Analysis of Frequency Agile Radar via Compressed Sensing

Tianyao Huang ⓘ, Yimin Liu ⓘ, *Member, IEEE*, Xingyu Xu ⓘ, Yonina C. Eldar ⓘ, *Fellow, IEEE*, and Xiqin Wang

Abstract—**Frequency agile radar (FAR) is known to have excellent electronic counter-countermeasures performance and the potential to realize spectrum sharing in dense electromagnetic environments. Many compressed sensing (CS) based algorithms have been developed for joint range and Doppler estimation in FAR. This paper considers theoretical analysis of FAR via CS algorithms. In particular, we analyze the properties of the sensing matrix, which is a highly structured random matrix. We then derive bounds on the number of recoverable targets. Numerical simulations and field experiments validate the theoretical findings and demonstrate the effectiveness of CS approaches to FAR.**

Index Terms—**Frequency agile radar, compressed sensing, performance guarantee, electronic counter-countermeasures.**

I. INTRODUCTION

FREQUENCY agile radars (FARs) are pulse-based radars, in which the carrier frequencies are varied in a random/pseudo-random manner from pulse to pulse as illustrated in Fig. 1. Each transmission occupies a narrow band (B_0). Pulse returns of different frequencies are processed coherently to synthesize a wider band ($B > B_0$), which generates high range resolution (HRR) profiles.

Since the works [2], [3], frequency agility has received increasing attention [4]–[10] in the radar community due to its multi-fold merits. First, frequency agility introduces good electronic counter-countermeasures (ECCM) performance, because the randomly varied frequencies of the pulses are difficult to track and predict. In addition, the flexibility of the narrow band transmission makes it easier to avoid and reject barrage jamming. Second, like stepped frequency radar, FAR can be used for two-dimensional (2D) imaging [8], [9], while requiring only a narrow band receiver, which significantly lowers the hardware

Fig. 1. An example of a FAR waveform. The boxes indicate the frequency band transmitted in the given time window.

system cost. Third, in contrast to linearly stepped frequency radar, FAR decouples the range-Doppler parameters and produces a thumbtack ambiguity function [4]. It also mitigates aliasing artifacts in synthetic aperture radar (SAR) [11] and extends the unambiguous Doppler window in inverse SAR (ISAR) imaging [5]. Finally, frequency agility can be utilized to exploit vacant spectral bands [10], and shows potential to increase spectrum efficiency and cope with spectrum sharing issues in a contested, congested and competitive electromagnetic environment.

We consider the problem of joint HRR profile and Doppler estimation of the target. When a traditional matched filter is used for joint range-Doppler estimation, sidelobe pedestal problems occur in FAR [4]. Therefore, weak targets could be masked by the sidelobe of dominant ones, which restricts the application of FAR in target detection and feature extraction [4]. By exploiting target sparsity, compressed sensing (CS) techniques [12], [13] have been applied in order to alleviate the sidelobe pedestal problem. *Liu et al.* [4] propose the RV-IAP algorithm for joint range-Doppler estimation, which is based on the Orthogonal Matching Pursuit (OMP) method [14]. Since then, many practical CS algorithms for FAR have been developed [6], [15].

This paper focuses on the theoretical analysis of CS methods for FAR in terms of reconstruction performance. Theoretical conditions that guarantee perfect recovery in general CS have been extensively studied. Randomness plays a key role in many theoretical results, and often leads to good empirical results [12]. Near optimal conditions for Gaussian, sub-Gaussian, Bernoulli and random Fourier matrices have been derived [16], [17] (and references therein). However, the measurement matrix in FAR differs from these random matrices, so that previous theoretical results are not directly applicable.

We start by studying the measurement matrix properties of a FAR system. This matrix is random, due to the randomness of the frequencies. We begin by deriving probability bounds on the spark and coherence of the measurement matrix, extending the results of [1]. Based on these bounds, we develop recovery guarantees for joint HRR profile and Doppler estimation using FAR systems. Theoretical results show that owing to the randomness of the carrier frequencies, with high probability, one can jointly obtain a HRR profile and Doppler of targets while transmitting narrow-band pulses. The number of recoverable targets is proved to be $N/2$ using ℓ_0 minimization, or on the order of $\sqrt{\frac{N}{\log(NB/B_0)}}$ using ℓ_1 minimization, where N is the number of pulses.

We next perform simulations and field experiments to demonstrate the reconstruction performance of FAR using CS methods. We build an X-band FAR prototype with a synthetic bandwidth of 1 GHz, and test the recovery performance in a real environment. The results show that both the HRR profiles and Doppler of the observed target (a moving car) are reconstructed with $N = 512$ and $B/B_0 = 32$.

The rest of paper is organized as follows. Section II introduces the signal model and problem formulation. In Section III, a brief review of CS algorithms and their performance guarantees is provided. We derive conditions for joint range-Doppler recovery using FAR in Section IV. Numerical simulation and field experiment results are shown in Section V. Section VI concludes the paper.

Throughout the paper we use the following notation. The sets $\mathbb{C}, \mathbb{R}, \mathbb{Z}, \mathbb{N}$ refer to complex, real, integer, and natural numbers. Notation $|\cdot|$ is used for the modulus, absolute value or cardinality for a complex, real valued number, or a set, respectively, and $j := \sqrt{-1}$. For $x \in \mathbb{R}$, $\lfloor x \rfloor$ (or $\lceil x \rceil$) is the largest (smallest) integer less (greater) than or equal to x. Uppercase boldface letters denote matrices (e.g., $\boldsymbol{A}$), and lowercase boldface letters denote vectors (e.g., $\boldsymbol{a}$). The m, n-th element of matrix $\boldsymbol{A}$ is written as $[\boldsymbol{A}]_{m,n}$, and $[\boldsymbol{a}]_n$ denotes the n-th entry of a vector. Given a matrix $\boldsymbol{A} \in \mathbb{C}^{M \times N}$, a number n (or a set of integers, Λ), $\boldsymbol{A}_n$ ($\boldsymbol{A}_\Lambda \in \mathbb{C}^{M \times |\Lambda|}$) denotes the n-th column of $\boldsymbol{A}$ (the submatrix consisting of the columns of $\boldsymbol{A}$ indexed by Λ). As for a vector $\boldsymbol{a} \in \mathbb{C}^N$, $\boldsymbol{a}_\Lambda \in \mathbb{C}^{|\Lambda|}$ denotes the sub-vector consisting of the elements of $\boldsymbol{a}$ indexed by Λ. The complex conjugate operator, transpose operator, and the complex conjugate-transpose operator are *, ^{T}, ^{H}, respectively. We use $\|\cdot\|_p$, $p = 1, 2$ as the ℓ_p norm of an argument, and $\mathbb{P}(\cdot)$ denotes the probability of an event. Operations $\mathrm{E}[\cdot]$ and $\mathrm{D}[\cdot]$ represent the expectation and variance of a random argument, respectively. The real and imaginary part of a complex valued argument are denoted by $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$, respectively.

II. SIGNAL MODEL

A. Radar Returns Model

In this section, we introduce FAR, following the presentation in [6]. A FAR system transmits monotone pulses, where the n-th transmitted pulse is written as

$$T_x(n, t) := \mathrm{rect}\left(\frac{t - nT_r}{T_p}\right) e^{j 2\pi f_n (t - nT_r)}, \quad (1)$$

$n = 0, 1, \ldots, N - 1$, where T_r and T_p are the pulse repetition interval and pulse duration, respectively, $T_r > T_p$, and $\mathrm{rect}(\cdot)$ represents the rectangular envelope of the pulse

$$\mathrm{rect}(x) := \begin{cases} 1, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

The frequency of the n-th pulse f_n is randomly varied as $f_n = f_c + d_n B$, where f_c is the initial frequency, d_n is the n-th random frequency-modulation code, $0 \le d_n \le 1$, and B is the synthetic bandwidth. For a single pulse, the bandwidth ($B_0 = 1/T_p$) is narrow, $B_0 < B$, and the coarse range resolution (CRR) is $\frac{T_p c}{2}$, where c is the speed of light. Synthesizing echoes of different frequencies refines the range resolution to $\frac{c}{2B}$. We denote the number of HRR bins inside a CRR bin as

$$M := \left\lceil \frac{T_p c}{2} \cdot \frac{2B}{c} \right\rceil = \lceil T_p B \rceil \in \mathbb{N}. \quad (3)$$

Received echoes are assumed delays of the transmissions. We begin by assuming that there is a single ideal scatterer with scattering coefficient $\beta \in \mathbb{C}$. The echo of the n-th pulse can then be written as

$$R_x(n, t) := \beta T_x\left(n, t - \frac{2r(t)}{c}\right), \quad (4)$$

where $r(t)$ denotes the range of the scatterer with respect to the radar at time instant t. We assume that the scatterer is moving along the line of sight at a constant speed v, so that $r(t) = r(0) + vt$. After down conversion, the echo becomes

$$\begin{aligned} R_d(n, t) &:= R_x(n, t) \cdot e^{-j 2\pi f_n (t - nT_r)} \\ &= \beta \mathrm{rect}\left(\frac{t - 2r(t)/c - nT_r}{T_p}\right) e^{j 2\pi f_n (t - 2r(t)/c - nT_r)} \\ &\quad \cdot e^{-j 2\pi f_n (t - nT_r)} \\ &= \beta \mathrm{rect}\left(\frac{t - 2r(t)/c - nT_r}{T_p}\right) e^{-j 2\pi f_n \frac{2r(t)}{c}}. \end{aligned} \quad (5)$$

Echoes are sampled at the Nyquist rate of a single pulse, $f_s = 1/T_p$, so that each echo pulse is sampled once. Every sample corresponds to a CRR bin, and data from all CRR bins are processed in the same way. Returns of N pulses from the same CRR bin are combined to a vector

$$[R_d(0, t), R_d(1, T_r + t), \ldots, R_d(N - 1, (N - 1)T_r + t)], \quad (6)$$

and processed to generate HRR profiles and Doppler estimates. During the coherent processing interval (CPI), i.e. NT_r, we assume that the scatterer does not cross a CRR bin, which means that

$$vNT_r < \frac{T_p c}{2}. \quad (7)$$

Without loss of generality, suppose that the l-th CRR bin contains the scatterer, $l = 0, 1, \ldots, \lfloor T_r f_s \rfloor$. The corresponding sampling instant for the n-th pulse is $t = nT_r + l/f_s$. Substituting $t = nT_r + l/f_s$ into (5), the sampled echoes are given

by

$$R_d(n, nT_r + l/f_s) = \beta e^{-j4\pi(f_c + d_n B)(r(0) + v(nT_r + l/f_s))/c}$$

$$\approx \beta e^{-j4\pi f_c \frac{r(0)+vl/f_s}{c}} e^{-j4\pi d_n B r(0)/c} e^{-j4\pi f_c v T_r n\zeta_n/c}, \quad (8)$$

where the approximation holds if the term $e^{-j4\pi d_n Bvl/(f_s c)} \approx 1$, which requires $e^{-j4\pi Bv T_r/c} \approx 1$. Here $\zeta_n := 1 + d_n B/f_c$. Generally, different carrier frequencies imply different Doppler shifts, unless the relative bandwidth B/f_c is negligible, i.e. $\zeta_n \approx 1$. However, in a (synthetic) wideband radar, this approximation does not usually hold, and could give rise to estimation performance deterioration in practice if applied. In the simulations and field experiments in Section V, the signal processing algorithms do not adopt this assumption. However, in the mathematical analysis in Section IV, we assume $\zeta_n \approx 1$ for theoretical convenience. The impact of the relative bandwidth will be discussed in the simulations.

