COMPRESSIVE RADAR WITH OFF-GRID AND EXTENDED TARGETS

ALBERT FANNJIANG AND HSIAO-CHIEH TSENG

ABSTRACT. Compressed sensing (CS) schemes are proposed for monostatic as well as synthetic aperture radar (SAR) imaging with chirps. In particular, a simple method is developed to improve performance with off-grid targets. Tomographic formulation of spotlight SAR is analyzed by CS methods with several bases and under various bandwidth constraints. Performance guarantees are established via coherence bound and the restricted isometry property. CS analysis provides a fresh and clear perspective on how to optimize temporal and angular samplings for spotlight SAR.

1. INTRODUCTION

Advances in compressed sensing (CS) and radar processing have provided tremendous impetus to each other. On the one hand, the two CS themes of sparse reconstruction and low-coherence, pseudo-randomized data acquisition are longstanding concepts in radar processing. On the other hand, CS contributes provable performance guarantees for sparse recovery algorithms and informs refinement of these algorithms. These and other important issues relevant to CS radar are thoroughly reviewed in [13,24] (see also the references therein).

Target sparsity, a main theme in CS, arises naturally in radar processing. According to the geometrical theory of diffraction [21], the scattering response of a target at radio frequencies can often be approximated as a sum of responses from individual reflectors. These scattering centers provide a concise, yet physically relevant, description of the object [18]. A spiky reconstruction of reflectivity may thus be highly valuable for automatic target recognition. More generally radar images are compressible by means of either parametric models of physical scattering behaviors or transform coding [24].

In the present work, we consider point as well as extended targets. For point targets, a main drawback, however, of the standard CS framework often neglected by previous works is the reliance on a underlying grid. In reality, the dominant scattering centers can not be assumed to be positioned exactly at the computational grid points. Indeed, CS-algorithms often break down when the targets are located between grid points [13,16]. In order to reduce gridding error, the grid has to be refined, giving rise to higher coherence of the measurement matrix which is detrimental to standard CS algorithms.

Can CS approach be extended to the case of arbitrarily located targets? Several approaches have been proposed to address this critical question [4, 11, 16, 17] and all of them focus on algorithmic improvements. In this paper we propose a simple, alternative approach, which focuses on improving the accuracy of the measurement matrix, and establish performance guarantee for the refined system (Sections 3 and 4).

We then turn to the spotlight mode of synthetic aperture radar (SAR). We consider the tomographic framework of spotlight SAR as formulated in [22] for different classes of targets, including point targets, localized extended targets and distributed extended targets. These different targets require different sparsifying bases which in turn demand different sampling

schemes. We propose several sampling schemes and establish performance guarantee for them (Section 5). In particular, we examine the realistic case of high-frequency, narrow-band signals and derive an error bound. In each case, we show that CS can achieve the standard resolution limit with sparse sampling. Finally we present numerical experiments demonstrating our results (Section 6) and draw conclusion (Section 7).

2. MONOSTATIC SIGNAL MODEL

Let us first formulate the signal model for the case of a monostatic radar with narrow fractional bandwidth and colocated transmitter and receiver antennas. A complex waveform f with the carrier frequency ω_0 is transmitted. Let r and v denote the range and the radial velocity, respectively. We parameterize the complex scene by the reflectivity function $\rho(\tau, u)$ where the delay $\tau = 2r/c_0$ is the round-trip propagation time and $u = 2v\omega_0/c_0$ is the Doppler shift. Under the far-field and narrow-band approximations [7], the scattered signal is given by

(1)
$$y(t) \ = \iint x(\tau, u) f(t - \tau) e^{-2\pi i u t} \, du \, d\tau + w(t)$$

where

$$x(\tau, u) = \rho(\tau, u) e^{-\pi i u \tau}$$

and $w(t)$ represents the circular white complex Gaussian baseband noise.

(1) For immobile targets, $\rho(\tau, u) = \rho(\tau)\delta(u)$, eq. (1) becomes

(2)
$$y(t) = \int_{-\infty}^{\infty} \rho(\tau) f(t - \tau) \, d\tau + w(t).$$

For immobile point targets located at $\{\tau_k\}$, eq. (2) becomes

(3)
$$y(t) = \sum_k \rho_k f(t - \tau_k) + w(t).$$

(2) For moving point targets located at $\{\tau_k\}$ with Doppler shifts $\{u_l\}$, we have

$$\rho(\tau, u) = \sum_{k,l} \rho_{k,l} \delta(\tau - \tau_k)\delta(u - u_l),$$

and

(4)
$$y(t) = \sum_{k,l} x_{k,l} f(t - \tau_k) e^{-2\pi i u_l t} + w(t).$$

For the transmitted signal, let I_T be the indicator function of duration $[0, T]$ of transmission. By far the most commonly used waveform is the linear FM chirp

(5)
$$f_{\mathrm{LC}}(t) = \exp \left[2\pi i \left(\frac{\alpha_1}{2} t^2 + \omega_0 t \right) \right] I_T(t)$$

owing to the simplicity in implementation. The bandwidth of linear chirp can be defined as $B = \alpha_1 T$.

Motivated by the low coherence of the Alltop sequences [1], we will also consider the quadratic FM chirp:

(6)
$$f_{\mathrm{QC}}(t) = \exp \left[2\pi i \left(\frac{\alpha_2}{3} t^3 + \frac{\alpha_1}{2} t^2 + \omega_0 t \right) \right] I_T(t)$$

2

whose bandwidth is given by $B = \alpha_2 T^2 + \alpha_1 T$. The Alltop sequences have been analyzed in the stylized CS-radar work in [19].

3. Immobile targets

3.1. **On-grid targets.** With linear chirp and immobile point targets, the signal model is given by

$$y(t_j) = \sum_k \rho_k \exp\left[2\pi i \left(\frac{\alpha_1}{2}(t_j - \tau_k)^2 + \omega_0(t_j - \tau_k)\right)\right]$$
$$= f_{\mathrm{LC}}(t_j) \sum_k \rho_k f_{\mathrm{LC}}(-\tau_k) \exp\left[-2\pi i \alpha_1 \tau_k t_j\right].$$

Suppose first that the targets are located exactly on the grid points. Let $\tau_k = k\Delta\tau$ where $\Delta\tau$ is the grid size and $k = 1, \ldots, n$, $n > m$. Let

$$(7) \qquad\qquad \alpha_1 = \frac{Q}{T\Delta\tau}, \quad Q \in \mathbb{N}$$

and

$$\widehat{t}_j = t_j/T \in [0, 1].$$

Then with

$$(8) \qquad Y_j \;=\; \frac{y(t_j)}{f_{\mathrm{LC}}(t_j)}$$

$$(9) \qquad X_k \;=\; \rho_k f_{\mathrm{LC}}(-\tau_k)$$

$$(10) \qquad F_{j,k} \;=\; \exp\left[-2\pi i \alpha_1 \tau_k t_j\right] = \exp\left[-2\pi i B k \Delta\tau \widehat{t}_j\right] = \exp\left[-2\pi i Q k \widehat{t}_j\right]$$

we can write the signal model as the linear system $Y = FX + E$ where $E = [w(t_j)] \in \mathbb{C}^m$ represents the measurement errors. The bandwidth is given by $B = \alpha_1 T = Q/\Delta\tau$.

A main thrust of CS is the performance guarantee for the basis pursuit denoising (BPDN):

$$\widehat{X} = \arg\min \|Z\|_1, \quad \|\mathbf{F}Z - Y\|_2 \le \epsilon$$

under the assumption of the restricted isometry property (RIP):

$$a(1 - \delta_k)\|Z\|_2 \le \|\mathbf{F}Z\|_2 \le a(1 + \delta_k)\|Z\|_2$$

for some constant $a > 0$ and all k-sparse Z where δ_k is the k-th order restricted isometry constant. More precisely, we have the following statement [3, 25].

Proposition 1. *Let $\widehat{t}_j \in [0, 1]$, $j = 1, 2, \ldots, m$ be independent uniform random variables. If*

$$(11) \qquad\qquad \frac{m}{\ln m} \ge Cs \ln^2 s \ln n \ln\frac{1}{\beta}, \quad \beta \in (0, 1)$$

for some universal constant C and sparsity level s, then the random Fourier measurements $[\frac{1}{\sqrt{m}} e^{2\pi i k \widehat{t}_j}]$, $k = 1, \ldots, n$, satisfy the RIP with $\delta_{2s} < \sqrt{2} - 1$ and the BPDN solution $\widehat{X}$ satisfies

$$\left\|\widehat{X} - X\right\|_2 \le C_0 \frac{1}{\sqrt{s}} \left\|X^{(s)} - X\right\|_1 + C_1 \|E\|_2 \;, \quad \left\|\widehat{X} - X\right\|_1 \le C_0 \left\|X^{(s)} - X\right\|_1 + C_1 \|E\|_2$$

for some constants $,C_1$, with probability at least $1 - \beta$. Here $X^{(s)}$ is the best s-sparse approximation of X.

3

According to the above result, the (normalized) sampling times should be chosen randomly and uniformly in $[0, 1]$ and the number of time samples m needs to be sufficiently large (11). This result is valid as long as the point targets are located exactly on the grid points which is a stringent and often unrealistic assumption.

3.2. **Off-grid targets.** In practice, the time delays $\{\tau_k\}$ may not sit exactly on the grid points $\{k\Delta\tau\}$. The mismatch between the actual signal and the signal model creates the gridding error which leads to poor performance in CS reconstruction [8, 16, 17].

