aligned} P \left( \|\hat{\mathbf{A}}\|_{op} \geq 2 \frac{(1 + \epsilon)}{\sqrt{1 - \epsilon_1}} \sqrt{\frac{N \log(N + M)}{M}} \right) \\ \leq P \left( \|\mathbf{A}\|_{op} \geq 2(1 + \epsilon) \sqrt{\frac{N \log(N + M)}{M}} \right) \\ + P \left( \min_m D_{m,m} \leq \sqrt{1 - \epsilon_1} \right) \\ \leq \left( \frac{1}{N + M} \right)^{\alpha_1 - 1} + 4N \exp(-dM\bar{\epsilon}^2), \end{aligned}$$

where

$$\begin{aligned} \alpha_1 &= \frac{2(1 + \epsilon)^2}{1 + \frac{2(1 + \epsilon)}{3} \sqrt{\frac{\log(N + M)}{N_c}}}, \\ N_c &\geq \frac{4}{9} \log(N + M), \\ \bar{\epsilon} &= \left( \epsilon_1 \frac{q^*}{\left( \frac{N_c}{N} \right)^{2/q^* - 1}} \right). \end{aligned}$$

Therefore,

$$\begin{aligned} \|\hat{\mathbf{A}}\|_{op} &\leq 2 \frac{(1 + \epsilon)}{\sqrt{1 - \epsilon_1}} \sqrt{\frac{N \log(N + M)}{M}} \\ \text{w.p. } p_2 &\geq 1 - \left( \frac{1}{N + M} \right)^{\alpha_1 - 1} + 4N \exp(-dM\bar{\epsilon}^2). \end{aligned} \quad (24)$$

This gives us the condition in (19). Using (24) in (32) we obtain the condition in (17).

Next, we establish that the measurement matrix does not reduce the absolute value of non-zero entries of the sparse vector  $\mathbf{x}$  below the noise level.

$$\begin{aligned} P \left( \min_i D_{i,i} |x_i| \leq 8\sigma \sqrt{2 \log N} \right) &\leq NP \left( D_{i,i} \leq \sqrt{1 - \epsilon_1} \right) \\ &\leq 4N \exp(-dM\bar{\epsilon}^2). \end{aligned}$$

Therefore, we have

$$\min_i |z_i| \geq 8\sigma\sqrt{2\log N} \quad (25)$$

w.p.  $p_3 \geq 1 - 4N \exp(-dM\bar{\epsilon}^2)$ .

We define the following events associated with a realization of measurement matrix  $\mathbf{A}$

$$\begin{aligned} \Xi_1 &: \mu(\mathbf{A}) = \mathcal{O}\left(\frac{1}{\log N}\right) \\ \Xi_2 &: \|\hat{\mathbf{A}}\|_{op}^2 \leq \frac{c_0 N}{K_{\max} \log N} \\ \Xi_3 &: \min_i |z_i| \geq 8\sigma\sqrt{2\log N}. \\ \Xi_4 &: \text{successful support recovery for a fixed sensing matrix.} \end{aligned}$$

Let  $\Xi$  be the event that the sampled measurement matrix satisfies the conditions required for successful recovery and recovers a  $K$ -sparse vector  $\mathbf{x}$  selected from the target model. This implies

$$P(\Xi) \geq P(\Xi_4 | \Xi_1 \cap \Xi_2 \cap \Xi_3) \times (1 - P(\Xi_1^c) - P(\Xi_2^c) - P(\Xi_3^c)). \quad (26)$$

Using result from Lemma 5 for  $P(\Xi_4 | \Xi_1 \cap \Xi_2 \cap \Xi_3)$ , (24), (25) and (23) in (26), we get the desired recovery guarantee. ■