For brevity, we omit the notation l, and write $R_d(n) := R_d(n, nT_r + l/f_s)$. We further introduce notations

$$\begin{cases} \tilde{\gamma} := \beta e^{-j4\pi f_c \frac{r(0)+vl/f_s}{c}}, \\ \tilde{p} := -4\pi B r(0)/(Mc), \\ \tilde{q} := -4\pi f_c v T_r/c. \end{cases} \quad (9)$$

With these definitions (8) becomes

$$R_d(n) \approx \tilde{\gamma} e^{j\tilde{p}Md_n + j\tilde{q}n\zeta_n}. \quad (10)$$

After the unknowns $\tilde{\gamma}$, $\tilde{p}$ and $\tilde{q}$ are estimated, the absolute intensity $|\beta|$, HRR range $r(0)$ and velocity v are inferred as $|\tilde{\gamma}|$, $-\frac{Mc\tilde{p}}{4\pi B}$ and $-\frac{c\tilde{q}}{4\pi f_c T_r}$, respectively.

When there are K scatterers occurring inside the CRR cell, radar returns are modeled as a combination of returns from all scatterers,

$$R_d(n) = \sum_{k=0}^{K-1} \tilde{\gamma}_k e^{j\tilde{p}_k Md_n + j\tilde{q}_k n\zeta_n}, \quad (11)$$

where $\tilde{\gamma}$, $\tilde{p}$ and $\tilde{q}$ in (10) are replaced with $\tilde{\gamma}_k$, $\tilde{p}_k$ and $\tilde{q}_k$ for the k-th scatterer, respectively.

To avoid grating lobes in the HRR profiles (which are also called ghost images in the literature) [18], [19], the frequency codes are required to satisfy $\min_{n\neq m}|d_n - d_m| \leq 1/M$, $n, m = 0, 1, \ldots, N-1$. When codes are discrete, we denote by $\mathcal{D}_d$ the set of available frequency codes and by $M^\star := |\mathcal{D}_d|$ the number of codes. The codes are often uniformly spaced, e.g. $d_n \in \mathcal{D}_d := \{\frac{m}{M^\star}|m = 0, 1, \ldots, M^\star - 1\}$. It is required that $M^\star \geq M$ and a typical choice is $M^\star = M$. When codes belong to a continuous set $\mathcal{D}_c := [0, 1)$ (for some of our theoretical results), the requirement is usually easy to satisfy with $N \geq M$. We assume that in both discrete and continuous cases, the codes $d_0, \ldots d_{N-1}$ are identically, independently, and uniformly distributed.

B. Signal Model in Matrix Form

We can rewrite (11) in matrix form as

$$y = \Phi x, \quad (12)$$

where the measurement vector $y \in \mathbb{C}^N$ has entries $[y]_n = R_d(n)$. The vector $x \in \mathbb{C}^{NM}$ corresponds to the scattering intensities $\tilde{\gamma}$. The pair $(\tilde{p}, \tilde{q})$ defines the key target parameters, range and Doppler, and belongs to a continuous 2D domain. The resolutions for $\tilde{p}$ and $\tilde{q}$ are $\frac{2\pi}{M}$ and $\frac{2\pi}{N}$, respectively. Consider the unambiguous continuous region $(p, q) \in [0, 2\pi)^2$, and discretize p and q at the Nyquist rates, $\frac{2\pi}{M}$ and $\frac{2\pi}{N}$, respectively. Thus, one obtains $p_m := \frac{2\pi m}{M}$ and $q_n := \frac{2\pi n}{N}$, $m = 0, 1, \ldots, M-1$, $n = 0, 1, \ldots, N-1$. Denote the sets containing HRR grids and Doppler grids as $\mathcal{P} := \{\frac{2\pi m}{M}|m = 0, 1, \ldots, M-1\}$ and $\mathcal{Q} := \{\frac{2\pi n}{N}|n = 0, 1, \ldots, N-1\}$, respectively, and assume that the targets are located precisely on the grid. Define the matrix $X \in \mathbb{C}^{M \times N}$ with entries

$$[X]_{m,n} = \begin{cases} \tilde{\gamma}_k, & \text{if } \exists k, (\tilde{p}_k, \tilde{q}_k) = (p_m, q_n), \\ 0, & \text{otherwise,} \end{cases} \quad (13)$$

representing the 2D scattering coefficients in the range-Doppler domain, $m = 0, 1, \ldots, M-1$ and $n = 0, 1, \ldots, N-1$. We vectorize X to obtain $x := \text{vec}(X^T)$ with entries $[x]_{n+mN} := [X]_{m,n}$.

To introduce the measurement matrix $\Phi \in \mathbb{C}^{N \times MN}$, we define the matrices $R \in \mathbb{C}^{N \times M}$ and $D \in \mathbb{C}^{N \times N}$, corresponding to HRR range and Doppler parameters, respectively, with entries

$$[R]_{n,m} := e^{jp_m Md_n}, \quad (14)$$

$$[D]_{n,l} := e^{jq_l n\zeta_n}, \quad (15)$$

$m = 0, 1, \ldots, M-1$, and $l, n = 0, 1, \ldots, N-1$. If $\zeta_n \approx 1$, then D is a Fourier matrix. Define $\Phi := (R^T \odot D^T)^T$, where $\odot$ denotes the Khatri-Rao product. Then the elements of Φ are given by

$$[\Phi]_{n,l+mN} := [R]_{n,m}[D]_{n,l} = e^{jp_m Md_n + jq_l n\zeta_n}, \quad (16)$$

$m = 0, 1, \ldots, M-1$ and $l, n = 0, 1, \ldots, N-1$. When echoes are corrupted by additive noise $w \in \mathbb{C}^N$, (12) becomes

$$y = \Phi x + w. \quad (17)$$

The sensing matrix Φ in (17) has more columns than rows, $MN \geq N$, which shows that joint range and Doppler estimation in FAR is naturally an under-determined problem. When x is K-sparse, which means there are K non-zeroes in x, and $K \ll MN$, CS algorithms can be used to solve (17). The targets' parameters can then be recovered from the support set of x.

C. Discussion on the Signal Model

Note that when there is only one scatterer observed, the matched filter that maximizes the signal to noise ratio (SNR) works well in FAR. However, when there are multiple scatterers, sidelobe pedestal problem occurs and weak targets can be masked by the dominant targets' sidelobe. The matched filter estimates the scattering intensities by

$$\hat{x} := \Phi^H y = \Phi^H \Phi x + \Phi^H w. \quad (18)$$

In such an under-determined model, $\Phi^H \Phi \neq I$, spurious responses emerge in $\hat{x}$ even if there is no noise, i.e. $w = 0$. These spurious responses are the sidelobe pedestal.

To better interpret the sidelobe pedestal problem, we compare the signal model of FAR with that of an instantaneous wideband radar (IWR). In such a hypothetical radar, we assume that the radar transmits/receives all of its M sub-bands (with $\mathcal{D}_d$ as the set of frequency codes, $|\mathcal{D}_d| = M^\star = M$), and processes the echoes individually for each band. In FAR, the same set $\mathcal{D}_d$ is also applied with $Md_n \in \mathbb{N}$. In analogy to (11), the return of the m-th frequency in the n-th pulse can be written as

$$R_{\text{IWR}}(m,n) := \sum_{k=0}^{K-1} \tilde{\gamma}_k e^{j\tilde{p}_k m + j\tilde{q}_k n \eta_m}, \tag{19}$$

where $\eta_m := 1 + \frac{mB}{Mf_c}$, $m = 0, 1, \ldots, M-1$, and $n = 0, 1, \ldots, N-1$. For notational brevity and simplicity, we assume $\eta_m \approx 1$ and $\zeta_n \approx 1$ for IWR and FAR, respectively. In this case, (19) can be rewritten in matrix form as

$$\boldsymbol{Z} = \boldsymbol{F} \boldsymbol{X} \boldsymbol{D}^T, \tag{20}$$

where $\boldsymbol{Z} \in \mathbb{C}^{M \times N}$ has entries $[\boldsymbol{Z}]_{m,n} = R_{\text{IWR}}(m,n)$, and $\boldsymbol{F} \in \mathbb{C}^{M \times M}$ is a Fourier matrix with entries $[\boldsymbol{F}]_{l,m} := e^{jp_m l}$, $m, l = 0, 1, \ldots, M-1$, $n = 0, 1, \ldots, N-1$. Equivalently,

$$\boldsymbol{z} = (\boldsymbol{F} \otimes \boldsymbol{D}) \boldsymbol{x}, \tag{21}$$

where $\boldsymbol{z} := \text{vec}(\boldsymbol{Z}^T) \in \mathbb{C}^{MN}$ and $\otimes$ denotes the Kronecker product. The sensing matrix in the IWR $\boldsymbol{\Psi} := \boldsymbol{F} \otimes \boldsymbol{D} \in \mathbb{C}^{MN \times MN}$ is orthogonal, i.e. $\frac{1}{MN} \boldsymbol{\Psi}^H \boldsymbol{\Psi} = \boldsymbol{I}$, and the sidelobe pedestal problem vanishes.

The measurements in FAR can be regarded as sampling[1] of the IWR measurements, i.e.

$$[\boldsymbol{y}]_n = [\boldsymbol{Z}]_{Md_n, n}, \ n = 0, 1, \ldots, N-1. \tag{22}$$

Only one sub-band data is acquired for each pulse. Therefore the sensing matrix of FAR consists of partial rows of that of the IWR, i.e.,

$$\left(\boldsymbol{\Phi}^T\right)_n = \left(\boldsymbol{\Psi}^T\right)_{n+Md_n N}, \ n = 0, 1, \ldots, N-1, \tag{23}$$

and becomes an under-determined matrix. This interpretation suggests that the sidelobe pedestal of FAR results from the information loss in the frequency domain. The spectral incompleteness leads to an under-determined problem (12). In Section IV, we prove that, owing to the randomness of the frequencies, the scatterers can still be correctly reconstructed via CS methods with high probability.