To deal with this problem we can make the model more accurate to compensate for the gridding error. Let $\tau_k = k\Delta\tau + \xi_k$ with $|\xi_k| < \frac{\Delta\tau}{2}$. Then

$$y(t_j) = \sum_k \rho_k f_{\mathrm{LC}}(t_j - \tau_k)$$

(12)
$$= f_{\mathrm{LC}}(t_j) \sum_k \rho_k f_{\mathrm{LC}}(-\tau_k) e^{-2\pi i \alpha_1 t_j \tau_k}$$

$$= f_{\mathrm{LC}}(t_j) \sum_k \rho_k f_{\mathrm{LC}}(-k\Delta\tau - \xi_k) e^{-2\pi i \alpha_1 t_j k\Delta\tau} e^{-2\pi i \alpha_1 \xi_k t_j}.$$

Suppose that $B \max |\xi_k|$ is small. Then we can write

$$e^{-2\pi i \alpha_1 \xi_k t_j} = e^{-2\pi i B \xi_k \widehat{t}_j} = e^{-\pi i B \xi_k} \left(1 - 2\pi i B \xi_k (\widehat{t}_j - 1/2) + \mathcal{O}\left(B^2 \max_k |\xi_k|^2\right)\right).$$

Let $Q = B\Delta\tau \in \mathbb{N}$ be the resolution-time-bandwidth product (RTBP). With

(13)
$$Y_j = \frac{y(t_j)}{f_{\mathrm{LC}}(t_j)}$$

(14)
$$X_k = \rho_k f_{\mathrm{LC}}(-k\Delta\tau - \xi_k) e^{-\pi i B \xi_k}, \quad X'_k = -2\pi i Q \widehat{\xi}_k X_k, \quad \widehat{\xi}_k = \frac{\xi_k}{\Delta\tau}$$

(15)
$$F_{j,k} = \exp\left[-2\pi i Q k \widehat{t}_j\right], \quad G_{j,k} = (\widehat{t}_j - 1/2) F_{jk}$$

the linear system takes the form

$$Y_j = \sum_k \left(F_{j,k} X_k + G_{j,k} X'_k\right) + E_j$$

or equivalently

(16)
$$Y = \begin{bmatrix} \mathbf{F} & \mathbf{G} \end{bmatrix} \begin{bmatrix} X \\ X' \end{bmatrix} + E.$$

After X and X' are solved from the system, we can estimate $\{\xi_k\}$ and $\{\rho_k\}$, respectively, by

$$\xi_k = \frac{i\Delta\tau X'_k}{2\pi Q X_k}$$

and

$$\rho_k = \frac{X_k e^{\pi i B \xi_k}}{f_{\mathrm{LC}}(-k\Delta\tau - \xi_k)}.$$

It is generally difficult to establish RIP for matrices other than random partial Fourier matrices and random matrices of independently and identically distributed (i.i.d.) entries such as $[\mathbf{F} \ \mathbf{G}]$. An alternative approach is via the notion of mutual coherence μ. The mutual

coherence of a matrix $\mathbf{A}$ is defined by the maximum normalized inner product between columns of $\mathbf{A}$:

$$(17) \qquad \mu(\mathbf{A}) = \max_{i \neq j} \frac{|A_i^* A_j|}{\|A_i\|_2 \|A_j\|_2}.$$

The lower the level of mutual coherence the better performance is with the system.

The following lemma states a coherence bound for the system (16).

Lemma 1. *Suppose* $2n < \delta \exp\left[K^2/2\right]$. *Then the mutual coherence* μ *of the combined sensing matrix* $\mathbf{A} = [\mathbf{F}\ \mathbf{G}]$ *satisfies*

$$(18) \qquad \mu \leq C \left[\frac{\sqrt{2}K}{\sqrt{m}} + \frac{1}{2\pi Q} \right]$$

for some universal constant C, *with probability greater than* $(1-\delta)^2 - 4e^{-m/18}$.

This lemma says that to reduce the mutual coherence of the sensing matrix one should increase the number of data and the RTBP. The proof of the lemma is given in appendix A.

We then obtain an error bound from Lemma 1 in conjunction with the following result (Theorem 3.1 of [12]).

Proposition 2. *Suppose the data is noisy and the* ℓ^2*-norm of the noise is less than* ϵ. *If the sparsity of* X *satisfies*

$$(19) \qquad s < \frac{1}{4}(1 + \frac{1}{\mu})$$

then the BPDN solution $\widehat{X}$ *satisfies the error bound*

$$(20) \qquad \left\| \widehat{X} - X \right\|_2^2 \leq \frac{4\epsilon^2}{1 - (4s-1)\mu}.$$

Observe that $\mathrm{supp}X = \mathrm{supp}X'$. This property can be utilized in the greedy pursuit (such as orthogonal matching pursuit) as follows. A common stage for any greedy pursuit is to choose the index corresponding to the columns of the maximum coherence with the residual vector. Since X, X' have the same sparse structure, one may utilize a priori information: choose both k-th columns from $\mathbf{F}$ and $\mathbf{G}$, and test the size of the projected vector from Y on the span of the two columns.

The sparsity constraint (19) can be relaxed if the targets are randomly distributed in distance (time delay) as stated in the following proposition [5].

Proposition 3. *Suppose the data noises are additive Gaussian noise of variance* σ^2. *Assume that the targets are uniformly randomly distributed and* s*-sparse such that*

$$(21) \qquad s \leq \frac{c_1 nm}{\|\mathbf{A}\|_2^2 \log n}.$$

Assume also

$$(22) \qquad \mu \leq \frac{c_2}{\log n}.$$

5

If

(23)
$$\min_{k \in S} |X_k| > 8\sigma\sqrt{2 \log n}$$

then the solution $\widehat{X}$ of the LASSO

(24)
$$\min_{Z} \; \lambda\sigma \|Z\|_1 + \frac{1}{2}\|Y - \mathbf{A}Z\|_2^2, \quad \lambda = 2\sqrt{2 \log n}$$

obeys $\mathrm{supp}(\widehat{X}) = \mathrm{supp}(X)$ with probability at least $1 - \mathcal{O}(n^{-1})$.

Once the target support is exactly recovered the target strengths can be calculated by a simple least-squares step to obtain a nearly optimal reconstruction.

On the one hand, (21) demands little on the mutual coherence. On the other hand, in order for (21) to improve over (19) we need a sufficiently strong spectral norm bound. We prove the following spectral norm bound in appendix B.

Lemma 2. *The sensing matrix in (15) satisfies the spectral norm bound*

$$\|\mathbf{A}\|_2^2 \le 2n$$

with probability greater than

$$\left(1 - \frac{C(m-1)}{n}\right)^{m(m-1)}$$

where C is an absolute constant.

Lemmas 1, 2 and Proposition 3 imply that LASSO can achieve nearly optimal reconstruction of randomly distributed targets whose number is on the order of the data number m, modulo a logarithmic factor. This is stated in the following theorem.

Theorem 1. *Suppose the data noises are additive Gaussian noise of variance σ^2. Assume that the targets are uniformly randomly distributed and s-sparse such that*

$$s \le \frac{c_1 m}{2 \log n}.$$

Suppose that

$$\frac{\sqrt{2}K}{\sqrt{m}} + \frac{1}{2\pi Q} \le \frac{c_2}{C \log n}.$$

If (23) holds, then the LASSO (24) recovers the target support exactly with probability at least $1 - \mathcal{O}(n^{-1})$.

4. Moving targets

Next we discuss the case of moving targets and include Doppler shift in the signal model.

4.1. **On-grid targets.** With $f = f_{\mathrm{LC}}(t)$ we obtain the signal model from (4)

$$y(t_j) = \sum_{k,l} x_{k,l} \exp\left[2\pi\mathrm{i}\left(\frac{\alpha_1}{2}(t_j - \tau_k)^2 + \omega_0(t_j - \tau_k)\right)\right] e^{-2\pi\mathrm{i}u_l t_j}$$

$$= f_{\mathrm{LC}}(t_j) \sum_{k,l} x_{k,l} f_{\mathrm{LC}}(-\tau_k) \exp\left[-2\pi\mathrm{i}\left(\alpha_1 \tau_k + u_l\right) t_j\right].$$

Let $\tau_k = k\Delta\tau$, $u_l = (l - n/2)\Delta u$ with $k, l = 1, \ldots, n$. With

$$\alpha_1 = \frac{\Delta u}{n\Delta\tau}.$$

we set up the following grid $\{\gamma_p\}$

$$\gamma_p = \tau_k + u_l/\alpha_1 = (k + n(l - n/2))\Delta\tau = (p + n - n^2/2)\Delta\tau$$

where $p = k + n(l - 1) \in \{1, \ldots, n^2\}$. By choosing

$$T = \frac{nQ}{\Delta u}, \quad Q \in \mathbb{N}$$

we can write the linear system $Y = \mathbf{F}X + E$ in terms of

$$(25) \qquad\qquad Y_j \;\; = \;\; \frac{y(t_j)}{f_{\mathrm{LC}}(t_j)} e^{\pi\mathrm{i}(2n - n^2)Q\widehat{t}_j}$$

$$(26) \qquad\qquad X_p \;\; = \;\; x_{k,l} f_{\mathrm{LC}}(-\tau_k)$$

$$(27) \qquad\qquad F_{j,p} \;\; = \;\; \exp\left[-2\pi\mathrm{i}Qp\widehat{t}_j\right] \in \mathbb{C}^{m \times n^2}$$

where, with the uniform random variables t_j over $[0, T]$, $j = 1, 2, \ldots, m$, the sensing matrix $\mathbf{F}$ becomes the random partial Fourier matrix of size $m \times n^2$ with which BPDN is guaranteed to perform according to Proposition 1. Note with the above choice of parameters the bandwidth $B = Q/\Delta\tau$ is the same as in (7).