*Proof of Theorem 2:* Using  $\delta_K \leq \delta + \epsilon$  in (2), it can be deduced that the eigen values of  $\mathbf{A}_\Gamma^* \mathbf{A}_\Gamma, \forall \Gamma \in [1 - \delta - \epsilon, 1 + \delta + \epsilon]$  such that  $\text{card}(\Gamma) \leq K$ . This can be translated into a condition on the elements of the Gramian matrix  $\mathbf{A}^* \mathbf{A}$  using lemma 4 given as the following events  $\Xi_5 : \bigcap_{\substack{m_1, m_2 \\ m_1 \neq m_2}} \{|\langle \mathbf{A}(\mathbf{m}_1), \mathbf{A}(\mathbf{m}_2) \rangle| \leq \frac{\delta}{K}\}$   $\Xi_6 : \bigcap_{m=1}^M \left\{ \left| \|\mathbf{A}(m)\|_2^2 - 1 \right| \leq \epsilon \right\}$ .

Let  $\Xi$  be the event denoting that the RIP condition of order  $K$  is satisfied. Then we have

$$P(\Xi^c) \leq 1 - \frac{N^2}{2} P\left(|\langle \mathbf{A}(\mathbf{m}_1), \mathbf{A}(\mathbf{m}_2) \rangle| \geq \frac{\delta}{K}\right) - NP\left(\left|\|\mathbf{A}(m)\|_2^2 - 1\right| \geq \epsilon\right) \quad (27)$$

Using the result from (21) and using the fact that we can choose  $N_c$  such that  $\epsilon \leq \frac{q^*}{\frac{N_c}{N}(\frac{2}{q^*}-1)}$ , we get

$$P\left(\left|\|\mathbf{A}(m)\|_2^2 - 1\right| \geq \epsilon\right) \leq 4 \exp\left(-d \left(\epsilon \frac{q^*}{\frac{N_c}{N}(\frac{2}{q^*}-1)}\right)^2 M\right) \quad (28)$$

Similarly, using results from lemma 3 we can obtain

$$P\left(|\langle \mathbf{A}(\mathbf{m}_1), \mathbf{A}(\mathbf{m}_2) \rangle| \geq \frac{\delta}{K}\right) \leq 4 \exp\left(-dM \frac{\delta^2}{K^2} \left(\frac{q^*}{\frac{N_c}{N}(\frac{2}{q^*}-1)}\right)^2\right). \quad (29)$$

Using (28), (29) and the condition that  $M \geq a\delta^{-2}K^2 \log N$  in (27) we get the desired results. ■

## VI. CONCLUSION

In this work we have shown that structured compressed sensing matrices with random components can be realized in radar (delay estimation) setting using an LFM waveform modulated by a random sparse multi-tone signal in transmit and a simple traditional analog receiver structure. We provide recovery guarantees for the proposed compressive sensor comparable to random Toeplitz/Circular matrices with much larger number of random elements. The proposed scheme is well matched to practical implementation utilizing small number of random parameters and uniform sampling ADCs on receive.

A potential direction for future research is to investigate the effectiveness of multi-chirp waveforms in the multiple output multiple input (MIMO) setting with multiple transmit and receive antenna elements for estimating support of targets in angle and range domain [13]. We note that in this setting each transmitter can use a single chirp with a random frequency offset and superposition is achieved at each receiver as the waveforms reflected from the scene is naturally summed at each receiver.

## APPENDIX

### SOME USEFUL LEMMAS

*Lemma 4:* Given a complex matrix  $M \in \mathbb{C}^{n \times n}$  with  $C_i = M_{i,i}$ , and  $R_i = \sum_{j \neq i} |M_{i,j}|$  then we have

$$\lambda_i \in \cup_{i=1}^n D(C_i, R_i), \forall i = 1, \dots, n, \quad (30)$$

where  $D(c, r)$  is a disc with center  $c$  and radius  $r$ , and  $\lambda_i$  are the eigen-values of  $M$ .

We restate the theorem given in [13], which was extended to the complex setting in [26], that give the conditions on the measurement matrix for successful recovery when using  $\ell_1$  penalized optimization methods.