III. REVIEW OF COMPRESSED SENSING

In Section IV, we prove that using CS methods, FAR can provably recover the HRR profiles and Doppler. Before deriving the results, we review some basic notions of CS [13].

Consider an under-determined linear regression problem, e.g. (12), where $\boldsymbol{x}$ is sparse. The sparsest solution can be obtained

---

[1]Since an instantaneous narrowband waveform is used in FAR, it naturally enjoys low data rate in comparison with IWR. However, this paper does not aim at minimizing the data rate. It may have the potential to further reduce the data rate by combining frequency agility with approaches like sub-Nyquist sampling in the fast-time domain [10], [20], [21], omitting some pulses or frequency bands [22], [23].

via

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0, \ \text{s.t. } \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \tag{P_0}$$

where $\| \cdot \|_0$ denotes ℓ_0 "norm" of a vector, i.e. the number of non-zeroes. This solution is the true vector, when the sensing matrix $\boldsymbol{A}$ has the spark property.

Definition 1 (Spark, [13]): Given a matrix $\boldsymbol{A}$, Spark($\boldsymbol{A}$) is the smallest possible number such that there exists a subgroup of columns from $\boldsymbol{A}$ that are linearly dependent.

Unique recovery of $\boldsymbol{x}$ can be ensured if the following condition is satisfied.

Theorem 2: The equation $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ is uniquely solved by (P_0) if and only if $\|\boldsymbol{x}\|_0 < \frac{\text{Spark}(\boldsymbol{A})}{2}$.

The above theorem provides a fundamental limit on the maximum sparsity that leads to successful recovery. In general, ℓ_0 optimization is NP-hard. A widely used alternative is basis pursuit, which solves the problem

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1, \ \text{s.t. } \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \tag{P_1}$$

In noisy cases, variants like basis pursuit denoising, LASSO and Dantzig selector can be applied. Many greedy methods have also been suggested to approximate (P_0).

Sufficient conditions that guarantee uniqueness using these methods are extensively studied. Bounds on the mutual incoherence property (MIP) and restricted isometry property (RIP) are widely applied conditions to ensure sparse recovery. In this paper, we rely on the MIP. A matrix $\boldsymbol{A}$ has MIP if its coherence is small, where coherence is defined as the maximum correlation between two columns, i.e.

$$\mu(\boldsymbol{A}) := \max_{l \neq k} \frac{|\boldsymbol{A}_l^H \boldsymbol{A}_k|}{\|\boldsymbol{A}_l\|_2 \|\boldsymbol{A}_k\|_2}. \tag{24}$$

Theorem 3 ([24]): If a matrix $\boldsymbol{A} \in \mathbb{C}^{N \times L}$ has coherence $\mu(\boldsymbol{A}) < \frac{1}{2K-1}$, then for any $\boldsymbol{x} \in \mathbb{C}^L$ of sparsity K, $\boldsymbol{x}$ is the unique solution to (P_1).

The condition in Theorem 3 ensures recovery in the presence of noise and also recovery using a variety of computationally efficient methods [12].

IV. SENSING MATRIX PROPERTIES OF FAR

In this section, we analyze the spark and MIP properties of the FAR's sensing matrix. These results are then used together with Theorems 2 and 3 to establish performance guarantees for FAR. In the following derivations, we assume that $\zeta_n \approx 1$.

A. Spark Property

The following theorem proves that the sensing matrix of FAR almost surely has the spark property.

Theorem 4: Consider $\boldsymbol{\Phi} \in \mathbb{C}^{N \times NM}$ defined in (16) with d_n drawn independently from a uniform continuous distribution over $\mathcal{D}_c = [0, 1)$, $n = 0, 1, \ldots, N-1$. Then, with probability 1, Spark($\boldsymbol{\Phi}$) = $N + 1$.

Proof: See Appendix A. ∎

Since $\boldsymbol{\Phi}$ has N rows, there must be a linearly dependent submatrix with $N + 1$ columns. Owing to the randomness of

the carrier frequencies, Theorem 4 shows that a sub-matrix built from any N columns of $\boldsymbol{\Phi}$ is of full rank almost surely. The result is based on the assumption of a continuous distribution on d_n. An immediate consequence of Theorems 2 and 4 is the following corollary.

Corollary 5: Consider a FAR whose frequency modulation codes are drawn independently from a uniform continuous distribution over $\mathcal{D}_c = [0, 1)$, $n = 0, 1, \dots, N-1$. Then, with probability 1, $K_{\max} = \frac{N}{2}$ scatterers can be exactly recovered by (P_0), where N is the number of pulses.

B. Mutual Incoherence Property

To obtain performance guarantees using ℓ_1 minimization or greedy CS methods under noiseless/noisy environments, we derive the MIP for FAR. We start by analyzing the asymptotic statistics of the FAR's sensing matrix. Then invoking Theorem 3, we obtain the maximum number of scatterers that FAR guarantees to exactly reconstruct with high probability.

Assume in this subsection that $d_n \sim U(\mathcal{D}_d)$, $|\mathcal{D}_d| = M^\star = M$, and recall that the parameters p and q are on a grid, i.e. $p \in \mathcal{P}$ and $q \in \mathcal{Q}$. First, consider the Gram matrix $\boldsymbol{\Phi}^H \boldsymbol{\Phi}$, which links to the coherence. Define $\boldsymbol{G} \in \mathbb{R}^{NM \times NM}$ as the modulus matrix of the Gram matrix, i.e.

$$[\boldsymbol{G}]_{k,l} = \left| \left[\boldsymbol{\Phi}^H \boldsymbol{\Phi} \right]_{k,l} \right|, k, l = 0, 1, \dots, NM-1. \quad (25)$$

We then have the following results, some of which are partially inspired by [25].

Lemma 6: The rows of the modulus matrix $\boldsymbol{G}$ are permutations of elements in its first row.

Proof: Denote by $\boldsymbol{\Phi}_{l_1}$ and $\boldsymbol{\Phi}_{l_2}$, $l_1, l_2 = 0, 1, \dots, MN-1$, two columns in $\boldsymbol{\Phi}$, corresponding to (p_{m_1}, q_{k_1}) and (p_{m_2}, q_{k_2}), respectively, $m_1, m_2 = 0, 1, \dots, M-1$, $k_1, k_2 = 0, 1, \dots, N-1$. Then

$$
\begin{aligned}
\boldsymbol{\Phi}_{l_1}^H \boldsymbol{\Phi}_{l_2} &= \sum_{n=0}^{N-1} e^{-j p_{m_1} M d_n - j q_{k_1} n} \cdot e^{j p_{m_2} M d_n + j q_{k_2} n} \\
&= \sum_{n=0}^{N-1} e^{-j(p_{m_1} - p_{m_2}) M d_n - j(q_{k_1} - q_{k_2}) n} \\
&= \sum_{n=0}^{N-1} e^{-j 2\pi (m_1 - m_2) d_n - j \frac{2\pi (k_1 - k_2)}{N} n}. \quad (26)
\end{aligned}
$$

Clearly (26) depends only on the difference between grid points, i.e. $m_1 - m_2 \in \{-M+1, \dots, M-1\}$ and $k_1 - k_2 \in \{-N+1, \dots, N-1\}$, and is independent of the particular indices l_1 and l_2. In addition, $|\boldsymbol{\Phi}_{l_1}^H \boldsymbol{\Phi}_{l_2}| = |\boldsymbol{\Phi}_{l_2}^H \boldsymbol{\Phi}_{l_1}|$. Therefore, for any element in $\boldsymbol{G}$, one can find an element in the first row of $\boldsymbol{G}$ with the same value. ∎

Consider now the l-th element in the 0-th row of $\boldsymbol{G}$, $l \neq 0$, which corresponds to the l-th column of $\boldsymbol{\Phi}$. Note that each column of $\boldsymbol{\Phi}$ relies on a specific parameter pair (p, q). For notational simplicity, we drop the subscripts of (p, q) related to

$\boldsymbol{\Phi}_l$, and define

$$\chi_l := \frac{1}{N} \boldsymbol{\Phi}_0^H \boldsymbol{\Phi}_l = \frac{1}{N} \sum_{n=0}^{N-1} e^{j p M d_n + j q n}, \, l = 1, \dots, NM-1. \quad (27)$$

We now analyze the mutual coherence, $\mu = \max_{l \neq 0} |\chi_l|$. Since d_n is random, χ_l is also random unless $p = 0$, in which case χ_l reduces to a constant $\frac{1}{N} \sum_{n=0}^{N-1} e^{jqn} = 0$ for $q \in \mathcal{Q} \backslash \{0\}$. This constant does do not affect the value of μ and is thus ignored. Define a set excluding these constants as

$$
\begin{aligned}
\Xi &:= \{1, 2, \dots, NM-1\} \backslash \{1, \dots, N-1\} \\
&= \{N, N+1, \dots, NM-1\}. \quad (28)
\end{aligned}
$$

Then χ_l is a random variable, $l \in \Xi$, and has the following statistical characteristics.

Lemma 7: As $N \to \infty$, the real and imaginary parts of χ_l, $\mathrm{Re}(\chi_l)$ and $\mathrm{Im}(\chi_l)$, $l \in \Xi$, have a joint Gaussian distribution,

$$
\begin{bmatrix} \mathrm{Re}(\chi_l) \\ \mathrm{Im}(\chi_l) \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2N} & 0 \\ 0 & \frac{1}{2N} \end{bmatrix} \right), \quad (29)
$$

except in the special case that the corresponding parameters $p = q = \pi$. In this setting, the joint Gaussian distribution becomes

$$
\begin{bmatrix} \mathrm{Re}(\chi_l) \\ \mathrm{Im}(\chi_l) \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{N} & 0 \\ 0 & 0 \end{bmatrix} \right). \quad (30)
$$

Proof: See Appendix B. ∎

The special case $p = q = \pi$ corresponds to a specific grid point of the (p, q) plane. Considering the generic case and this special case separately leads to the following conclusions.