4.2. **Off-grid targets.** Suppose the actual time delays $\{\tau_k\}$ and Doppler shifts $\{u_l\}$ are not exactly on the grid: $\tau_k = k\Delta\tau + \xi_k$, $u_l = (l - n/2)\Delta u + \eta_l$. Eq. (4) with linear chirp now becomes

$$y(t_j) = \sum_{k,l} x_{k,l} f_{\mathrm{LC}}(t_j - \tau_k) e^{-2\pi\mathrm{i}u_l t_j}$$

$$= f_{\mathrm{LC}}(t_j) \sum_{k,l} x_{k,l} f_{\mathrm{LC}}(-\tau_k) e^{-2\pi\mathrm{i}(\alpha_1 \tau_k + u_l) t_j},$$

$$= f_{\mathrm{LC}}(t_j) \sum_{k,l} x_{k,l} f_{\mathrm{LC}}(-k\Delta\tau - \xi_k) e^{-2\pi\mathrm{i}(\alpha_1 k\Delta\tau + (l - n/2)\Delta u) t_j} e^{-2\pi\mathrm{i}(\alpha_1 \xi_k + \eta_l) t_j}.$$

With $\Delta\tau$, Δu, and γ_p as above, the approximation

$$e^{-2\pi\mathrm{i}(\alpha_1 \xi_k + \eta_l) t_j} = e^{-2\pi\mathrm{i}B\zeta_p \widehat{t}_j} \approx e^{-\pi\mathrm{i}B\zeta_p}\left(1 - 2\pi\mathrm{i}B\zeta_p(\widehat{t}_j - 1/2)\right), \quad \zeta_p = \xi_k + \eta_l/\alpha_1$$

7

for small $B|\zeta_p|$ now leads to (16) with

$$(28) \qquad Y_j \;=\; \frac{y(t_j)}{f_{\mathrm{LC}}(t_j)} e^{\pi i (2n - n^2) Q \widehat{t}_j}$$

$$(29) \qquad X_p \;=\; x_{k,l} f_{\mathrm{LC}}(-k\Delta\tau - \xi_k) e^{-\pi i B \zeta_p} \;, \quad X'_p = -2\pi i B \zeta_p X_p$$

$$(30) \qquad F_{j,p} \;=\; \exp\left[-2\pi i Q p \widehat{t}_j\right] \;, \quad G_{j,p} = (\widehat{t}_j - 1/2) F_{jp}.$$

The sensing matrix $\mathbf{A} = [\mathbf{F}\ \mathbf{G}]$ obeys the same coherence bound in Lemma 1 and thus has the performance guarantee given in Proposition 2. Once X and X' are reconstructed, we can compute $\{\zeta_p\}$ by

$$\zeta_p = \frac{i X'_p}{2\pi B X_p}.$$

Even though we are unable to determine $\{\xi_k\}$ and $\{\eta_l\}$ separately, the sparsity patterns of X and X determine the time delays and Doppler shifts up to the precision of $\Delta\tau$ and Δu, respectively. Moreover, the magnitude of the reflectivity can be recovered since $|\rho_{k,l}| = |x_{k,l}| = |X_p|$.

4.3. Quadratic chirp.
For purpose of comparison, let us the signal model (4) with $f = f_{\mathrm{QC}}$. The received signal takes the form

$$y(t_j) \;=\; \sum_{k,l=1}^{n} x_{k,l} \exp\left[2\pi i \left(\frac{\alpha_2}{3}(t_j - \tau_k)^3 + \frac{\alpha_1}{2}(t_j - \tau_k)^2 + \omega_0 (t_j - \tau_k)\right)\right] e^{-2\pi i u_l t_j}$$

$$(31) \qquad =\; f_{\mathrm{QC}}(t_j) \sum_{k,l=1}^{n} x_{k,l} f_{\mathrm{QC}}(-\tau_k) \exp\left[2\pi i \left(\alpha_2 t_j(\tau_k^2 - t_j \tau_k) - t_j(\alpha_1 \tau_k - u_l)\right)\right].$$

Instead of randomly time sampling, we use uniform sampling $t_j = j\Delta t$, $j = 1, 2, \ldots, m$ with $\Delta t = T/m$. With the combined index $p = k + n(l-1)$ as before the linear system becomes

$$(32) \qquad Y_j \;=\; \frac{y(t_j)}{f_{\mathrm{QC}}(t_j)}$$

$$(33) \qquad X_p \;=\; x_{k,l} f_{\mathrm{QC}}(-\tau_k)$$

$$(34) \qquad F_{jp} \;=\; \exp\left[2\pi i \left(\alpha_2 t_j(-t_j \tau_k + \tau_k^2) - t_j(\alpha_1 \tau_k - u_l)\right)\right]$$

Lemma 3. *With the on-grid targets $\tau_k = k\Delta\tau, u_l = l\Delta u$, the uniform sampling $t_j = j\Delta t, j = 1, 2, \ldots, m$, as well as the parameters*

$$(35) \qquad \alpha_2 \;=\; \frac{m}{T^2 \Delta\tau} \quad \text{(or its integer multiple)}$$

$$(36) \qquad \alpha_1 \;=\; \frac{1}{T\Delta\tau} \quad \text{(or its integer multiple)}$$

$$(37) \qquad \Delta u \;=\; \frac{1}{T} \quad \text{(or its integer multiple)}$$

$$(38) \qquad T \;=\; m\Delta\tau$$

$$m \;=\; n$$

8

the mutual coherence of $\mathbf{F} \in \mathbb{C}^{n \times n^2}$ *in (34) becomes*

$$(39) \qquad F_{j,p} = \exp\left[-2\pi i \left(k \frac{j^2}{n} + (-k^2 + k + l)\frac{j}{n}\right)\right], \quad p = k + n(l-1)$$

and satisfies the bound

$$(40) \qquad \mu(\mathbf{F}) \leq \frac{1}{\sqrt{n}}$$

for any prime number $n > 2$.

The proof of Lemma 3 is given in appendix C. Lemma 3 in conjunction with Proposition 2 then implies that up to $\mathcal{O}(\sqrt{n})$ targets can be recovered stably by BPDN with QC. As we will see in the numerical examples below, QC with random sampling and uniform sampling have essentially the same numerical performance even though we do not have the performance guarantee for QC with random sampling (Fig. 4 and 5).

Note that with the choice of parameters in Lemma 3 the bandwidth becomes $B = (n + 1)/\Delta\tau$. The coherence bounds for LC (Lemma 1) and QC (Lemma 3) are then comparable with $Q = n + 1$. And indeed, their numerical performances are similar to each other when their bandwidth are same (Fig. 4 and 5) .

For off-grid targets we apply the same linearization technique with the parameters (35)-(38) and obtain the following linear system:

$$(41) \qquad Y_j = \frac{y(t_j)}{f_{\mathrm{QC}}(t_j)}$$

$$(42) \qquad X_p = x_{k,l} f_{\mathrm{QC}}(-k\Delta\tau) e^{\pi i(2k\widehat{\xi}_k + \widehat{\xi}_k^2)} e^{\pi i m \widehat{\xi}_k/2} e^{-2\pi i m \widehat{\xi}_k \widehat{t}_j} e^{-\pi i \widehat{\xi}_k} e^{-\pi i \widehat{\eta}_l}$$

$$(43) \qquad X'_p = -2\pi i \widehat{\xi}_k X_p, \quad X''_p = -2\pi i \widehat{\eta}_l X_p, \quad \widehat{\xi}_k = \frac{\xi_k}{\Delta\tau}, \quad \widehat{\eta}_l = \frac{\eta_l}{\Delta u}$$

$$(44) \qquad F_{jp} = \exp\left[2\pi i \left(\widehat{t}_j(k^2 - m\widehat{t}_j k) + \widehat{t}_j(l - k)\right)\right]$$

$$(45) \qquad G_{jp} = (\widehat{t}_j - 1/2)(-2k + m(\widehat{t}_j - 1/2) + 1)F_{jp}$$

$$(46) \qquad H_{jp} = (\widehat{t}_j - 1/2)F_{jp}$$

Because of the complexity of the system we are unable to prove the incoherence property for (44)-(46).

5. Compressive spotlight SAR

In this section, we consider the spotlight synthetic aperture radar (SAR) for a stationary scene, represented by the reflectivity $\rho(\mathbf{r})$ of spatial variables $\mathbf{r}$ only. For simplicity of the presentation, we focus on the case of two dimensions $\mathbf{r} = (r_1, r_2)$. The adaption to three dimensions is straightforward.

In standard radar processing, the received signal, upon receive, is typically deramped by mixing the echo with the reference transmitted chirp [20]. Under the start-stop approximation and a far-field assumption the deramp processing produces samples of the Fourier transform of the Radon projection, orthogonal to the radar look direction, of the scene reflectivity multiplied by a quadratic phase term. Furthermore, if the time-bandwidth product

$TB = \alpha_1 T^2$ is significantly larger than the total number n of resolution cells, the quadratic phase term can be neglected and the deramped signal can be written simply as [22]

$$(47) \qquad y(\nu(t), \theta) = \mathcal{F}[\rho](\nu(t) \cos \theta, \nu(t) \sin \theta), \quad \nu(t) = \frac{2}{c_0}(\omega_0 + \alpha_1(t - \tau_0))$$

where $\mathcal{F}$ stands for the 2-d Fourier transform, θ is the look angle and τ_0 the round-trip travel time to the scene center. For a sufficiently small scene, $\nu(t)$ is effectively restricted to

$$(48) \qquad \nu \in \frac{2}{c_0}[\omega_0, \omega_0 + \alpha_1 T].$$

We will first discuss the ideal case of full spatial frequency band

$$(49) \qquad \nu \in [0, \nu_*]$$

and consider the partial bandwidth case in Section 5.3. We will return to the narrow band case (48) with

$$\nu_0 = 2\omega_0/c_0 \gg 2\alpha_1 T/c_0$$

in Section 5.4.