*Lemma 5:* For a measurement model  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$ , where  $\mathbf{A} \in \mathbb{C}^{M \times N}$  has unit  $\ell_2$  norm columns,  $\mathbf{x}$  is drawn from a  $K$ -sparse model in complex domain and  $w_i \sim \mathcal{CN}(0, \sigma^2)$ , if the following conditions are satisfied

$$\mu(\mathbf{A}) \leq \frac{\alpha_0}{\log N}, \quad (31)$$

where  $\alpha_0 > 0$  is a constant independent of the dimensions of the problem; also, if

$$K \leq K_{\max}(\mathbf{A}) = \frac{\alpha_1 N}{\|\mathbf{A}\|_{op}^2 \log(N)}, \quad (32)$$

for some  $\alpha_1 > 0$ , and

$$\min_{k \in I} |x_k| > 8\sigma\sqrt{2\log(N)}, \quad (33)$$

then the solution  $\hat{\mathbf{x}}$  of (12) has the same support as the unknown sparse vector  $\mathbf{x}$  and relative error is bounded as shown below

$$\text{supp}(\hat{\mathbf{x}}) = \text{supp}(\mathbf{x}) \quad (34)$$

$$Pr\left(\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \frac{\sigma\sqrt{3N}}{\|\mathbf{y}\|_2}\right) \geq 1 - 2N^{-1}(2\pi \log(N) + KN^{-1}) - \mathcal{O}(N^{-2\log 2}). \quad (35)$$

*Lemma 6 (Matrix Bernstein inequality [40]):* Let  $\mathbf{A}_i$  be a sequence of i.i.d. random matrices. For a random matrix expressed as  $\mathbf{A} = \sum_i \mathbf{A}_i$  we have

$$P\left(\|\mathbf{A}\|_{op} \geq t\right) \leq (d_1 + d_2) \exp\left(\frac{-t^2/2}{\frac{Lt}{3} + \nu(\mathbf{A})}\right), \quad (36)$$

$$\begin{aligned} \|\mathbf{A}_i\|_{op} &\leq L, \forall i = 1, \dots, D \\ \nu(\mathbf{A}) &= \max(\mathbf{E}(\mathbf{A}\mathbf{A}^*), \mathbf{E}(\mathbf{A}^*\mathbf{A})), \end{aligned}$$

where  $\mathbf{A}_i \in \mathbb{C}^{d_1 \times d_2}$ . The expected value of the operator norm of  $\mathbf{A}$  is bounded by

$$\mathbf{E}\left(\|\mathbf{A}\|_{op}\right) \leq \sqrt{2\nu(\mathbf{A}) \log(d_1 + d_2)} + \frac{L \log(d_1 + d_2)}{3}. \quad (37)$$

*Lemma 7:* Given a matrix  $\mathbf{A}$ , the operator norm of the matrix  $\hat{\mathbf{A}} = \mathbf{A}\mathbf{D}^{-1}$  is bounded by the following inequality

$$\frac{\|\mathbf{A}\|_{op}}{\min_i D_{i,i}} \geq \|\hat{\mathbf{A}}\|_{op} \geq \frac{\|\mathbf{A}\|_{op}}{\max_i D_{i,i}}, \quad (38)$$

where  $\mathbf{D}$  is a diagonal matrix with positive diagonal elements.

*Proof:* For any vector  $\mathbf{v}$  such that  $\|\mathbf{v}\|_2 = 1$ , we have

$$\begin{aligned} \bar{\mathbf{A}}\mathbf{v} &= \frac{\mathbf{A}\mathbf{D}^{-1}\mathbf{v}}{\|\mathbf{D}^{-1}\mathbf{v}\|_2} \|\mathbf{D}^{-1}\mathbf{v}\|_2 = \mathbf{A}\mathbf{u}(\mathbf{v}) \|\mathbf{D}^{-1}\mathbf{v}\|_2 \\ \implies \|\bar{\mathbf{A}}\mathbf{v}\|_2 &= \|\mathbf{A}\mathbf{u}(\mathbf{v})\|_2 \|\mathbf{D}^{-1}\mathbf{v}\|_2, \end{aligned}$$

where  $\mathbf{u}(\mathbf{v}) = \frac{\mathbf{D}^{-1}\mathbf{v}}{\|\mathbf{D}^{-1}\mathbf{v}\|_2}$ . Since  $\|\mathbf{v}\|_2 = 1$ , we bound the Euclidean norm of  $\mathbf{D}^{-1}\mathbf{v}$  as follows

$$\frac{1}{\min_i D_{i,i}} \geq \|\mathbf{D}^{-1}\mathbf{v}\|_2 \geq \frac{1}{\max_i D_{i,i}}.$$