Corollary 8: When $N \to \infty$ and $l \in \Xi$, with $N\epsilon^2 > 2/\pi$,

$$\mathbb{P}(|\chi_l| > \epsilon) \leq e^{-N\epsilon^2/2}. \quad (31)$$

Proof: Lemma 7 proves that when p and q do not equal π simultaneously, the real and imaginary parts of χ_l asymptotically obey $\mathcal{N}(0, \frac{1}{2N})$ independently. Therefore, the magnitude $|\chi_l|$ obeys a Rayleigh distribution with probability density function $f(x) = 2Nx e^{-Nx^2}$, $x \geq 0$, and cumulative distribution function $F(x) = 1 - e^{-Nx^2}$, $x \geq 0$. Thus, $\mathbb{P}(|\chi_l| > \epsilon) = 1 - F(\epsilon)$, which yields

$$\mathbb{P}(|\chi_l| > \epsilon) = e^{-N\epsilon^2} \leq e^{-N\epsilon^2/2}. \quad (32)$$

In the special case that $p = q = \pi$, the real part of χ_l asymptotically obeys $\mathcal{N}(0, \frac{1}{N})$ independently and the imaginary part vanishes. Then elementary estimates of the Gaussian error function yield

$$\mathbb{P}(|\chi_l| > \epsilon) \leq \sqrt{\frac{2}{\pi N \epsilon^2}} e^{-N\epsilon^2/2}, \quad (33)$$

which is less than $e^{-N\epsilon^2/2}$ if $\frac{2}{\pi N \epsilon^2} \leq 1$, i.e. $N\epsilon^2 > \frac{2}{\pi}$. ∎

Lemma 9: The maximum $\mu = \max_l |\chi_l|$, $l \in \Xi$, satisfies the following as $N \to \infty$,

$$\mathbb{P}(\mu > \epsilon) \leq (MN - N) e^{-N\epsilon^2/2}. \quad (34)$$

Proof: For fixed $\epsilon > 0$, we have $N\epsilon^2 > \frac{2}{\pi}$ as $N \to \infty$. According to the union bound

$$\mathbb{P}(\mu > \epsilon) \leq \sum_{l \in \Xi} \mathbb{P}(|\chi_l| > \epsilon)$$

$$\leq (MN - N)e^{-N\epsilon^2/2}, \qquad (35)$$

since there are $NM - N$ indices in Ξ. ∎

We next derive a condition for FAR to meet the requirement of Theorem 3 $\mu(\mathbf{\Phi}) < \frac{1}{2K-1}$ with high probability.

Theorem 10: The coherence of $\mathbf{\Phi}$, defined in (16), obeys $\mu(\mathbf{\Phi}) < \frac{1}{2K-1}$ with a probability higher than $1 - \delta$, when

$$K \leq \frac{1}{2\sqrt{2}} \sqrt{\frac{N}{\log(MN - N) - \log \delta}} + \frac{1}{2}. \qquad (36)$$

Proof: Let $\epsilon = \frac{1}{2K-1}$. From (36), we have that

$$N\epsilon^2 \geq 2 \left(\log(MN - N) - \log \delta \right). \qquad (37)$$

Assume that $0 < \delta < 1$ and $MN - N = N(M - 1) \geq 2$. Then

$$N\epsilon^2 \geq 2 \log 2 > \frac{2}{\pi}. \qquad (38)$$

Using (34) and (37), we finally obtain

$$\mathbb{P}(\mu \leq \epsilon) > 1 - (MN - N)e^{-N\epsilon^2/2} \geq 1 - \delta, \qquad (39)$$

completing the proof. ∎

Theorem 10 shows that the sensing matrix of FAR has the MIP; thus, according to Theorem 3, HRR range-Doppler reconstruction is guaranteed if the number of targets satisfies $K = O(\sqrt{\frac{N}{\log MN}})$, where N in the numerator represents the number of measurements, and MN in the denominator links to the number of grid points. The work [25] for direction of arrival estimation using random array MIMO radar also proposes a bound based on MIP, which guarantees recovery of a number of targets on the order of $K = O(\frac{\sqrt{L}}{\log G})$, where L and G are the number of measurements and grid points, respectively. In [25], a bound for non-uniform recovery based on RIPless theory is also provided, and K is relaxed to $K = O(\frac{L}{\log^2 G})$. However, RIPless is not directly applicable to FAR, because for each row of $\mathbf{\Phi}$, i.e. $\boldsymbol{a} = (\mathbf{\Phi}^T)_n \in \mathbb{C}^{MN}$, $\mathrm{E}[\boldsymbol{a}\boldsymbol{a}^H]$ is rank deficient and the isotropy property $\mathrm{E}[\boldsymbol{a}\boldsymbol{a}^H] = \boldsymbol{I}$ does not hold. Dorsch and Rauhut [26] analyze the joint angle-delay-Doppler recovery performance using MIMO based on RIP. They assume a periodic random probing signal with N_t independent samples. In this case the recoverable number of scatterers is on the order of $K = O(\frac{N_t}{\log^2 G})$. Though the RIP leads to a tighter bound than MIP, the RIP of the FAR system matrix is still an open question.

Sufficient conditions that guarantee uniform recovery are usually pessimistic. It is well known that CS algorithms often outperform the theoretical uniform recovery guarantees. In the next section, we evaluate the practical performance of FAR using CS methods.

V. SIMULATION AND EXPERIMENTAL RESULTS

In this section, simulations and field experiments are executed to demonstrate the properties of the sensing matrix of FAR and
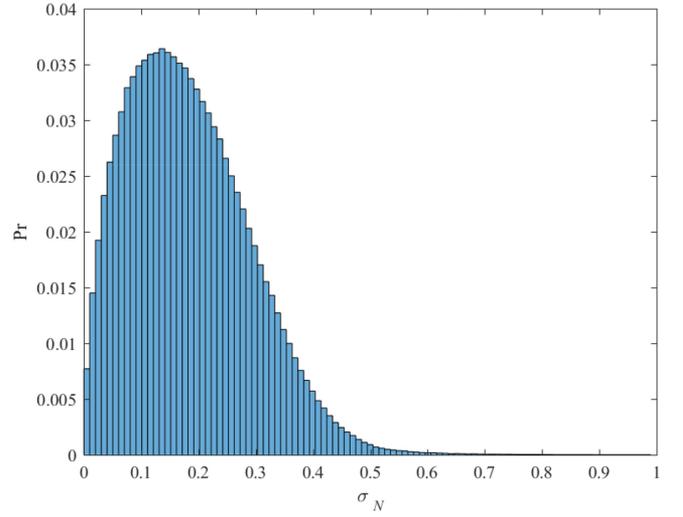
Fig. 2. Histogram of σ_N (normalized by $\sqrt{N}$), the minimum singular value of each $N \times N$ sub-matrix. The histogram is obtained using 2000 Monte-Carlo trials and $\binom{NM}{N}$ sub-matrices in each Monte-Carlo trial.

the effectiveness of CS algorithms to reconstruct the targets' HRR range and Doppler.

A. Spark Property

First, the spark property of the sensing matrix $\mathbf{\Phi} \in \mathbb{C}^{N \times NM}$ is discussed. We construct a sub-matrix $\mathbf{\Phi}_\Omega \in \mathbb{C}^{N \times N}$ of $\mathbf{\Phi}$, where the set $\Omega \subset \{0, 1, \dots, NM - 1\}$ and $|\Omega| = N$, and calculate the minimum singular value of $\mathbf{\Phi}_\Omega$. We check whether it is equal to zero (rank deficient). Concretely, we set $N = 6$ and $M = 3$, which are small to make it possible to enumerate all the $\binom{NM}{N}$ sub-matrices of $\mathbf{\Phi}$. We record the minimum singular value σ_N (normalized by $\sqrt{N}$) of each sub-matrix; thus, we obtain $\binom{NM}{N}$ results, among which the minimum is denoted as σ_Ω. The frequency codes d_n are distributed uniformly on a continuous set. We further assume the relative bandwidth satisfies $B/f_c \approx 0$. We perform 2000 Monte-Carlo trials. The histograms of σ_N and σ_Ω are depicted in Fig. 2 and Fig. 3, respectively. The minimum of σ_N and σ_Ω is $1.28 \times 10^{-6} > 0$. The results indicate that a continuous distribution of codes results in good properties of the sensing matrix. For comparison, we also perform simulations with codes distributed on the discrete set $\mathcal{D}_d$, and count the number of minimum singular values σ_N, i.e. σ_Ω, less than $\epsilon_{\mathrm{SVD}} = 1 \times 10^{-15}$, which leads to $\Pr(\sigma_\Omega < \epsilon_{\mathrm{SVD}}) \approx 0.358$. Therefore, a continuous distribution leads to better spark performance than a discrete distribution.

B. MIP

Next, we consider the MIP of the sensing matrix. The parameters are set to $N = 64$ and $M = 16$. The frequency codes are uniformly distributed over the discrete set $\mathcal{D}_d$. The relative bandwidths are set to $B/f_c = \{0, 0.1, 0.5\}$, where $B/f_c = 0$ means that the assumption $\zeta_n \approx 1$ holds. We also simulate the continuous case with $d_n \sim \mathcal{D}_c$ and $\zeta_n \approx 1$. Curves are obtained with 10^6 Monte-Carlo trials. For each trial, we calculate the mutual
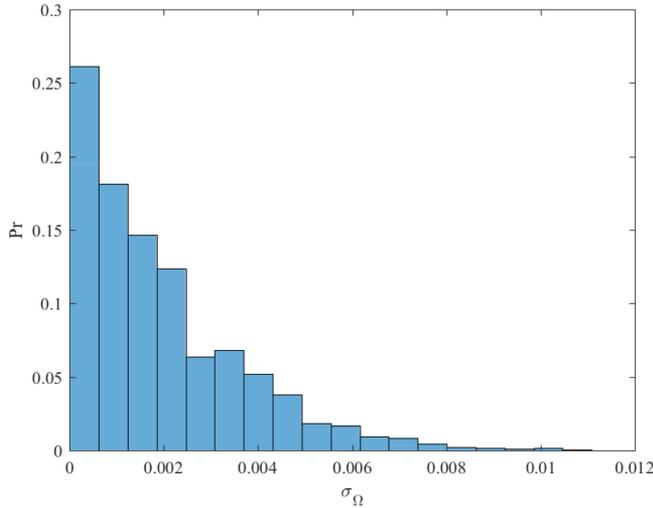
Fig. 3. Histogram of σ_Ω, the minimum σ_N among all sub-matrices of a sensing matrix $\boldsymbol{\Phi}$. The histogram is obtained using 2000 Monte-Carlo trials.

Fig. 5. Exact support set recovery probabilities using basis pursuit (P_1) and matched filter (18) in noiseless cases.

Fig. 4. The cumulative distribution functions of the mutual coherence μ obtained from 10^6 Monte-Carlo trials.

coherence μ of the sensing matrix and depict the corresponding cumulative distribution function. The theoretical bound in (34) is also displayed. The results are shown in Fig. 4. It can be seen that the theoretical upper bound (34) is tight under the assumption that the relative bandwidth is negligible (thus $\zeta_n \approx 1$ holds). When the relative bandwidth is large, the actual mutual coherence could exceed the predicted one, e.g., in the case that $B/f_c = 0.1$. However, the curve of $B/f_c = 0.5$ is under that of $B/f_c = 0.1$, which indicates that a larger relative bandwidth does not necessarily result in worse mutual incoherence.