5.1. Off-grid point targets.
Let the computation domain be the finite square lattice

$$(50) \qquad \mathcal{L} = \left\{ \ell(p_1, p_2) : p_1, p_2 = 1, ..., \sqrt{n} \right\}.$$

The total number of cells n is a perfect square. For the off-grid targets represented by

$$\rho(\mathbf{r}) = \sum_{\mathbf{p} \in \mathbb{Z}^2} \rho_{\mathbf{p}} \delta(\mathbf{r} - \ell\mathbf{p} - \ell\mathbf{h_p}), \quad \mathbf{h_p} = (h_{1\mathbf{p}}, h_{2\mathbf{p}}), \ |h_{1\mathbf{p}}|, |h_{2\mathbf{p}}| < 1/2$$

the signal model (47) becomes

$$y(\nu, \theta) = \sum_{\mathbf{p} \in \mathbb{Z}^2} \rho_{\mathbf{p}} e^{-2\pi i \ell \nu \widehat{\mathbf{d}} \cdot (\mathbf{p} + \mathbf{h_p})}$$

where $\widehat{\mathbf{d}} = (\cos \theta, \sin \theta)$ denotes the direction of look. Following the same perturbation technique

$$e^{-2\pi i \ell \nu \widehat{\mathbf{d}} \cdot (\mathbf{p} + \mathbf{h_p})} = e^{-2\pi i \ell \nu \widehat{\mathbf{d}} \cdot \mathbf{p}}(1 - 2\pi i \ell \nu \widehat{\mathbf{d}} \cdot \mathbf{h_p}) + \mathcal{O}(|\ell \nu \widehat{\mathbf{d}} \cdot \mathbf{h_p}|^2)$$

we consider the signal model

$$(51) \qquad y(\nu, \theta) = \sum_{\mathbf{p} \in \mathbb{Z}^2} \rho_{\mathbf{p}} e^{-2\pi i \ell \nu \widehat{\mathbf{d}} \cdot \mathbf{p}}(1 - 2\pi i \ell \nu \widehat{\mathbf{d}} \cdot \mathbf{h_p}) + w(\nu, \theta)$$

where w is the error.

Let $(a_k, b_k), k = 1, \ldots, m$, be i.i.d. uniform random variables on $[-1/2, 1/2]^2$. Set ν_k, θ_k such that

$$(52) \qquad \ell \nu_k \cos \theta_k = a_k, \quad \ell \nu_k \sin \theta_k = b_k$$

and let $Y_k = y(\nu_k, \theta_k)$. For (52) to have a solution in (49) it is necessary that $\ell \nu_* \geq 1/2$ or equivalently $\ell \geq 1/(2\nu_*)$. In the small scene limit $\nu_* = \omega_0 + \alpha_1 T$, the resolution limit

$$(53) \qquad \ell \geq \frac{c_0}{4(\omega_0 + \alpha_1 T)}$$

is half of the standard resolution limit for passive radars owing to the two-way travel of the active radar signal.

10

Define the primary and secondary target vectors by

$$X_l = \rho_{\mathbf{p}}, \quad X_l' = -2\pi i \ell \mathbf{h}_{\mathbf{p}} X_l, \quad l = (p_2 - 1)\sqrt{n} + p_1$$

where $(X_l) \in \mathbb{C}^n, (X_l') \in \mathbb{C}^{2n}$. The approximate signal model takes the form (16) with

$$F_k = e^{-2\pi i \ell \nu_k \widehat{\mathbf{d}}_k \cdot \mathbf{p}}, \quad G_k = e^{-2\pi i \ell \nu_k \widehat{\mathbf{d}}_k \cdot \mathbf{p}} \nu_k \widehat{\mathbf{d}}_k$$

which is the two-dimensional version of (30).

The coherence bound given in Lemma 1 holds true with a slight adjustment of the constants. Thus this system has the performance guarantee stated in Proposition 2.

5.2. **Localized target.** Instead of point targets, we now consider extended targets that are sparse in the pixel basis. We pixelate the scene with n pixels of size ℓ. The centers of the pixels are identified as the finite square lattice $\mathcal{L}$. Let $\square_{\mathbf{p}}$ denote the $\mathbf{p}$-th pixel, $\forall \mathbf{p} \in \mathbb{N}^2$.

Suppose the scene is represented in the pixel basis as

$$\rho = \sum_{\mathbf{p} \in \mathbb{N}^2} I_{\mathbf{p}} \rho_{\mathbf{p}}$$

where $I_{\mathbf{p}}$ is the indicator function of $\square_{\mathbf{p}}$. Then (47) becomes

$$(54) \qquad y(\nu, \theta) = \ell^2 \sum_{\mathbf{p}} \rho_{\mathbf{p}} e^{2\pi i \ell \nu \widehat{\mathbf{d}} \cdot \mathbf{p}} \frac{\sin(\pi \ell \nu \cos \theta)}{\pi \ell \nu \cos \theta} \frac{\sin(\pi \ell \nu \sin \theta)}{\pi \ell \nu \sin \theta}, \quad \mathbf{p} \in \mathbb{N}^2.$$

We shall reconstruct $\{\rho_j\}$ from eq. (54) with the measurement data $y(\nu_k, \theta_k), k = 1, \ldots, m$ where ν_k is the spatial frequency corresponding to the look angle θ_k.

Let ν_k, θ_k be determined as in (52) with i.i.d. uniform random variables $(a_k, b_k), k = 1, \ldots, m$, on $[-1/2, 1/2]^2$ and

$$Y_k = \frac{y(\nu_k, \theta_k)}{\ell^2 \operatorname{sinc} a_k \cdot \operatorname{sinc} b_k}.$$

Then we can rewrite (54) as the standard CS problem with a random partial Fourier matrix of two dimensions

$$Y_k = \sum_j \rho_j e^{2\pi i (a_k p_1 + b_k p_2)}, \quad j = (p_2 - 1)\sqrt{n} + p_1$$

which has the same form as the one for on-grid targets and thus the CS performance guarantee given in Proposition 1 if the target is sufficiently sparse in the pixel basis.

The performance guarantee can be extended to the case of piecewise constant targets if the ℓ_1 minimization in BPDN is replaced by the total-variation (TV) minimization. A version of the performance guarantee for TV-minimization is given in [15].

5.3. **Distributed target.** Finally we consider the case of distributed targets which do not have a sparse representation in the pixel basis. Although the Fourier basis may seem at first a natural choice for representing such targets it is not suitable for our purpose since the scene of a sparse Fourier representation will yield a sparse set of data which will be missed with high probability by any sparse sampling scheme.

Instead, we represent the reflectivity function in the Littlewood-Paley basis with the mother wavelet

$$(55) \qquad \psi(\mathbf{r}) = (\pi^2 r_1 r_2)^{-1} (\sin(2\pi r_1) - \sin(\pi r_1)) \cdot (\sin(2\pi r_2) - \sin(\pi r_2))$$

which has the following Fourier transform

(56)
$$\widehat{\psi}(\xi, \eta) = \begin{cases} 1, & \frac{1}{2} \leq |\xi|, |\eta| \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

The following set of functions

(57)
$$\psi_{\mathbf{p},\mathbf{q}}(\mathbf{r}) = 2^{-(p_1+p_2)/2}\psi(2^{-\mathbf{P}}\mathbf{r} - \mathbf{q}), \quad \mathbf{p}, \mathbf{q} \in \mathbb{Z}^2$$

with the notation

$$2^{-\mathbf{P}}\mathbf{r} = (2^{-p_1}r_1, 2^{-p_2}r_2)$$

forms an orthonormal basis in $L^2(\mathbb{R}^2)$ [10]. In the Littlewood-Paley basis we have the expansion

(58)
$$\rho(r_1, r_2) = \sum_{\mathbf{p},\mathbf{q}\in\mathbb{Z}^2} \rho_{\mathbf{p},\mathbf{q}}\psi_{\mathbf{p},\mathbf{q}}(r_1, r_2)$$

for any square integrable function.