Using this we get,

$$\frac{\|\mathbf{A}\mathbf{u}(\mathbf{v})\|_2}{\min_i D_{i,i}} \geq \|\bar{\mathbf{A}}\mathbf{v}\|_2 \geq \frac{\|\mathbf{A}\mathbf{u}(\mathbf{v})\|_2}{\max_i D_{i,i}}.$$

By taking supremum over  $\mathbf{v}$  in the space of unit norm vectors we obtain the desired result.  $\blacksquare$

*Lemma 8:* For the matrix  $c_i \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i \in \mathbb{C}^{M \times N}$  in (8), we have

$$\|c_i \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i\|_{op} \leq \sqrt{\frac{N}{MN_c}}, \quad (39)$$

$\forall i = 1, \dots, N$

*Proof:* By sub-multiplicativity property of the operator norm we have,

$$\|c_i \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i\|_{op} \leq |c_i| \|\mathbf{H}_i\|_{op} \|\bar{\mathbf{A}}\|_{op} \|\mathbf{D}_i\|_{op}.$$

Since  $\mathbf{H}_i$  and  $\mathbf{D}_i$  are diagonal matrices with complex exponential entries, it can be shown that  $\|\mathbf{H}_i\|_{op} = \|\mathbf{D}_i\|_{op} = 1$ . Also, we assume that  $c_i$  are sub-Gaussian random variables with  $|c_i| \leq 1$ . In order to find the operator norm of  $\bar{\mathbf{A}}$ , we define  $\mathbf{G} = \bar{\mathbf{A}}\bar{\mathbf{A}}^* \in \mathbb{C}^{M \times M}$  since it is full rank and  $\|\bar{\mathbf{A}}\|_{op}^2 = \|\mathbf{G}\|_{op}$ . The entries of matrix  $\mathbf{G}$  are as follows

$$\begin{aligned} G(k, k) &= \frac{N}{MN_c} \\ G(k, l) &= \frac{N}{MN_c} \frac{1}{N} \sum_{m=0}^{N-1} \exp\left(2\pi j \frac{l-k}{N} m\right) \end{aligned}$$

$$= \frac{N}{MN_c} D_N \left(\frac{l-k}{N}\right) = 0,$$

$\forall k, l = 0, \dots, M-1$ , such that  $k \neq l$ , and  $D_N \left(\frac{l-k}{N}\right) = \frac{1}{N} \sum_{m=0}^{N-1} \exp(2\pi j \frac{l-k}{N} m)$  is the discrete Dirichlet Kernel. The second term is zero because the discrete Dirichlet kernel is being evaluated at it's zeros, which are the Fourier frequency bins  $(\frac{n}{N}), n \in \mathbb{Z}$ . Therefore,

$$\mathbf{G} = \frac{N}{MN_c} \mathbf{I}.$$

This implies that the  $\|\mathbf{G}\|_{op} = \frac{N}{MN_c}$  and leads to the result in (39).  $\blacksquare$

*Lemma 9:* For the matrices  $\mathbf{P}_i = c_i \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i \in \mathbb{C}^{M \times N}$  given by (8), we have

$$\left\| \sum_{i=1}^N \mathbf{E}(\mathbf{P}_i^* \mathbf{P}_i) \right\| \leq \frac{N}{M}, \quad (40)$$

$$\left\| \sum_{i=1}^N \mathbf{E}(\mathbf{P}_i \mathbf{P}_i^*) \right\| = \frac{N}{M}. \quad (41)$$