C. Recovery Performance in Noiseless Cases

In Fig. 5, we consider the recovery performance in noiseless cases. In particular, we plot the probability of exact recovery using the basis pursuit algorithm (P_1) and matched filter (18), where exact recovery means the support set of the unknown vector $\boldsymbol{x}$ is exactly estimated. In the simulations, the pulse

number is $N = 64$, the number of frequencies is $M = 8$, and the amplitudes of scattering coefficients are all set to 1. The number of scatterers, K, is varied. The initial carrier frequency is $f_c = 10$ GHz and the bandwidth is $B = 64$ MHz. For each point on the curve, we perform 200 Monte-Carlo trials, where the frequency codes are randomly drawn obeying $U(\mathcal{D}_d)$, the support set of $\boldsymbol{x}$ is random, and the phases of non-zeros in $\boldsymbol{x}$ are i.i.d $U([0, 2\pi])$. We solve (P_1) using CVX [27], [28]. In both methods, we assume the number of scatterers, K, is known, and the support set is obtained as the indices of the largest K magnitudes in $\boldsymbol{x}$. The magnitudes are also compared with a threshold $\epsilon = 10^{-2}$. Those not exceeding the threshold are removed from the support set. From Fig. 5, it is seen that CS dramatically outperforms the traditional matched filter. When $K > 5$, the support set recovery probabilities using matched filter drops significantly, because some of the scatterers are masked by the sidelobes. When basis pursuit is applied, the region leading to exact support set recovery in FAR is fairly broad. However, the theoretical bound (36) is 1.5 with $\delta = 0.1$, and is quite pessimistic.

D. Recovery Performance in Noisy Cases

We next consider noisy cases, and choose the probability of successful recovery to evaluate the performance of different CS algorithms. A successful recovery is defined as exact recovery of the support set. In the simulations, $f_c = 10$ GHz, $B = 64$ MHz, $N = 64$, $M = 8$ and the number of scatterers $K = 3$. The scatterers have identical amplitudes of 1 with random phases. The noise $\boldsymbol{w}$ in (17) is assumed Gaussian white noise with a covariance matrix $\sigma^2 \boldsymbol{I}$, and σ^2 varies from -15 dB to 15 dB. Subspace pursuit [29] and Lasso are compared. In subspace pursuit, the number of scatterers K is assumed known a priori. The Lasso algorithm solves

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_1 \qquad (40)$$
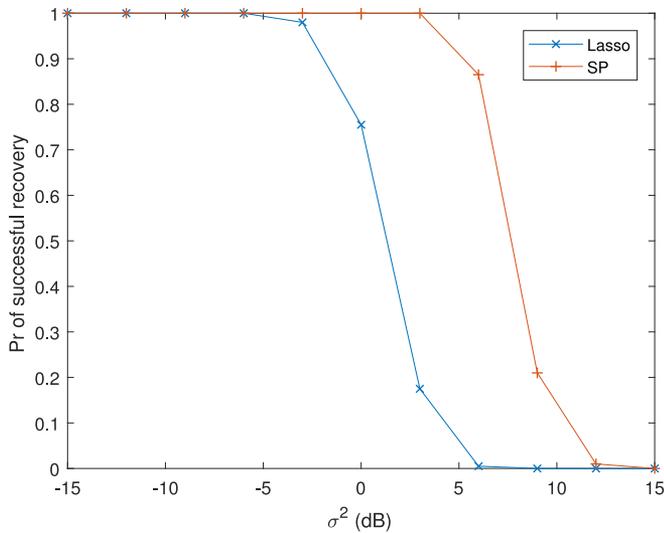
Fig. 6.   Probabilities of successful recovery using subspace pursuit and Lasso algorithms.

with $\lambda = 3\sigma^2$, and is implemented with CVX. When the magnitude of the estimate is larger than $\epsilon = 0.2$, the corresponding index is put into the estimated support set. The results are shown in Fig. 6 with 200 Monte-Carlo trials. Both algorithms have high successful recovery probabilities for an FAR in high SNR, $\sigma^2 \leq 0$ dB. Subspace pursuit outperforms Lasso with the genie-aided information on the cardinal number of the support set. We also note that the selection of the parameters λ and ϵ has a significant impact on the performance of Lasso.

E. Field Experiments

Next, we show field experiments from a true FAR prototype. We use separated antennas for transmitting and receiving, respectively, so that returns from short range objects are not eclipsed by the transmission. The radar works at an initial frequency of $f_c = 9$ GHz. Frequencies are varied pulse by pulse. There are $M^\star = 64$ frequencies with a minimum gap of 16 MHz, which results in a synthetic bandwidth of $B = 1024$ MHz. In each pulse, the carrier frequency $f_n = f_c + d_n B$ is randomly chosen. Specifically, $d_n \in \mathcal{D}_d = \{0, 1/M^\star, \ldots, (M^\star - 1)/M^\star\}$ and $d_n \sim U(\mathcal{D}_d)$. The HRR is $c/2B \approx 0.15$ m. We set $T_r = 0.2$ ms and the equivalent pulse duration $T_p = 31.25$ ns. The CRR is $cT_p/2 \approx 4.7$ m, and the number of HRR bins in a CRR bin is $M = \lceil T_p B \rceil = 32 < M^\star$, which satisfies the condition to eliminate ghost images [18], [19]. The number of pulses is $N = 512$. The moving target does not cross a CRR bin during the CPI, which requires $v < \frac{T_p c}{2NT_r} \approx 366.2$ m/s. For a slowly moving car, the velocity is lower than 10 m/s.

Field experiments are executed to evaluate performance. The target is a household car (see Fig. 7) with two corner reflectors and four small metal spheres upon its roof to enhance the SNR. When the radar is operated, the car moves in front of the radar at a nearly constant speed along the road, surrounded by static objects including a big stone, roadside trees and iron barriers. Returns from all scatterers located in $(0, cT_r/2]$, i.e. $(0, 30]$ km,

Fig. 7.   Field experiment scenario.
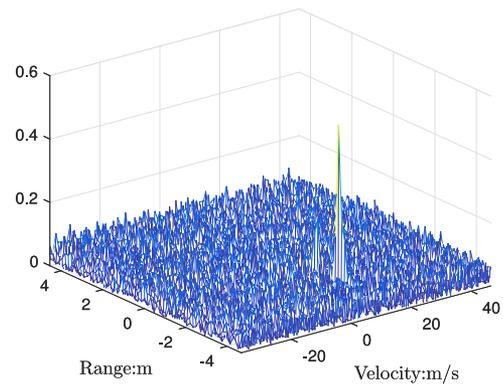
Fig. 8.   Field experiment result using a matched filter.
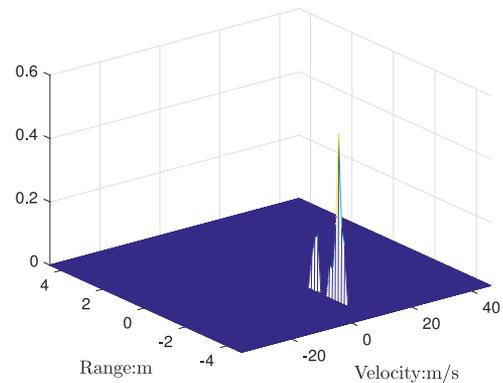
Fig. 9.   Field experiment result using OMP.

are collected. There are $\lfloor T_r/T_p \rfloor = 6400$ CRR bins. We perform static clutter canceling [30] over all CRR bins. Then in these CRR bins, we find the CRR bin which has the maximum amplitude of radar echoes, and infer that the car is located in that bin. With data in that CRR bin, we perform range-Doppler processing of the target. We apply the matched filter (18) and the OMP algorithm to jointly estimate the HRR profile and Doppler of the target. In OMP, the algorithm iterates 50 times, i.e., assuming $K = 50$. The results are shown in Fig. 8 and Fig. 9, respectively. To better demonstrate the sidelobe, we project the three dimensional images in Fig. 8 and Fig. 9 onto amplitude-range dimensions; see Fig. 10. In Fig. 10, only the maximum K amplitudes in $\hat{x}$ using matched filter are found and shown. In both methods, the velocity estimation is 6.9 m/s, and the span of the car is around 1.9 m. Comparing the recovery performance,
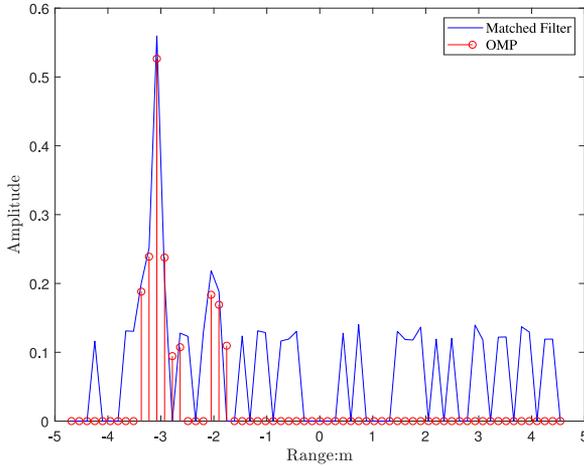
Fig. 10.   Field experiment results projected onto amplitude-range dimensions.

we see that the matched filter suffers from sidelobe pedestal while OMP demonstrates a clearer reconstruction.

VI. CONCLUSION

In this paper, sparse recovery for a frequency agile radar with random frequency codes is studied. We analyzed the spark and mutual incoherence properties of the radar sensing matrix, which guarantee reliable reconstruction of the targets. Using ℓ_0 minimization, FAR exactly recovers $K = \frac{N}{2}$ scatterers in noise-less cases almost surely, where N is the number of pulses. When we apply ℓ_1 minimization or greedy CS methods and there is noise, the number of scatterers that is guaranteed to be reliably reconstructed by FAR is on the order of $K = O(\sqrt{\frac{N}{\log M N}})$, where M is the number of HRR bins in a CRR bin. Numerical simulations and field experiments were executed to validate the theoretical results and also demonstrate the practical recovery performance of FAR.

APPENDIX A
PROOF OF THEOREM 4

To avoid confusion, in this appendix we will index the row of a matrix by ξ and the column by η. We need to prove that any N columns of $\boldsymbol{\Phi}$ are almost surely linearly independent. Following the form in (16), we first fix some N columns $\boldsymbol{\Phi}_{l_\eta + m_\eta N}$, with η running through $\{0, 1, \ldots, N-1\}$. These N columns constitute a new $N \times N$ matrix $\boldsymbol{A}$ whose elements are $[\boldsymbol{A}]_{\xi,\eta} = [\boldsymbol{\Phi}]_{\xi, l_\eta + m_\eta N}$. We need to show that $\boldsymbol{A}$ is almost surely invertible. For brevity, we set

$$z_\xi := \exp(j 2\pi d_\xi), \quad c_{\xi,\eta} := \exp(j 2\pi l_\eta \xi / N). \quad (41)$$

With the notation above, we may write $[\boldsymbol{A}]_{\xi,\eta} = c_{\xi,\eta} z_\xi^{m_\eta}$. The proof follows from the following three lemmas. The first one, being more abstract, is a strengthened version of the well-known fact that the set of zeros of a nonzero polynomial has measure zero.