Let $\widehat{\mathbf{d}}_k = (\cos\theta_k, \sin\theta_k), k = 1, \ldots, m$ be the direction of look and let ν_k be the spatial frequency corresponding to the look angle θ_k. Taking the Fourier transform we obtain from (58) that

(59)
$$y(\nu_k, \theta_k) = 2\pi \sum_{\mathbf{p},\mathbf{q}\in\mathbb{Z}^2} 2^{(p_1+p_2)/2}\rho_{\mathbf{p},\mathbf{q}}e^{-2\pi i\nu_k 2^{\mathbf{P}}\widehat{\mathbf{d}}_k\cdot\mathbf{q}}\widehat{\psi}(\nu_k 2^{\mathbf{P}}\widehat{\mathbf{d}}_k), \quad k = 1, \ldots, n.$$

Now we describe how to select $(\theta_k, \nu_k), k = 1, \ldots, m$. First we define some notation. Let

$$l = \sum_{\mathbf{j}=(-p_*,-p_*)}^{\mathbf{p}-(1,1)} (2n_\mathbf{j} + 1)^2 + (q_1 + n_\mathbf{p})(2n_\mathbf{p} + 1) + (q_2 + n_\mathbf{p} + 1), \quad |\mathbf{q}|_\infty \leq n_\mathbf{p}, |\mathbf{p}|_\infty \leq p_*,$$

$$k = \sum_{\mathbf{j}=(-p_*,-p_*)}^{\mathbf{p}'-(1,1)} (2m_\mathbf{j} + 1)^2 + (q_1' + m_{\mathbf{p}'})(2m_{\mathbf{p}'} + 1) + (q_2' + m_{\mathbf{p}'} + 1), \quad |\mathbf{q}'|_\infty \leq m_{\mathbf{p}'}, |\mathbf{p}'|_\infty \leq p_*$$

be the column and row indexes, respectively, of the sensing matrix (see below) for some $m_\mathbf{p}, n_\mathbf{p}$ which are the numbers of row and columns, respectively of, the block $\mathbf{p}$ (see below) and $p_* \in \mathbb{N}$ which determines the maximum number, $2p_*+1$, of dyadic scales. It is important to keep in mind the relationship between k, l and $(\mathbf{p}', \mathbf{q}'), (\mathbf{p}, \mathbf{q})$ in order to follow the scheme described below.

The main idea is to select $(\theta_k, \nu_k), k = 1, \ldots, m$ such that the sensing matrix becomes block-diagonal with each block being a random partial Fourier matrix in *two* dimensions which has the CS performance guarantee given in Proposition 1. To this end, $\nu_k\widehat{\mathbf{d}}_k$ should be randomly and uniformly distributed on $[-1/2, 1/2]^2$. This can be realized by the following scheme which is a modification of the sampling scheme presented in [14].

For fixed $\mathbf{p}' = (p_1', p_2') \in \mathbb{Z}^2$, let (a_k, b_k) be independent, uniform random variables on $[-1/2, 1/2]^2$ and define

$$\widetilde{a}_k = \begin{cases} 1/2 + a_k, & a_k \in [0, 1/2] \\ -1/2 + a_k, & a_k \in [-1/2, 0] \end{cases}$$

$$\widetilde{b}_k = \begin{cases} 1/2 + b_k, & b_k \in [0, 1/2] \\ -1/2 + b_k, & b_k \in [-1/2, 0] \end{cases}.$$

Clearly,

(60) $$\widetilde{a}_k, \widetilde{b}_k \in [-1, -1/2] \cup [1/2, 1].$$

Let ν_k, θ_k such that

(61) $$\nu_k 2^{p'_1} \cos \theta_k = \widetilde{a}_k$$

(62) $$\nu_k 2^{p'_2} \sin \theta_k = \widetilde{b}_k.$$

The constraint (60) then implies

(63) $$\nu_k \geq \max\left(2^{-p'_1-1}, 2^{-p'_2-1}\right)$$

meaning that the spatial frequency must be greater than half of the inverse of the resolved scale. Altogether, (60)-(62) implies that for the Littlewood-Paley scheme to work for any specific scale, we need about 50% bandwidth instead of the full bandwidth such as given in (49).

Define the sensing matrix $\mathbf{A} = [A_{k,l}]$ with

(64) $$A_{k,l} = \frac{1}{2n_{\mathbf{p}} + 1} e^{-\mathrm{sgn}(a_k)\pi i q_1} e^{-\mathrm{sgn}(b_k)\pi i q_2} \widehat{\psi}(\nu_k 2^{\mathbf{P}} \widehat{\mathbf{d}}_k) e^{-2\pi i \nu_k 2^{\mathbf{P}} \widehat{\mathbf{d}}_k \cdot \mathbf{q}}.$$

By (56), (61) and (62) it is clear that $A_{k,l}$ are zero if $\mathbf{p} \neq \mathbf{p}'$. Indeed, we have

$$\widehat{\psi}(\nu_k 2^{\mathbf{P}} \widehat{\mathbf{d}}_k) e^{-2\pi i \nu_k 2^{\mathbf{P}} \widehat{\mathbf{d}}_k \cdot \mathbf{q}} = \widehat{\psi}(2^{\mathbf{P}-\mathbf{P}'}(a_k, b_k)) e^{2\pi i 2^{\mathbf{P}-\mathbf{P}'}(a_k, b_k) \cdot \mathbf{q}}$$

which vanishes if $\mathbf{p}' \neq \mathbf{p}$ thanks to the support constraint of $\widehat{\psi}$. Consequently the sensing matrix is the block-diagonal matrix with each block (indexed by $\mathbf{p}$) in the form of random Fourier matrix

(65) $$A_{k,l} = \frac{1}{2m_{\mathbf{p}} + 1} e^{2\pi i (q_1 a_k + q_2 b_k)}$$

where $\mathrm{sgn}(a_k)$ and $\mathrm{sgn}(b_k)$ are, respectively, the signs of a_k and b_k. The random phase term $e^{\mathrm{sgn}(\xi_k)\pi i q_1} e^{\mathrm{sgn}(\eta_k)\pi i q_2}$ can be further absorbed in the definition of data and plays no role in the performance analysis. The rest of the sensing elements form a random partial Fourier matrix.

Let $X = (X_l) \in \mathbb{C}^n$ be the target vector with

$$X_l = 2\pi (2m_{\mathbf{p}} + 1) 2^{(p_1+p_2)/2} \rho_{\mathbf{p},\mathbf{q}} e^{\mathrm{sgn}(a_k)\pi i q_1} e^{\mathrm{sgn}(b_k)\pi i q_2}$$

and

$$n = \sum_{j_1=-p_*}^{p_*} \sum_{j_2=-p_*}^{p_*} (2n_{\mathbf{j}} + 1)^2.$$

The block-diagonal structure of $\mathbf{A}$ implies that the target structures of different dyadic scales are decoupled and can be determined separately by CS techniques.

Under the assumptions of Proposition 1 we obtain the following theorem.

Theorem 2. *Suppose that for a given* $\mathbf{p}, |\mathbf{p}| \leq p_*$, *(11) is satisfied for* $n_{\mathbf{p}}, m_{\mathbf{p}}$ *and the target sparsity* $s_{\mathbf{p}}$ *on the scale* $2^{\mathbf{P}}$. *Suppose*

(66) $$\nu_k \geq 2^{\max(-p_1-1, -p_2-1)}$$

and that the data are contaminated by noise of ℓ^2-*norm* $\epsilon_{\mathbf{p}}$.

13

Let a_l, b_l be independently and uniformly distributed on $[-1/2, 1/2]$. Let the look directions and the spatial frequencies be determined by (61)-(62). Then the BPDN solution $\widehat{X}_\mathbf{p}$ satisfies

$$(67) \qquad \|\widehat{X}_\mathbf{p} - X_\mathbf{p}\|_2 \leq C_1 s_\mathbf{p}^{-1/2} \|X_\mathbf{p} - X_\mathbf{p}^{(s_\mathbf{p})}\|_1 + C_2 \epsilon_\mathbf{p}$$

with probability at least $1 - \beta$ where C_1 and C_2 are absolute constants.

Remark 1. *In the small scene limit, the spatial frequency range (48) and (66) imply the resolution limit*

$$2^\mathbf{P} \geq \frac{c_0}{4(\omega_0 + \alpha_1 T)}$$

which is the same as (53).

5.4. **High frequency/narrow band limit.** The preceding sampling scheme for SAR may be a bit cumbersome and unpractical. A far more practical sampling scheme is to independently select $\theta_k, k = 1, \ldots, m_1$ independently from $[0, 2\pi]$ and $\nu_l, l = 1, \ldots, m_2$ from (48). We also want to obey the spatial frequency constraint (48) with $\alpha_1 T \ll \omega_0$. In particular we assume that look angles θ_k are i.i.d. with the probability density function $f(\theta)$. In what follows we shall consider the high frequency limit and let

$$(68) \qquad \nu_0 \ell \gg 1, \quad \nu_0 = 2\omega_0/c_0$$

and $\alpha_1 T$ be fixed. We will consider on-grid targets only. We prove in Appendix C the following coherence bound.

Lemma 4. *Suppose*

$$(69) \qquad n \leq \frac{\delta}{8} e^{K^2/2}, \quad \delta, K > 0.$$

Then the sensing matrix satisfies the coherence bound

$$(70) \qquad \mu < \bar{\mu} + \frac{\sqrt{2}K}{\sqrt{m_1}}$$

with probability greater than $(1 - \delta)^2$ where in general χ^{i} (resp. χ^{s}) satisfies the bound

$$(71) \qquad \bar{\mu} \leq c_\gamma (1 + \nu_0 \ell)^{-1/2} \|f\|_{\gamma, \infty}$$

where $\| \cdot \|_{\gamma, \infty}$ is the Hölder norm of order $\gamma > 1/2$ and the constant c_γ depends only on γ.

Note that the above coherence bound does not require full, circular view of the scene, but the smoothness of the sampling density function f which depends indirectly on the size of the support of f. Lemma 4 in conjunction with Proposition 2 yield an error bound for BPDN which depends entirely on the sampling angles and high base frequency and thus is a very crude estimate.