*Proof:* First, we compute the norm of  $\mathbf{P}_i \mathbf{P}_i^*$  as it is a full-rank matrix using  $\mathbf{D}_i \mathbf{D}_i^* = \mathbf{I}$ ,  $\mathbf{H}_i \mathbf{H}_i^* = \mathbf{I}$  and  $\bar{\mathbf{A}} \bar{\mathbf{A}}^* = \frac{N}{MN_c} \mathbf{I}$  we have

$$\begin{aligned} \mathbf{P}_i \mathbf{P}_i^* &= c_i c_i^* \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i \mathbf{D}_i^* \bar{\mathbf{A}}^* \mathbf{H}_i^* \\ &= c_i c_i^* \frac{N}{MN_c} \mathbf{I}. \end{aligned}$$

Using the probabilistic model for  $c_i$  given in (5), we get

$$\mathbf{E}(c_i c_i^*) = \frac{N_c}{N}.$$

We have

$$\sum_{i=1}^N \mathbf{E}(\mathbf{P}_i \mathbf{P}_i^*) = \frac{N}{M} \mathbf{I}.$$

Applying the operator norm yields the result in (41). Similarly, using the sub-additivity of the operator norm and  $\|\mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i \mathbf{D}_i^* \bar{\mathbf{A}}^* \mathbf{H}_i^*\|_{op} = \|\mathbf{D}_i^* \bar{\mathbf{A}}^* \mathbf{H}_i^* \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i\|_{op}$  we get

$$\begin{aligned} \left\| \sum_{i=1}^N \mathbf{E}(\mathbf{P}_i^* \mathbf{P}_i) \right\|_{op} &= \left\| \sum_{i=1}^N \mathbf{E}(c_i^* c_i) \mathbf{D}_i^* \bar{\mathbf{A}}^* \mathbf{H}_i^* \mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i \right\|_{op} \\ &\leq \sum_{i=1}^N |\mathbf{E}(c_i c_i^*)| \|\mathbf{H}_i \bar{\mathbf{A}} \mathbf{D}_i \mathbf{D}_i^* \bar{\mathbf{A}}^* \mathbf{H}_i^*\|_{op} \\ &= \frac{N}{M}. \end{aligned}$$

*Lemma 10:* Let  $\mathbf{B} = \mathbf{G}_m^* \mathbf{F}^* \mathbf{E}_m^* \mathbf{E}_m \mathbf{F} \mathbf{G}_m$ , and  $\bar{\mathbf{B}} = \mathbf{G}_{m_1}^* \mathbf{F}^* \mathbf{E}_{m_1}^* \mathbf{E}_{m_2} \mathbf{F} \mathbf{G}_{m_2}$  then we have

$$\|\mathbf{B}\|_{op} \leq \left\lceil \frac{N}{M} \right\rceil \frac{1}{N_c}, \quad \|\mathbf{B}\|_F \leq \left\lceil \frac{N}{M} \right\rceil \frac{\sqrt{M}}{N_c}, \quad (42)$$

$$\|\bar{\mathbf{B}}\|_{op} \leq \left\lceil \frac{N}{M} \right\rceil \frac{1}{N_c}, \quad \|\bar{\mathbf{B}}\|_F \leq \left\lceil \frac{N}{M} \right\rceil \frac{\sqrt{M}}{N_c}, \quad (43)$$

where  $\lceil x \rceil = z \in \mathbb{Z}$  such that  $z \geq x$ ,  $\forall x \in \mathbb{R}$ .