Lemma 11: Let $P \in \mathbb{C}[z_1, \cdots, z_n]$ be a nonzero complex polynomial in n-variables. Denote by $\mathcal{N} \subset \mathbb{C}^n$ the set of zeros of P. Consider the n-torus $\mathbb{T}^n = \underbrace{S^1 \times \cdots \times S^1}_{n}$ with its obvious

embedding $\mathbb{T}^n \subset \mathbb{C}^n$. Let σ_n be the Haar measure on $\mathbb{T}^n$. We have

$$\sigma_n(\mathcal{N} \cap \mathbb{T}^n) = 0. \quad (42)$$

Proof: The left hand side of the above equality is well-defined, since $\mathcal{N}$ is closed and, consequently, $\mathcal{N} \cap \mathbb{T}^n$ is a closed set in $\mathbb{T}^n$. Note that σ_n coincides with the product measure $\underbrace{\sigma_1 \times \cdots \times \sigma_1}_{n}$. This enables us to prove the result by induction via Fubini's theorem.

For $n = 1$ the set $\mathcal{N}$ is discrete, thus the proposition is trivial. Suppose for $n = k$ the proposition is true. Let $\pi : \mathbb{C}^{k+1} \to \mathbb{C}$ be the projection onto the first component. We now make the natural identification $\mathbb{C}[z_1, \cdots, z_{k+1}] \approx (\mathbb{C}[z_2, \cdots, z_{k+1}])[z_1]$. With the result for $n = k$ in hand, we can find a σ_k-negligible set (a set of measure zero) $\mathcal{O} \subset \mathbb{T}^k$ such that for any $(z_2, \cdots, z_{k+1}) \in \mathbb{T}^k \setminus \mathcal{O}$, the polynomial P (viewed as a polynomial in z_1) has at least one nonzero coefficient, i.e. is a nonzero polynomial in z_1. Then by the result for $n = 1$, for $\mathbf{z}^{(k)} \in \mathbb{T}^k \setminus \mathcal{O}$, the measure of the slice

$$\sigma_1 \left(\pi \left(\mathcal{N} \cap \mathbb{T}^{k+1} \cap \left\{ (z_2, \cdots, z_{k+1}) = \mathbf{z}^{(k)} \right\} \right) \right) = 0. \quad (43)$$

We write this fact as $\sigma_1(\pi(\mathcal{N}|_{\mathbf{z}^{(k)}})) = 0$. Bearing in mind that $\sigma_{k+1} = \sigma_k \times \sigma_1$, we apply Fubini's theorem to obtain

$$
\begin{aligned}
\sigma_{k+1}(\mathcal{N} \cap \mathbb{T}^{k+1}) &= \int_{\mathbb{T}^{k+1}} 1_\mathcal{N} d\sigma_{k+1} \\
&= \int_{\mathbb{T}^k \setminus \mathcal{O}} d\sigma_k \int_{S^1} 1_\mathcal{N} \cdot \mu(dz_1) \\
&\quad + \int_{\mathcal{O}} d\sigma_k \int_{S^1} 1_\mathcal{N} \cdot \sigma_1(dz_1) \\
&= \int_{\mathbb{T}^k \setminus \mathcal{O}} \sigma_1(\pi(\mathcal{N}|_{\mathbf{z}^{(k)}})) \sigma_k(d\mathbf{z}^{(k)}) + 0 \\
&= 0, \quad (44)
\end{aligned}
$$

which completes the proof.   ∎

Lemma 12: Let $\{d_\xi\}$ be independent random variables with continuous distributions for $\xi = 0, \ldots, N-1$. Fix some complex numbers $c_{\xi,\eta}$, positive real constants A_ξ and nonnegative integers m_η, $\xi, \eta = 0, \ldots, N-1$. Let $z_\xi = \exp(j A_\xi d_\xi)$. Then the following two statements are equivalent:

i) The $N \times N$ random matrix $\boldsymbol{A}$ with elements

$$[\boldsymbol{A}]_{\xi,\eta} = c_{\xi,\eta} z_\xi^{m_\eta} \quad (45)$$

is almost surely invertible.

ii) There exists a vector $\mathbf{w} \in \mathbb{C}^N$ such that the $N \times N$ deterministic matrix $\boldsymbol{A}^{(\mathbf{w})}$ with elements

$$\left[\boldsymbol{A}^{(\mathbf{w})} \right]_{\xi,\eta} = c_{\xi,\eta} w_\xi^{m_\eta} \quad (46)$$

is invertible.

Proof: (i)⇒(ii) is obvious. We prove that (ii)⇒(i). Note that $\det \boldsymbol{A}$ is a polynomial P in N variables $z_0, \ldots, z_{N-1}$. By (ii), this polynomial is nonzero, since $P(w_1, \ldots, w_{N-1}) = \det \boldsymbol{A}^{(\mathbf{w})} \neq 0$. Let $\mathcal{N}$ be the set of zeros of P. Lemma 11 now implies that $\sigma_N(\mathcal{N} \cap \mathbb{T}^N) = 0$. On the other hand, the map $\phi : \mathbb{R}^N \to \mathbb{T}^N$ defined by $\phi(x_0, \ldots, x_{N-1}) =$

$(e^{j2\pi x_0}, \ldots, e^{j2\pi x_{N-1}})$ is obviously absolutely continuous. By the assumption that d_ξ are independent and absolutely continuous, the map $(d_0, \ldots, d_{N-1}) : \Omega \to \mathbb{R}^N$ is also absolutely continuous. Thus the probability of the event $(d_1, \ldots, d_{N-1})^{-1}\phi^{-1}(\mathcal{N} \cap \mathbb{T}^N)$ is 0, as desired. ∎

Lemma 13: Let l_η, m_η be nonnegative integers and $c_{\xi,\eta} = \exp(j2\pi l_\eta \xi / N)$. Furthermore, assume that the map $\eta \mapsto (\exp(j2\pi l_\eta / N), m_\eta)$ is injective when η takes value in $\{0, \ldots, N-1\}$. Then the statement (ii) in Lemma 12 is true.

Proof: Choose some real number b which is not a rational multiple of π. Take $w_\xi = \exp(jb\xi)$. Then $\boldsymbol{A}^{(\mathbf{w})}$ becomes a Vandermonde matrix

$$\left[\boldsymbol{A}^{(\mathbf{w})}\right]_{\xi,\eta} = e^{j\xi(2\pi l_\eta/N + bm_\eta)}. \tag{47}$$

By the determinant of a Vandermonde matrix, it suffices to prove that $\eta \mapsto \exp(j(2\pi l_\eta/N + bm_\eta))$ is injective. Suppose to the contrary that for some $\eta \neq \eta'$ we have $(2\pi l_\eta/N + bm_\eta) - (2\pi l_{\eta'}/N + bm_{\eta'}) = 2k\pi$, $k \in \mathbb{Z}$. Since b is not a rational multiple of π, this implies $m_\eta = m_{\eta'}$, henceforth $2\pi l_\eta/N - 2\pi l_{\eta'}/N = 2k\pi$. But then $(\exp(2\pi l_\eta/N), m_\eta) = (\exp(2\pi l_{\eta'}/N), m_{\eta'})$, contradicting the injectivity of $\eta \mapsto (\exp(j2\pi l_\eta/N), m_\eta)$. This proves that

$$\det \boldsymbol{A}^{(\mathbf{w})} = \prod_{0 \leq \eta < \eta' < N} \left(e^{j\frac{2\pi}{N}l_{\eta'}+jbm_{\eta'}} - e^{j\frac{2\pi}{N}l_\eta+jbm_\eta}\right) \neq 0. \tag{48}$$

In other words, our choice of $\mathbf{w} \in \mathbb{C}^N$ makes the matrix in (46) invertible. ∎

The map $\eta \mapsto (\exp(j2\pi l_\eta/N), m_\eta)$ is injective since (l_η, m_η) is pairwise distinct and $0 \leq l_\eta < N$. Then one may apply Lemma 13 and Lemma 12 successively and, by the consequence of Lemma 12, conclude the proof of Theorem 4.

We conclude with a remark on the robustness of our proof. The approximation we take here is $\zeta_\xi \approx 1$. Since ζ_ξ is absorbed in $c_{\xi,\eta}$, only the conclusion of Lemma 13 will be affected if $\zeta_\xi \neq 1$. However, from the proof of Lemma 13, for ζ_ξ sufficiently close to 1, its conclusion remains true, thus Theorem 4 still holds.

APPENDIX B
PROOF OF LEMMA 7

We divide the proof into two parts: in the first part, we use Lyapunov's condition to prove that the real and imaginary parts of χ_l have a joint Gaussian distribution asymptotically; in the second part, the expectation and variance are calculated. For conciseness, we omit the subscript l in χ_l. The parameters p and q belong to specific grids as stated in Subsection II-B, respectively. Note that $l \in \Xi$, which means $p \neq 0$. Also recall the assumption that random frequency codes $d_n \sim U(\mathcal{D}_d)$, and are independent from each other.

A. Asymptotic Distribution

First, consider the case $p \neq \pi$. Introduce a constant λ, and define a random variable

$$X_n := \mathrm{Re}(e^{jpMd_n+jqn}) + \lambda\mathrm{Im}(e^{jpMd_n+jqn})$$

$$= \cos(pMd_n + qn) + \lambda\sin(pMd_n + qn)$$
$$= T_\lambda \cos(pMd_n + qn + z_\lambda)$$
$$= T_\lambda \cos(pMd_n + \theta), \tag{49}$$

where $T_\lambda := \sqrt{1+\lambda^2}$, $\sin z_\lambda = -\frac{\lambda}{T_\lambda}$, $\cos z_\lambda = \frac{1}{T_\lambda}$ and $\theta := qn + z_\lambda$. Define

$$Y_N := \sum_{n=0}^{N-1} X_n. \tag{50}$$

Note that $Y_N = \chi$ when $\lambda = j$. Thus, it is equivalent to prove that for any real value λ, as $N \to \infty$, it holds that

$$\frac{Y_N - \mathrm{E}[Y_N]}{S_N} \sim \mathcal{N}(0, 1), \tag{51}$$

where S_N is the standard variance, obeying

$$S_N^2 = \mathrm{E}\left[\left(\sum_{n=0}^{N-1}(X_n - \mathrm{E}[X_n])\right)^2\right]$$
$$= \sum_{n=0}^{N-1} \mathrm{E}\left[(X_n - \mathrm{E}[X_n])^2\right], \tag{52}$$

where independence between the X_n is used. For $p \in (0, 2\pi)$ and $M > 1$, it holds that $S_N^2 > 0$.