To see how spatial frequency sampling can improve the coherence bound, consider the case of circular SAR with randomly and uniformly distributed looks on $[0, 2\pi]$. Then the argument in Appendix C shows that

$$(72) \quad \bar{\mu} = \max_{\mathbf{p} \neq \mathbf{p}'} \frac{1}{m_1 m_2} \sum_{j=1}^{m_2} \mathbb{E}\left[\sum_{k=1}^{m_1} e^{-2\pi \mathrm{i} \ell \nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')}\right] = \max_{\mathbf{p} \neq \mathbf{p}'} \frac{1}{m_2} \sum_{j=1}^{m_2} J_0(2\pi \ell \nu_j |\mathbf{p} - \mathbf{p}'|)$$

where the zeroth order Bessel function J_0 has the large-argument asymptotic

$$(73) \qquad J_0(z) = \sqrt{\frac{2}{\pi z}} \left\{ \cos\left(z - \pi/4\right) + \mathcal{O}(1/|z|) \right\}, \quad z \gg 1.$$

On the one hand, the worst case bound is

$$(74) \qquad J_0(2\pi \ell \nu_j |\mathbf{p} - \mathbf{p}'|) < \frac{c}{\sqrt{\ell \nu_0}}, \quad \forall \mathbf{p} \neq \mathbf{p}', \quad \forall j$$

for some $c > 0$. On the other hand, if the standard resolution criterion

$$(75) \qquad \ell \alpha_1 T / c_0 \geq 1/2$$

is satisfied, then the argument $2\pi \ell \nu_j |\mathbf{p} - \mathbf{p}'|$ ranges over at least one period, with ν_j in (48) and $\mathbf{p}, \mathbf{p}' \in \mathcal{L}$, and one can deliberately select a sequence of spatial frequencies to minimize the right hand side of (72) in view of the sinusoidal nature of the leading asymptotic in (73). In principle, the leading order term in (73) can be made to cancel out which yields the improved bound

$$\bar{\mu} = \mathcal{O}((\ell \nu_0)^{-1}).$$

Likewise, the fluctuation part of μ, which is bounded by the second term of the right hand side of (70), can also be reduced by an irregular sampling of spatial frequencies.

More generally, let f be represented by the Fourier series

$$f(\theta) = \sum_l a_l e^{\mathrm{i} l \theta}.$$

Then

$$\bar{\mu} = \max_{\mathbf{p} \neq \mathbf{p}'} 2\pi \sum_l a_l e^{\mathrm{i} l (\theta_{\mathbf{p}'-\mathbf{p}} + \pi/2)} J_l(2\pi \ell \nu_j)$$

where J_l is the Bessel function of order l and $\theta_{\mathbf{p}'-\mathbf{p}}$ the angle of $\mathbf{p}' - \mathbf{p}$ (cf. (91)). The large argument asymptotic for J_l

$$(76) \qquad J_l(z) = \sqrt{\frac{2}{\pi z}} \left\{ \cos\left(z - l\pi/2 - \pi/4\right) + \mathcal{O}(|z|^{-1}) \right\}, \quad z \gg 1$$

suggests a similar reduction of $\bar{\mu}$ under (75) by judicious choice of spatial frequencies. The $|z|^{-1/2}$ decay of the asymptotic (76) indicates the spatial frequencies should weigh (slightly) more heavily on large spatial frequencies.

6. NUMERICAL EXAMPLES

In the following simulations, we use YALL1 [27] to solve BPDN. BPDN, however, does not take advantage of the common support of X and X'. The off-grid perturbations $\{\xi_k\}$ are i.i.d. uniform random variables in $[-0.4\Delta\tau, 0.4\Delta\tau]$.

In the first set of simulations, we test the case of immobile targets which produce noiseless signals according to (12) with the parameters $m = 64, n = 128$, SNR $= 100$, and $Q = 2$.

Fig. 1 shows the results of BPDN reconstruction of 10 well separated targets. Reconstruction based on eq. (8)-(10) is clearly superior to that based on eq. (13)-(15). Fig. 2 shows an even greater improvement when the targets are closely spaced.
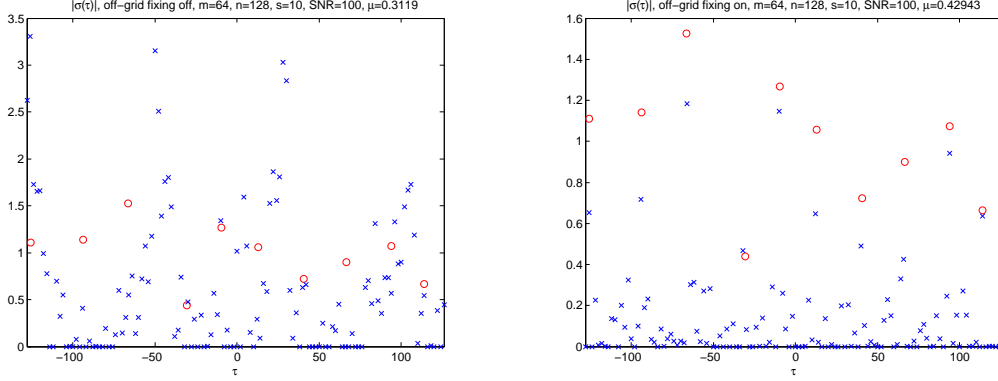
15

FIGURE 1. BPDN reconstruction (in blue) of off-grid targets with (left) eq. (8)-(10) and (right) eq. (13)-(15). Red circles represent the exact values.
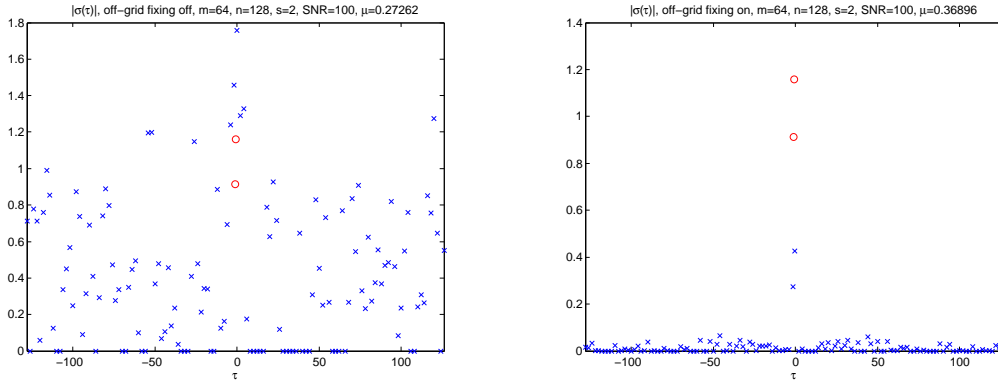

FIGURE 2. BPDN reconstruction (in blue) of closely spaced targets separated by $0.2\Delta\tau$ with (left) eq. (8)-(10) and (right) eq. (13)-(15). Red circles represent the exact values.

For the second set of simulations, we test the case of moving targets. The parameters are chosen as $m = 150$, $n = 625$, $s = 10$, SNR $= 100$, and $Q = 2$.

Fig. 3 shows significant improvement of reconstruction based on eq. (28)-(30) over that based on eq. (25)-(27).

In FIgure 4 and 5 we present numerical comparisons between LC and QC with the following parameters

- For LC, $\Delta\tau = 1$, $\Delta u = 1/m$, $m = n = 47$, $T = 1/\Delta u = m$, and the corresponding bandwidth is $B = 47$.
- For QC, $\Delta\tau = 1$, $\Delta u = 1/m$, $m = n = 47$, $\Delta t = \Delta\tau$, $T = n\Delta t$. By setting $\alpha_1 = 0$, the corresponding bandwidth is $B = 47$.

In both cases we set the carrier frequency $\omega_0 = 1000$. In addition to equally-spaced sampling in (34) we also consider random sampling with QC as in the case of LC.
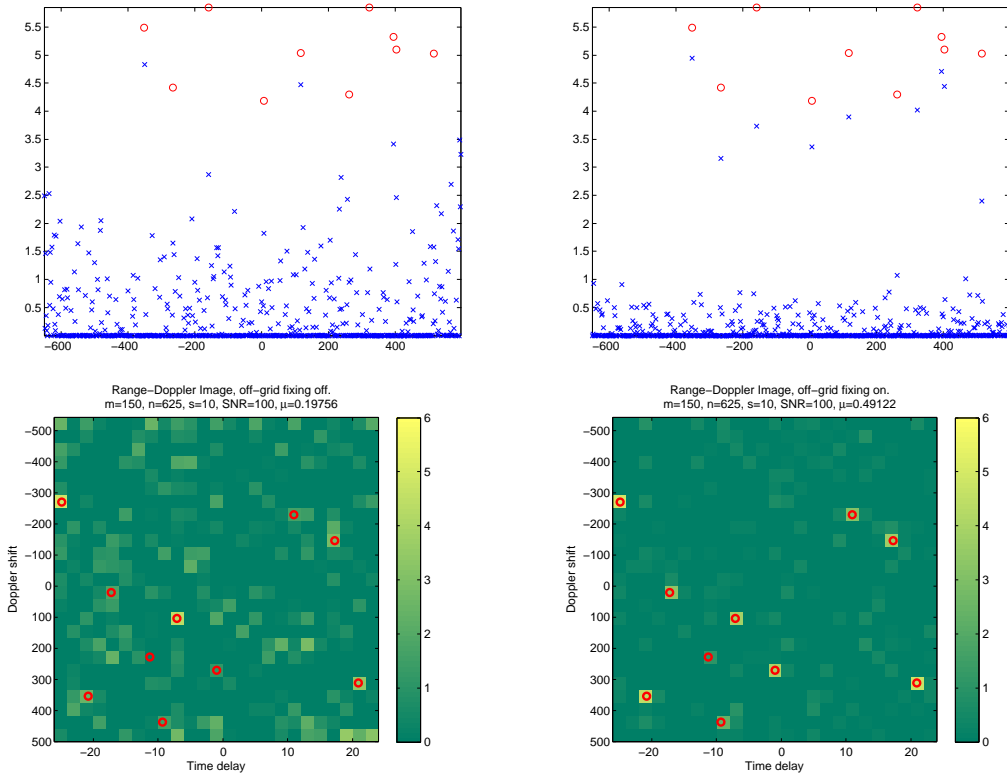
FIGURE 3. BPDN reconstruction of $|x_{k,l}|$ with (left) eq. (25)-(27) and (right) eq. (28)-(30). The top plots are the combined vector view and the bottom are the range-Doppler view.