*Proof:* We note that  $\text{Rank}(\mathbf{B}) = \text{Rank}(\bar{\mathbf{B}}) = M$ , and we can obtain a bound on the Frobenius norm of the matrices as shown below

$$\begin{aligned}\|\mathbf{B}\|_F &\leq \sqrt{M} \|\mathbf{B}\|_{op}, \\ \|\bar{\mathbf{B}}\|_F &\leq \sqrt{M} \|\bar{\mathbf{B}}\|_{op}.\end{aligned}$$

In order to find the bound on the operator norm we see that

$$\begin{aligned}\|\mathbf{B}\|_{op} &\leq \|\mathbf{G}_m\|_{op} \|\mathbf{E}_m\|_{op} \|\mathbf{F}\|^2, \\ \|\bar{\mathbf{B}}\|_{op} &\leq \|\mathbf{G}_{m_1}\|_{op} \|\mathbf{G}_{m_2}\|_{op} \|\mathbf{E}_{m_1}\|_{op} \|\mathbf{E}_{m_2}\|_{op} \|\mathbf{F}\|^2.\end{aligned}$$

Since  $\mathbf{G}_m$  and  $\mathbf{E}_m$  are diagonal matrices with complex exponentials along the principal diagonal, it can be shown that  $\|\mathbf{G}_m\|_{op}^2 = \|\mathbf{E}_m\|_{op}^2 = 1$ . In order to estimate the  $\|\mathbf{F}\|_{op}$ , we see that

$$\mathbf{F}^* \mathbf{F} = \frac{1}{N_c} \begin{bmatrix} \mathbf{V}_{1,1} & \cdots & \mathbf{V}_{1, \lceil \frac{N}{M} \rceil} \\ \mathbf{V}_{2,1} & \cdots & \mathbf{V}_{2, \lceil \frac{N}{M} \rceil} \\ \vdots & \vdots & \vdots \\ \mathbf{V}_{\lceil \frac{N}{M} \rceil, 1} & \cdots & \mathbf{V}_{\lceil \frac{N}{M} \rceil, \lceil \frac{N}{M} \rceil} \end{bmatrix}$$

Since  $p$  is co-prime with  $M$ , we observe that the  $N$  possible frequency tones circularly get mapped onto  $M$  possible aliased sinusoids. Therefore, if neither  $i, j \neq \lceil \frac{N}{M} \rceil$ ,

$$\mathbf{V}_{i,j} = \mathbf{I}.$$

If either  $i = \lceil \frac{N}{M} \rceil$  or  $j = \lceil \frac{N}{M} \rceil$ , then  $\mathbf{V}_{i,j} \in \mathbb{C}^{M \times M}$  is a partial identity matrix. It can easily be verified that  $\|\mathbf{F}\|_{op}^2 = \frac{1}{N_c} \lceil \frac{N}{M} \rceil$ . Using this result along-with the estimate on the bound for the Frobenius norm, we get the desired results. ■

*Lemma 11:* For the sub-Gaussian random variables described in (5) and (4), the Sub-Gaussian norm [41] is as follows

$$\begin{aligned}\|C_i\|_{\Psi_2} &= \sup_{q \geq 1} \mathbf{E}(|C_i|^q)^{\frac{1}{q}} \frac{1}{\sqrt{q}} \\ \|C_i\|_{\Psi_2} &= \left(\frac{N_c}{N}\right)^{\frac{1}{q^*}} \frac{1}{\sqrt{q^*}}, \\ q^* &= \max\left(1, 2 \log\left(\frac{N}{N_c}\right)\right).\end{aligned}\quad (44)$$

*Proof:* For the probability models given in (5) and (4), we have

$$\begin{aligned}|C_i| &= \begin{cases} 0 & \text{with probability } 1 - \frac{N_c}{N} \\ 1 & \text{w.p. } \frac{N_c}{N}. \end{cases} \\ \mathbf{E}(|C_i|^q)^{\frac{1}{q}} &= \left(\frac{N_c}{N}\right)^{\frac{1}{q}}. \\ \implies \|C_i\|_{\Psi_2} &= \sup_{q \geq 1} \left(\frac{N_c}{N}\right)^{\frac{1}{q}} \frac{1}{\sqrt{q}}.\end{aligned}$$