According to Lyapunov's central limit theorem [31], (51) holds, if for some $\delta > 0$,

$$\lim_{N \to \infty} \frac{\sum_{n=0}^{N-1} \mathrm{E}\left[|X_n - \mathrm{E}[X_n]|^{2+\delta}\right]}{S_N^{2+\delta}} = 0. \tag{53}$$

We consider $\delta = 1$. In the following, we calculate $\mathrm{E}[|X_n - \mathrm{E}[X_n]|^3]$ and S_N^3 to verify that (53) holds assuming $d_n \sim U(\mathcal{D}_d)$.

To calculate S_N^3, derive

$$\mathrm{E}[X_n] = \frac{T_\lambda}{M}\sum_{m=0}^{M-1} \cos(pm + \theta)$$

$$= \frac{T_\lambda}{2M\sin\frac{p}{2}}\left(\sin\left(\frac{M-1}{2}p + \theta\right) - \sin\left(-\frac{p}{2} + \theta\right)\right)$$

$$= \frac{T_\lambda}{M\sin\frac{p}{2}}\sin\frac{Mp}{2}\cos\left(\frac{M-1}{2}p + \theta\right), \tag{54}$$

where we assume $p \neq 0$. In addition,

$$\mathrm{E}[X_n^2] = \frac{T_\lambda^2}{M}\sum_{m=0}^{M-1}\cos^2(pm + \theta)$$

$$= \frac{T_\lambda^2}{2M}\sum_{m=0}^{M-1}(\cos(2pm + 2\theta) + 1)$$

$$= \frac{T_\lambda^2}{2} + \frac{T_\lambda^2\sin((2M-1)p + 2\theta) - \sin(2\theta - p)}{4M\sin p}$$

$$= \frac{T_\lambda^2}{2} + \frac{T_\lambda^2\sin(Mp)\cos((M-1)p + 2\theta)}{2M\sin p}. \tag{55}$$

Therefore,

$$
\begin{aligned}
\mathrm{D}[X_n] &= \mathrm{E}[X_n^2] - (\mathrm{E}[X_n])^2 \\
&= \frac{T_\lambda^2}{2} + \frac{T_\lambda^2 \sin(Mp)\cos((M-1)p+2\theta)}{2M\sin p} \\
&\quad - \frac{T_\lambda^2 \sin^2 \frac{Mp}{2} \cos^2 (\frac{M-1}{2}p+\theta)}{M^2 \sin^2 \frac{p}{2}} \\
&= \frac{T_\lambda^2}{2} + \frac{T_\lambda^2 \sin(Mp)\cos((M-1)p+2\theta)}{2M\sin p} \\
&\quad - \frac{T_\lambda^2 \sin^2 \frac{Mp}{2}\cos((M-1)p+2\theta)}{2M^2\sin^2\frac{p}{2}} - \frac{T_\lambda^2 \sin^2 \frac{Mp}{2}}{2M^2\sin^2\frac{p}{2}}.
\end{aligned}
\tag{56}
$$

Applying $p \in \{\frac{2\pi}{M}, \frac{2\pi\cdot2}{M}, \dots, \frac{2\pi(M-1)}{M}\}$, we have $\sin(Mp) = \sin^2\frac{Mp}{2} = 0$ and $\sin p \neq 0$, $\sin\frac{p}{2} \neq 0$. Then

$$
S_N^2 = \sum_{n=0}^{N-1} \mathrm{D}[X_n] = \frac{NT_\lambda^2}{2}.
\tag{57}
$$

We conclude that $S_N^2 = O(N)$, and $S_N^3 = O(N^{\frac{3}{2}})$.

To calculate the numerator in (53), note that

$$
|X_n| < C_1,
\tag{58}
$$

$$
|\mathrm{E}[X_n]| < C_2,
\tag{59}
$$

where C_1 and C_2 are positive constants not related to N. Then,

$$
\mathrm{E}\left[|X_n - \mathrm{E}[X_n]|^3\right] \leq \mathrm{E}\left[(|X_n| + |\mathrm{E}[X_n]|)^3\right] \leq (C_1 + C_2)^3.
\tag{60}
$$

Combing $S_N^3 = O(N^{\frac{3}{2}})$ and (60), we have

$$
\lim_{N\to\infty} \frac{\sum_{n=0}^{N-1}\mathrm{E}\left[|X_n - \mathrm{E}[X_n]|^3\right]}{S_N^3} \leq \lim_{N\to\infty}\frac{N(C_1+C_2)^3}{O(N^{\frac{3}{2}})} = 0.
\tag{61}
$$

Thus, (53) holds.

When $p = \pi$, $e^{jpMd_n+jqn} = e^{j\pi Md_n+jqn} = (-1)^{Md_n}e^{jqn}$. Define a random variable

$$
X_n' := T_\lambda(-1)^{Md_n}\cos\theta.
\tag{62}
$$

Following similar steps as above, we find that (53) still hold for X_n'.

According to Lyapunov's central limit theorem, as $N \to \infty$, $\mathrm{Re}(\chi)$ and $\mathrm{Im}(\chi)$ have an asymptotic joint Gaussian distribution.

B. Expectation and Variance

In this subsection, we calculate the expectations and variances of $\mathrm{Re}(\chi)$ and $\mathrm{Im}(\chi)$. Denote the variances of the real and imaginary parts and the correlation coefficient as σ_1^2, σ_2^2, and σ_{12}, respectively, i.e.

$$
\begin{bmatrix} \mathrm{Re}(\chi) \\ \mathrm{Im}(\chi) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathrm{Re}(\mathrm{E}[\chi]) \\ \mathrm{Im}(\mathrm{E}[\chi]) \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \right).
\tag{63}
$$

We start by analyzing the expectation of the complex valued χ,

$$
\mathrm{E}[\chi] = \mathrm{E}\left[\frac{1}{N}\sum_{n=0}^{N-1} e^{jpMd_n+jqn} \right].
\tag{64}
$$

Since $\mathrm{Pr}(d_n = \frac{m}{M}) = \frac{1}{M}$, it holds that

$$
\mathrm{E}[\chi] = \sum_{m=0}^{M-1}\frac{1}{MN}\sum_{n=0}^{N-1} e^{jpm+jqn}.
\tag{65}
$$

Exchanging the order of summations,

$$
\begin{aligned}
\mathrm{E}[\chi] &= \sum_{n=0}^{N-1}\frac{1}{MN}e^{jqn}\sum_{m=0}^{M-1}e^{jpm} \\
&= \frac{1}{MN}\frac{1-e^{jpM}}{1-e^{jp}}\sum_{n=0}^{N-1}e^{jqn} \\
&= 0,
\end{aligned}
\tag{66}
$$

where the last equality holds because $p \in \{\frac{2\pi}{M}, \frac{2\pi\cdot2}{M}, \dots, \frac{2\pi(M-1)}{M}\}$, which implies $e^{jpM} = 1$ while $e^{jp} \neq 1$ and hence $\frac{1-e^{jpM}}{1-e^{jp}} = 0$.

Next, we calculate the variances σ_1^2, σ_2^2 and σ_{12}. According to [32], it holds that

$$
\mathrm{E}[\chi^2] = \sigma_1^2 - \sigma_2^2 + 2j\sigma_{12},
\tag{67}
$$

$$
\mathrm{E}[|\chi|^2] = \sigma_1^2 + \sigma_2^2.
\tag{68}
$$

The left hand side of (67) satisfies

$$
\begin{aligned}
\mathrm{E}[\chi^2] &= \mathrm{E}\left[\frac{1}{N^2}\sum_{n=0}^{N-1}e^{jpMd_n+jqn}\sum_{k=0}^{N-1}e^{jpMd_k+jqk} \right] \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{k=0,k\neq n}^{N-1}e^{jq(n+k)}\mathrm{E}\left[e^{jpM(d_n+d_k)}\right] \\
&\quad + \frac{1}{N^2}\sum_{n=0}^{N-1}e^{j2qn}\mathrm{E}\left[e^{j2pMd_n}\right].
\end{aligned}
\tag{69}
$$

Applying $\mathbb{P}(d_n = \frac{m}{M}) = \frac{1}{M}$ and independence between d_n,

$$
\begin{aligned}
\mathrm{E}[\chi^2] &= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{k=0,k\neq n}^{N-1}e^{jq(n+k)}\sum_{m_1=0}^{M-1}\sum_{m_2=0}^{M-1}\frac{1}{M^2}e^{jp(m_1+m_2)} \\
&\quad + \frac{1}{N^2}\sum_{n=0}^{N-1}e^{j2qn}\sum_{m=0}^{M-1}\frac{1}{M}e^{jp2m} \\
&= \frac{1}{N^2M^2}\frac{(1-e^{jpM})^2}{(1-e^{jp})^2}\sum_{n=0}^{N-1}\sum_{k=0,k\neq n}^{N-1}e^{jq(n+k)}
\end{aligned}
$$

$$+ \frac{1}{N^2 M} \frac{1 - e^{j2pM}}{1 - e^{j2p}} \sum_{n=0}^{N-1} e^{j2qn}$$

$$= \frac{1}{N^2 M^2} \frac{(1 - e^{jpM})^2}{(1 - e^{jp})^2} \left(\frac{(1 - e^{jNq})^2}{(1 - e^{jq})^2} - \frac{1 - e^{j2qN}}{1 - e^{j2q}} \right)$$

$$+ \frac{1}{N^2 M} \frac{1 - e^{j2pM}}{1 - e^{j2p}} \frac{1 - e^{j2qN}}{1 - e^{j2q}}. \tag{70}$$

According to the assumption $p \in \left\{ \frac{2\pi}{M}, \frac{2\pi \cdot 2}{M}, \ldots, \frac{2\pi (M-1)}{M} \right\}$, we have $\frac{1 - e^{jpM}}{1 - e^{jp}} = 0$ and thus the first term in (70) equals zero. Note that

$$\lim_{x \to \pi} \frac{1 - e^{j2Mx}}{1 - e^{j2x}} = M. \tag{71}$$

We conclude that

$$\mathrm{E}\left[\chi^2\right] = \begin{cases} \dfrac{1}{N}, & \text{if } p = q = \pi, \\ 0, & \text{otherwise.} \end{cases} \tag{72}$$

Similarly, as for the left side of (68), we have

$$\mathrm{E}\left[|\chi|^2\right] = \mathrm{E}\left[\frac{1}{N^2} \sum_{n=0}^{N-1} e^{jpM d_n + jqn} \sum_{k=0}^{N-1} e^{-jpM d_k - jqk} \right]$$

$$= \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{k=0,k\neq n}^{N-1} e^{jq(n-k)} \mathrm{E}\left[e^{jpM(d_n - d_k)} \right] + \frac{\sum_{n=0}^{N-1} \mathrm{E}[1]}{N^2}$$

$$= \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{k=0,k\neq n}^{N-1} e^{jq(n-k)} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \frac{e^{jp(m_1 - m_2)}}{M^2} + \frac{1}{N}$$

$$= \frac{(1 - e^{jpM})(1 - e^{-jpM})}{N^2 M^2 (1 - e^{jp})(1 - e^{-jp})} \sum_{n=0}^{N-1} \sum_{k=0,k\neq n}^{N-1} e^{jq(n-k)} + \frac{1}{N}$$

$$= \frac{1}{N^2 M^2} \frac{|1 - e^{jpM}|^2}{|1 - e^{jp}|^2} \left(\frac{|1 - e^{jNq}|^2}{|1 - e^{jq}|^2} - N \right) + \frac{1}{N}$$

$$= \frac{1}{N}. \tag{73}$$

Substituting $\mathrm{E}\left[\chi^2\right] = 0$ and $\mathrm{E}\left[|\chi|^2\right] = \frac{1}{N}$ into (67) and (68), respectively, one finds that $\sigma_1^2 = \sigma_2^2 = \frac{1}{2N}$ and $\sigma_{12} = 0$. As for the case $p = q = \pi$, $\mathrm{E}\left[\chi^2\right] = \frac{1}{N}$ and $\mathrm{E}\left[|\chi|^2\right] = \frac{1}{N}$, it holds that $\sigma_1^2 = \frac{1}{N}$ and $\sigma_2^2 = \sigma_{12} = 0$.