In addition to BPDN, the greedy algorithm, Subspace Pursuit (SP) [9], is also used for reconstruction as the greedy algorithm tends to be much faster than solving BPDN. Moreover, one can easily modify the SP algorithm to take advantage of the prior information $\text{supp}X = \text{supp}X'$ in the case of off-grid targets (see (14) & (29)). Two definitions of "success" in recovery are used: the first one is defined by the relative mean square error (MSE) less than 0.1 and the second one is defined by exact recovery of the target support.

Fig. 4 and 5 show essentially no difference in performance between LC and QC as long as their bandwidth is the same. There is also no difference in performance between equally-spaced sampling and random sampling (labelled as "randt" in Fig. 4, 5) with QC. Fig. 4 shows that SP slightly outperforms BPDN (as implemented by YALL1).

For the case of localized extended targets, we use the 250×200 aerial image of Boston's Logan airport (Fig 6 left). The image with 19757 nonzero pixels is not strictly sparse but may be considered compressible. For the sake of speed we use SP with sparsity parameters much smaller than the actual sparsity, 19757, of the scene. The reconstruction with $m = 10000$ and sparsity parameter 2000 captures main features of the scene (Fig 6 middle). When the number of data increases to $m = 15000$ the reconstruction with sparsity parameter 4000 shows more fine-scale details (Fig 6 right).
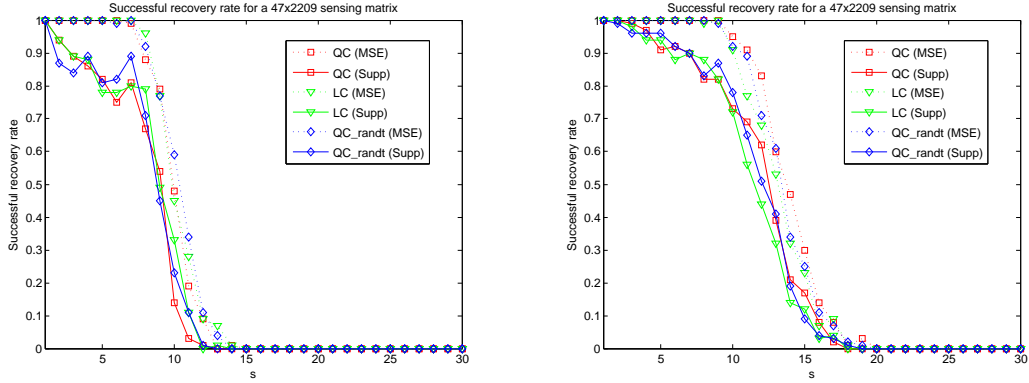
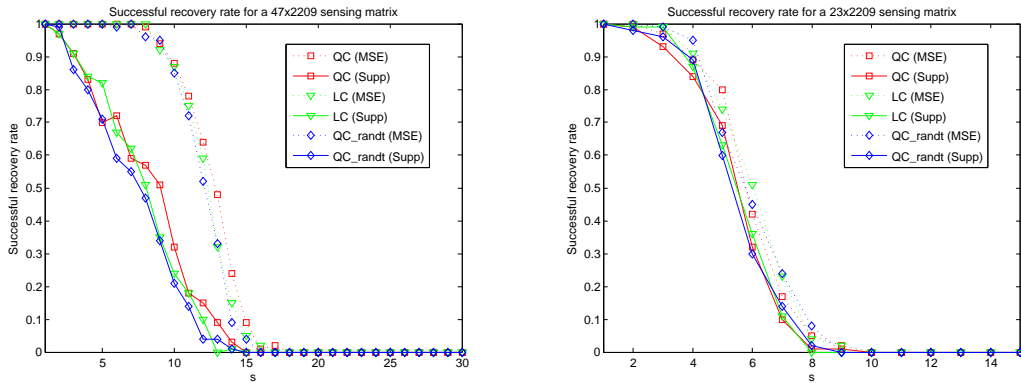FIGURE 4. Success probabilities of (left) BPDN and (right) SP reconstructions with 1% noise.

FIGURE 5. Success probabilities of SP reconstruction with (left) 5% and the full data $m = 47$ noise and (right) 1% noise (right) and an half data $m = 23$.
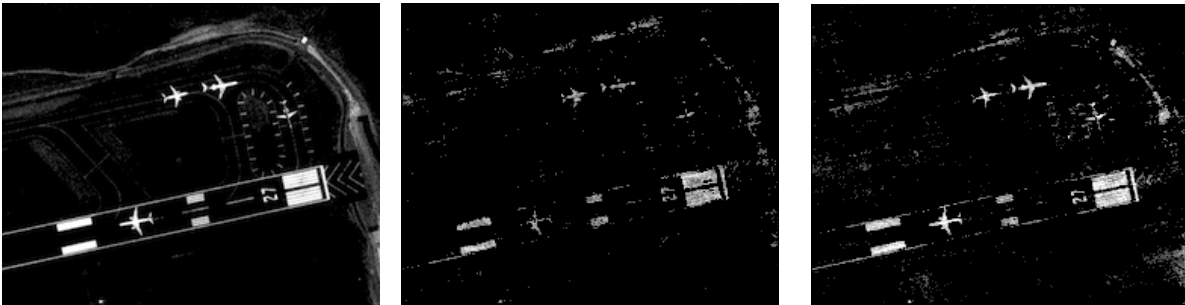

FIGURE 6. The 250×200 scene (left) and the SP reconstruction with (middle) $m = 10000$ and (right) $m = 15000$.

For the case of distributed targets, we simulate the data with the reflectivity function which depends only on one variable $\rho(r_1)$ and has a sparse representation in the Littlewood-Paley basis. Figure 7 shows the result of reconstruction with the following parameters: $p_* = 2, n_p = 100, \forall p$; for $p = -2, -1, 0, 1, 2$, $s_p = 12, 24, 13, 24, 23$, $m_p = 36, 64, 36, 64, 64$, $\epsilon = 0.5134, 1.4849, 0.6520, 1.0274, 1.3681$ equivalent to the relative noise level = 0.0500, 0.0649,
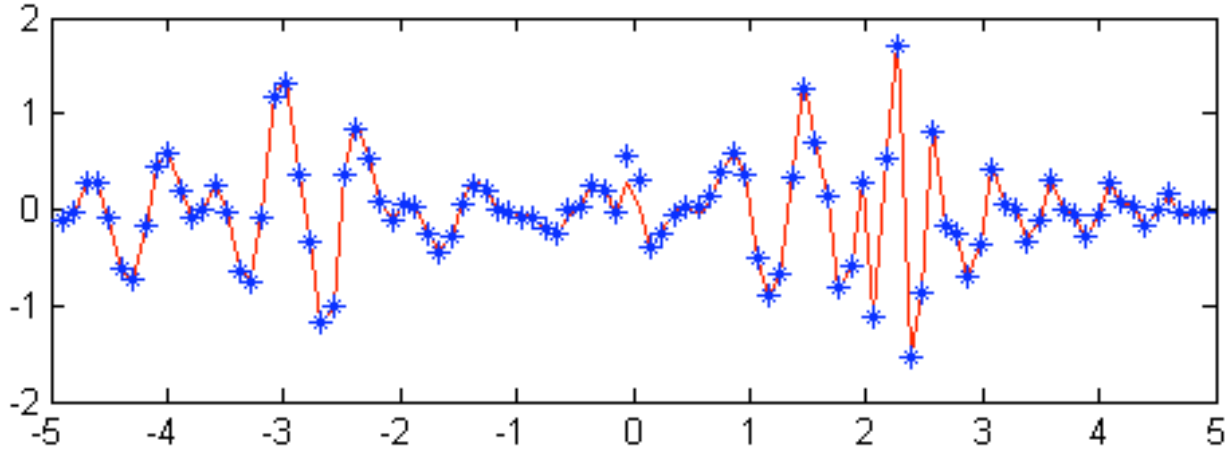
FIGURE 7. Imaging of a continuous scene in the presence of roughly 6% noise: The red-solid curve (top) is the exact profile and the blue ∗ shows the reconstructed profile.

$0.0494, 0.0640, 0.0698$. The resulting relative errors of reconstruction are $0.1941, 0.1240, 0.2498, 0.2361, 0.1502$.

7. CONCLUSION

We have explored compressed sensing approach to monostatic radar as well as spotlight SAR with chirped signal. In particular, we have proposed a simple method for dealing with off-grid targets and established a performance guarantee for the method.

We have extended our approach to the spotlight SAR in the tomographic formulation. We have proposed sampling schemes to deal with several classes of sparse targets under various bandwidth constraints and showed that the CS schemes can achieve standard resolution limits with sparse sampling. Our numerical experiments are consistent with the theoretical results.