The solution to this optimization problem can be found by taking the logarithm and solving the unconstrained optimization problem which is given as

$$q^* = 2 \log\left(\frac{N}{N_c}\right).$$

In order to satisfy the constraint, the solution is lower bounded by 1. ■

*Lemma 12:* Given a zero mean real random vector  $\mathbf{c}$  composed of independent and sub-Gaussian random variables  $c_i, i = 1, \dots, N$  such that  $\|c_i\|_{\Psi_2} \leq K, i = 1, \dots, N$ , we have

$$\begin{aligned}Pr(|\mathbf{c}^T \mathbf{B} \mathbf{c} - \mathbf{E}(\mathbf{c}^T \mathbf{B} \mathbf{c})| > t) \\ \leq 4 \exp\left(-d \min\left(\frac{t^2}{K^4 \|\mathbf{B}\|_F^2}, \frac{t}{K^2 \|\mathbf{B}\|_{op}}\right)\right),\end{aligned}\quad (45)$$

where  $\mathbf{B} \in \mathbb{C}^{N \times N}$ ,  $\mathbf{c} \in \mathbb{R}^N$ , for some absolute constant  $d > 0$  and  $\forall t > 0$ .

*Proof:* Let  $\mathbf{B} = \mathbf{B}_R + i\mathbf{B}_{Im}$ , where  $\mathbf{B}_R, \mathbf{B}_{Im} \in \mathbb{R}^{N \times N}$ . Therefore, using the Hanson-Wright inequality for real matrices given in [36] and the fact that  $\|\mathbf{B}_R\|_{op} \leq \|\mathbf{B}\|_{op}$ ,  $\|\mathbf{B}_{Im}\|_{op} \leq \|\mathbf{B}\|_{op}$ , and  $\|\mathbf{B}_R\|_F \leq \|\mathbf{B}\|_F$ ,  $\|\mathbf{B}_{Im}\|_F \leq \|\mathbf{B}\|_F$  we have

$$\begin{aligned}Pr(|\mathbf{c}^T \mathbf{B} \mathbf{c} - \mathbf{E}(\mathbf{c}^T \mathbf{B} \mathbf{c})| > t) &\leq \\ Pr\left(|\mathbf{c}^T \mathbf{B}_R \mathbf{c} - \mathbf{E}(\mathbf{c}^T \mathbf{B}_R \mathbf{c})| > \frac{t}{\sqrt{2}}\right) &+ \\ Pr\left(|\mathbf{c}^T \mathbf{B}_{Im} \mathbf{c} - \mathbf{E}(\mathbf{c}^T \mathbf{B}_{Im} \mathbf{c})| > \frac{t}{\sqrt{2}}\right).\end{aligned}$$

This gives us the inequality in (45). ■

#### ACKNOWLEDGMENT

This research was partially supported by Army Research Office grant W911NF-11-1-0391 and NSF Grant IIS-1231577.

#### REFERENCES

- [1] M. A. Richards, *Fundamentals of Radar Signal Processing*. New York, NY, USA: McGraw-Hill, 2014.
- [2] R. Middleton, "Dechirp-on-receive linearly frequency modulated radar as a matched-filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2716–2718, July 2012.
- [3] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [4] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [5] E. Candes and T. Tao, "The dantzig selector: Statistical estimation when  $p$  is much larger than  $n$ ," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, Dec. 2007.
- [6] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.
- [7] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 73, no. 3, pp. 273–282, 1996.
- [8] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [9] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [10] L. Potter, E. Ertin, J. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1006–1020, June 2010.
- [11] K. Gedalyahu and Y. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3017–3031, June 2010.
- [12] M. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2275–2284, June 2009.