ACKNOWLEDGMENT

REFERENCES

[1] T. Huang and Y. Liu, "Compressed sensing for a frequency agile radar with performance guarantees," in *Proc. IEEE China Summit Int. Conf. Signal Inf. Process.*, Jul. 2015, pp. 1057–1061.

[2] Z. Liu and S. Zhang, "Velocity estimation for hopped-frequency radar," *Signal Process.* (in Chinese), vol. 16, no. 2, pp. 97–100, 2000.

[3] S. R. J. Axelsson, "Analysis of random step frequency radar and comparison with experiments," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 890–904, Apr. 2007.

[4] Y. Liu, H. Meng, G. Li, and X. Wang, "Range-velocity estimation of multiple targets in randomised stepped-frequency radar," *Electron. Lett.*, vol. 44, no. 17, pp. 1032–1034, 2008.

[5] T. Huang, Y. Liu, G. Li, and X. Wang, "Randomized stepped frequency ISAR imaging," in *Proc. IEEE Radar Conf.*, May 2012, pp. 0553–0557.

[6] T. Huang, Y. Liu, H. Meng, and X. Wang, "Cognitive random stepped frequency radar with sparse recovery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 858–870, Apr. 2014.

[7] L. Zhang, Z.-J. Qiao, M. Xing, Y. Li, and Z. Bao, "High-resolution ISAR imaging with sparse stepped-frequency waveforms," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4630–4651, Nov. 2011.

[8] J. Yang, J. Thompson, X. Huang, T. Jin, and Z. Zhou, "Random-frequency SAR imaging based on compressed sensing," *IEEE Trans. Geosc. Remote Sens.*, vol. 51, no. 2, pp. 983–994, Feb. 2013.

[9] Z. Liu, X. Wei, and X. Li, "Decoupled ISAR imaging using RSFW based on twice compressed sensing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 3195–3211, Oct. 2014.

[10] D. Cohen, K. V. Mishra, and Y. C. Eldar, "Spectrum sharing radar: Coexistence via Xampling," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 3, pp. 1279–1296, Jun. 2018.

[11] J. E. Luminati, T. B. Hale, M. A. Temple, M. J. Havrilla, and M. E. Oxley, "Doppler aliasing artifact filtering in SAR imagery using randomised stepped-frequency waveforms," *Electron. Lett.*, vol. 40, no. 22, pp. 1447–1448, 2004.

[12] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[13] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[14] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[15] T. Huang, Y. Liu, H. Meng, and X. Wang, "Adaptive matching pursuit with constrained total least squares," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, 2012, Art. no. 76.

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

[17] F. Krahmer and H. Rauhut, "Structured random measurements in signal processing," *GAMM-Mitteilungen*, vol. 37, no. 2, pp. 217–238, 2014.

[18] Y. Liu, H. Meng, H. Zhang, and X. Wang, "Eliminating ghost images in high-range resolution profiles for stepped-frequency train of linear frequency modulation pulses," *IET Radar, Sonar Navigat.*, vol. 3, no. 5, pp. 512–520, 2009.

[19] Y. Liu, T. Huang, H. Meng, and X. Wang, "Fundamental limits of HRR profiling and velocity compensation for stepped-frequency waveforms," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4490–4504, Sep. 2014.

[20] O. Bar-Ilan and Y. C. Eldar, "Sub-Nyquist radar via Doppler focusing," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1796–1811, Apr. 2014.

[21] F. Xi, S. Chen, and Z. Liu, "Quadrature compressive sampling for radar signals," *IEEE Trans. Signal Process.*, vol. 62, no. 11, pp. 2787–2802, Jun. 2014.

[22] Y. Hu, Y. Liu, H. Meng, and X. Wang, "Extended range profiling in stepped-frequency radar with sparse recovery," in *Proc. IEEE RadarCon*, May 2011, pp. 1046–1049.

[23] R. Cohen and Y. C. Eldar, "Sparse Doppler sensing based on nested arrays," Aug. 2018. [Online]. Available: http://arxiv.org/abs/1710.00542

[24] J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1341–1344, Jun. 2004.

[25] M. Rossi, A. M. Haimovich, and Y. C. Eldar, "Spatial compressive sensing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 419–430, Jan. 2014.

[26] D. Dorsch and H. Rauhut, "Refined analysis of sparse MIMO radar," *J. Fourier Anal. Appl.*, vol. 23, no. 3, pp. 485–529, 2017.

[27] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," Mar. 2014, [Online]. Available: http://cvxr.com/cvx

[28] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control* (Lecture Notes in Control and Information Sciences), V. Blondel, S. Boyd, and H. Kimura, Eds. Berlin, Germany: Springer-Verlag, 2008, pp. 95–110. [Online]. Available: http://stanford.edu/ boyd/graph_dcp.html

[29] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[30] S. R. J. Axelsson, "Suppression of noise floor and dominant reflectors in random noise radar," in *Proc. Int. Radar Symp.*, May 2006, pp. 1–4.

[31] A. A. Borovkov, *Probability Theory* (Universitext). London, U.K.: Springer, 2013.

[32] Y. Lo, "A mathematical theory of antenna arrays with randomly spaced elements," *IEEE Trans. Antennas Propag.*, vol. 12, no. 3, pp. 257–268, May 1964.

Tianyao Huang received the B.S. degree in telecommunication engineering from Harbin Institute of Technology, Harbin, China, in 2009, and the Ph.D. degree in electronics engineering from Tsinghua University, Beijing, China, in 2014.

From 2014 to 2017, he was a Radar Researcher with Aviation Industry Corporation of China. Since July 2017, he has been an Assistant Professor with the Intelligent Sensing Lab, Department of Electronic Engineering, Tsinghua University. His current research interests include radar signal processing, radar waveform design, and compressive sensing.

Yimin Liu (M'12) received the B.S. and Ph.D. degrees (both with honors) in electronics engineering from Tsinghua University, Beijing, China, in 2004 and 2009, respectively.

From 2004, he was with the Intelligence Sensing Lab, Department of Electronic Engineering, Tsinghua University. He is currently an Associate Professor with Tsinghua University, where his field of activity is study on new concept radar and other microwave sensing technologies. His current research interests include radar theory, statistic signal processing, compressive sensing, and their applications in radar, spectrum sensing, and intelligent transportation systems.

Xingyu Xu is an undergraduate student with the Department of Electronic Engineering, Tsinghua University, Beijing, China. His research interests include topological and statistical data analysis.

Yonina C. Eldar (S'98–M'02–SM'07–F'12) received the B.Sc. degree in physics in 1995 and the B.Sc. degree in electrical engineering in 1996 both from Tel-Aviv University, Tel-Aviv, Israel, and the Ph.D. degree in electrical engineering and computer science in 2002 from Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. She is currently a Professor with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel, where she holds the Edwards Chair in Engineering. She is also a Visiting Professor with the Research Laboratory of Electronics, MIT, an Adjunct Professor with the Duke University, and was a Visiting Professor with Stanford University, Stanford, CA, USA. She is the author of the book *Sampling Theory: Beyond Bandlimited Systems* and the co-author of the books *Compressed Sensing* and *Convex Optimization Methods in Signal Processing and Communications*, all published by Cambridge University Press. Her research interests are in the broad areas of statistical signal processing, sampling theory and compressed sensing, optimization methods, and their applications to biology and optics.

Dr. Eldar is a member of the Israel Academy of Sciences and Humanities (elected 2017) and a EURASIP Fellow. She was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. She was a member of the Young Israel Academy of Science and Humanities and the Israel Committee for Higher Education. She is the Editor-in-Chief of *Foundations and Trends in Signal Processing*, a member of the IEEE Sensor Array and Multichannel Technical Committee, and serves on several other IEEE committees. In the past, she was a Signal Processing Society Distinguished Lecturer, member of the IEEE Signal Processing Theory and Methods and Bio Imaging Signal Processing technical committees, and was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the *EURASIP Journal of Signal Processing*, the *SIAM Journal on Matrix Analysis and Applications*, and the *SIAM Journal on Imaging Sciences*. She was the Co-Chair and Technical Co-Chair of several international conferences and workshops. She was the recipient of many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award (2013), the IEEE/AESS Fred Nathanson Memorial Radar Award (2014), and the IEEE Kiyo Tomiyasu Award (2016), the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, the Wolf Foundation Krill Prize for Excellence in Scientific Research, the Henry Taub Prize for Excellence in Research (twice), the Hershel Rich Innovation Award (three times), the Award for Women with Distinguished Contributions, the Andre and Bella Meyer Lectureship, the Career Development Chair at the Technion, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technions Award for Excellence in Teaching (two times). She was also the recipient of several best paper awards and best demo awards together with her research students and colleagues including the SIAM outstanding Paper Prize, the UFFC Outstanding Paper Award, the Signal Processing Society Best Paper Award, and the IET Circuits, Devices and Systems Premium Award. She was selected as one of the 50 most influential women in Israel.

Xiqin Wang received the B.S. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1991 and 1996, respectively. He was a Visiting Scholar and Assistant Research Engineer of PATH/ITS, UC Berkeley, Berkeley, CA, USA, from 2000 to 2003. He is currently a Professor with the Department of Electronic Engineering, Tsinghua University, where he is also the Vice-President. His current research interests include radar and communications signal processing, image processing, compressed sensing and cognitive signal processing, and electronic system design. He is also interested in structure of knowledge and curricula reforming in electronic engineering.