- [13] T. Strohmer and B. Friedlander, "Analysis of sparse mimo radar," *Applied and Computational Harmonic Analysis*, vol. 37, no. 3, pp. 361–388, 2014.
- [14] T. Strohmer and H. Wang, "Accurate imaging of moving targets via random sensor arrays and kerdock codes," *Inverse Problems*, vol. 29, no. 8, p. 085001, 2013.
- [15] M. Hgel, H. Rauhut, and T. Strohmer, "Remote sensing via  $\ell_1$ -minimization," *Foundations of Computational Mathematics*, vol. 14, no. 1, pp. 115–150, 2014.
- [16] M. Rossi, A. Haimovich, and Y. Eldar, "Spatial compressive sensing in mimo radar with random arrays," in *Information Sciences and Systems (CISS), 2012 46th Annual Conference on*, Mar. 2012, pp. 1–6.
- [17] M. Shastry, R. Narayanan, and M. Rangaswamy, "Sparsity-based signal processing for noise radar imaging," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 314–325, Jan. 2015.
- [18] E. Baransky, G. Itzhak, N. Wagner, I. Shmuel, E. Shoshan, and Y. Eldar, "Sub-nyquist radar prototype: Hardware and algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 809–822, Apr. 2014.
- [19] J. Yoo, C. Turnes, E. Nakamura, C. Le, S. Becker, E. Sovero, M. Wakin, M. Grant, J. Romberg, A. Emami-Neyestanak, and E. Candes, "A compressed sensing parameter extraction platform for radar pulse signal acquisition," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 626–638, Sep. 2012.
- [20] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010.
- [21] M. Duarte and R. Baraniuk, "Spectral compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 35, no. 1, pp. 111–129, 2013.
- [22] A. Gilbert, M. Strauss, and J. Tropp, "A tutorial on fast fourier sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 57–66, Mar. 2008.
- [23] E. Ertin, "Frequency diverse waveforms for compressive radar sensing," in *Waveform Diversity and Design Conference (WDD), 2010 International*, Aug. 2010, pp. 000216–000219.
- [24] E. Ertin, L. Potter, and R. Moses, "Sparse target recovery performance of multi-frequency chirp waveforms," in *19th European Signal Processing Conference, 2011*, Aug 2011, pp. 446–450.
- [25] T. T. Cai and T. Jiang, "Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices," *The Annals of Statistics*, vol. 39, no. 3, pp. 1496–1525, June 2011.
- [26] E. Candes and Y. Plan, "Near-ideal model selection by  $\ell_1$  minimization," *The Annals of Statistics*, vol. 37, no. 5A, pp. 2145–2177, Oct. 2009.
- [27] K. Davidson and S. Szarek, "Local operator theory, random matrices and banach spaces," *Handbook on the Geometry of Banach spaces*, vol. 1, pp. 317–366, 2001.
- [28] W. Bajwa, "Geometry of random toeplitz-block sensing matrices: bounds and implications for sparse signal processing," *Proc. SPIE*, vol. 8365, pp. 836505–836505–7, 2012.
- [29] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9–10, pp. 589–592, May 2008.
- [30] R. Baraniuk, M. Davenport, R. A. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, Dec. 2008.
- [31] J. Haupt, W. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5862–5875, Nov. 2010.
- [32] J. Romberg, "Compressive sensing by random convolution," *SIAM J. Img. Sci.*, vol. 2, no. 4, pp. 1098–1128, Nov. 2009.
- [33] H. Rauhut, J. Romberg, and J. A. Tropp, "Restricted isometries for partial random circulant matrices," *Applied and Computational Harmonic Analysis*, vol. 32, no. 2, pp. 242–254, 2012.
- [34] F. Krahermer, S. Mendelson, and H. Rauhut, "Suprema of chaos processes and the restricted isometry property," *Communications on Pure and Applied Mathematics*, vol. 67, no. 11, pp. 1877–1904, 2014.
- [35] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Birkhauser Basel, 2013.
- [36] M. Rudelson and R. Vershynin, "Hanson-wright inequality and sub-gaussian concentration," *Electron. Commun. Probab.*, vol. 18, pp. no. 82, 1–9, 2013.
- [37] J. Tropp, "Norms of random submatrices and sparse approximation," *Comptes Rendus Mathematique*, vol. 346, no. 2324, pp. 1271–1274, 2008.
- [38] E. van den Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, <http://www.cs.ubc.ca/labs/scl/spgl1>.
- [39] —, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.
- [40] J. Tropp, "An Introduction to Matrix Concentration Inequalities," *ArXiv e-prints*, Jan. 2015.
- [41] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices -Compressed sensing, 210–268*. Cambridge, USA: Cambridge Univ. Press, 2012.