APPENDIX A. PROOF OF LEMMA 1

Proof. We prove the coherence bound for the matrix $\mathbf{A} = [\mathbf{F} \ \mathbf{G}]$.

The k-th column vector A_k of $\mathbf{A}$ is given by

$$A_{jk} = \begin{cases} e^{-2\pi iQk\widehat{t}_j}, & k \le n \\ (\widehat{t}_j - 1/2)e^{-2\pi iQ(k-n)\widehat{t}_j}, & k > n. \end{cases}$$

Note that $\|A_k\|^2 = m, k \le n$ and $\mathbb{E}[\|A_k\|^2] = \frac{m}{12}, k > n$. Consequently, the scalar product of two distinct columns of $\mathbf{A}$ has three possible forms:

$$b_{kk'} = \sum_{j=1}^{m} a(\widehat{t}_j) \exp\left[2\pi iQ(k - k')\widehat{t}_j\right] \ , \quad a(t_j) = 1, \ (\widehat{t}_j - 1/2), \ \text{or} \ (\widehat{t}_j - 1/2)^2 \ ,$$

19

for $k, k' = 1, \ldots, n$. When $a(t_j) = 1$ or $(\hat{t}_j - 1/2)^2$ both columns are drawn from $\mathbf{F}$ or $\mathbf{G}$ and thus $k \neq k'$. When $a(t_j) = \hat{t}_j - 1/2$, one column is drawn from $\mathbf{F}$ and the other from $\mathbf{G}$. In the last case, k and k' are arbitrary.

Let $S_m = \sum_{j=1}^{m} U_j$, $T_m = \sum_{j=1}^{m} V_j$ where

$$U_j = a(\hat{t}_j) \cos\left[2\pi Q(k-k')\hat{t}_j\right] , \quad V_j = a(\hat{t}_j) \sin\left[2\pi Q(k-k')\hat{t}_j\right]$$

are independent (for different j) random variables in $[-1, 1]$. We have

$$|b_{kk'}| \leq |b_{kk'} - \mathbb{E}(b_{kk'})| + |\mathbb{E}(b_{kk'})|$$
$$= |S_m + iT_m - \mathbb{E}S_m - i\mathbb{E}T_m| + |\mathbb{E}(S_m + iT_m)|.$$

Recall the Hoeffding inequality.

Proposition 4. *Let $U_1, \ldots, U_m$ be independent random variables, and $S_m = \sum_{j=1}^{m} U_j$. Assume that $U_j \in [u, v]$, $j = 1, 2, \ldots, m$ almost surely, then we have*

$$(77) \qquad \mathbb{P}(|S_m - \mathbb{E}S_m| \geq mt) \leq 2\exp\left[-\frac{2m^2 t^2}{\sum_j (v-u)^2}\right]$$

for all positive t.

Choosing $t = K/\sqrt{m}$ for some constant K, we have

$$\mathbb{P}(|S_m - \mathbb{E}S_m| \geq \sqrt{m}K) \leq 2\exp\left[-K^2/2\right].$$

Note that the quantities S_m depend on $k - k'$ but there are at most $n - 1$ different values. The union bound yields

$$\mathbb{P}(\max_{k \neq k'}|S_m - \mathbb{E}S_m| \geq \sqrt{m}K) \leq 2(n-1)\exp\left[-K^2/2\right],$$

and similarly

$$\mathbb{P}(\max_{k \neq k'}|T_m - \mathbb{E}T_m| \geq \sqrt{m}K) \leq 2(n-1)\exp\left[-K^2/2\right].$$

We have

$$\mathbb{P}(\max_{k \neq k'}|b_{kk'} - \mathbb{E}b_{kk'}| < \sqrt{2m}K)$$
$$= \mathbb{P}(\max_{k \neq k'}|S_m + iT_m - \mathbb{E}S_m - i\mathbb{E}T_m| < \sqrt{2m}K)$$
$$> \left(1 - 2(n-1)\exp\left[-K^2/2\right]\right)^2 > (1-\delta)^2$$

if $\delta > 2n\exp\left[-K^2/2\right]$.

Now let us estimate the mean $\mathbb{E}(b_{kk'})$ or $\mathbb{E}(S_m + iT_m)$ for $k \neq k'$. Note that $\hat{t}_j$, $j = 1, \ldots, m$, are independently and uniformly distributed in $[0, 1]$. We have three different cases:

(1) For $a(\hat{t}_j) = 1$,

$$\mathbb{E}(b_{kk'}) = m\int_0^1 e^{2\pi i Q(k-k')t}\, \mathrm{d}t = 0.$$

(2) For $a(\widehat{t_j}) = \widehat{t_j} - 1/2$,

$$\mathbb{E}(b_{kk'}) = m \int_0^1 (t - 1/2)e^{2\pi iQ(k-k')t} \, dt = me^{\pi iQ(k-k')}\frac{(-1)^{(k-k')Q}}{2\pi i(k-k')Q}, \quad k \neq k',$$

and thus

$$|\mathbb{E}(b_{kk'})| \leq \frac{m}{2\pi Q}, \quad k \neq k'.$$

On the other hand,

$$\mathbb{E}(b_{kk}) = m \int_0^1 (t - 1/2) \cdot 1 \, dt = 0.$$

(3) For $a(\widehat{t_j}) = (\widehat{t_j} - 1/2)^2$,

$$\mathbb{E}(b_{kk'}) = m \int_0^1 (t - 1/2)^2 e^{2\pi iQ(k-k')t} \, dt = me^{\pi iQ(k-k')}\frac{(-1)^{(k-k')Q}}{(2\pi(k-k')Q)^2}, \quad k \neq k'$$

and thus

$$|\mathbb{E}(b_{kk'})| \leq \frac{m}{(2\pi Q)^2}.$$

For $k, k' \leq n$, since $\|A_k\|^2 = m$,

$$(78) \qquad b_{kk'} \leq \frac{C}{m}\sqrt{2m}K = C\frac{\sqrt{2}K}{\sqrt{m}}, \quad k, k' \leq n$$

for some universal constant C, with probability greater than $(1 - \delta)^2$.

On the other hand, for $k > n$,

$$\|A_k\|_2^2 = \sum_j (t_j - 1/2)^2$$

which is a sum of m i.i.d. random variables of mean $1/12$ on $[0, 1/4]$. Applying Hoeffding inequality with $t = 1/24$, we have

$$\mathbb{P}\Big(|\|A_k\|_2^2 - \frac{m}{12}| \geq \frac{m}{24}\Big) \leq 2e^{-m/18}$$

and thus

$$\mathbb{P}\Big(\|A_k\|_2^2 \leq \frac{m}{24}\Big) \leq 2e^{-m/18}.$$

We conclude from these observations that

$$(79) \qquad b_{kk'} \leq \frac{C}{m}\Big[\sqrt{2m}K + \frac{m}{2\pi Q}\Big] = C \cdot \Big[\frac{\sqrt{2}K}{\sqrt{m}} + \frac{1}{2\pi Q}\Big], \quad k \leq n < k'$$

$$(80) \qquad b_{kk'} \leq \frac{C}{m}\Big[\sqrt{2m}K + \frac{m}{(2\pi Q)^2}\Big] = C \cdot \Big[\frac{\sqrt{2}K}{\sqrt{m}} + \frac{1}{(2\pi Q)^2}\Big], \quad k, k' > n$$

with probability at least $(1 - \delta)^2 - 4e^{-m/18}$. (78)-(80) are what we set out to prove.

$\square$

Appendix B. Proof of Lemma 2

Proof. Observe that $\mathbf{AA}^* = \mathbf{FF}^* + \mathbf{GG}^*$ with

$$(\mathbf{FF}^*)_{i,j} = \sum_{k=1}^{n} e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)k} = e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)} \cdot \frac{1 - e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)n}}{1 - e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)}}$$

$$(\mathbf{GG}^*)_{i,j} = \sum_{k=1}^{n} (\widehat{t}_i - 1/2)(\widehat{t}_j - 1/2)e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)k} = (\widehat{t}_i - 1/2)(\widehat{t}_j - 1/2)e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)} \frac{1 - e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)n}}{1 - e^{2\pi i Q(\widehat{t}_j - \widehat{t}_i)}}.$$

Thus

$$(81) \qquad |(\mathbf{AA}^*)_{i,j}| = \left(1 + (\widehat{t}_i - 1/2)(\widehat{t}_j - 1/2)\right) \left| \frac{\sin\left[\pi Q(\widehat{t}_j - \widehat{t}_i)n\right]}{\sin\left[\pi Q(\widehat{t}_j - \widehat{t}_i)\right]} \right|$$

which can be controlled if the random variables $Z_{ij} = Q(\widehat{t}_j - \widehat{t}_i)$ are bounded away from integers. Z_{ij} has the density

$$f_Z = \mathrm{I}_{[0,Q]} * \mathrm{I}_{[0,Q]}$$

since $\widehat{t}_i, \widehat{t}_j$ are independently and uniformly distributed in $[0,1]$. Clearly f_Z is a triangular function on $[0, 2Q]$ and satisfies $|f_Z| \leq 1/Q$. Let

$$\zeta = \min_{i \neq j} \min_{k \in \mathbb{Z}} \{|Z_{ij} - k|\}.$$

Hence, for small $b > 0$,

$$\mathbb{P}(\zeta > b) > (1 - C_1 b)^{m \cdot (m-1)}$$

where the power counts for all possible pairs $(\widehat{t}_i, \widehat{t}_j)$, $1 \leq i \neq j \leq m$. Thus, with probability greater than $(1 - C_1 b)^{m \cdot (m-1)}$ we have

$$|(\mathbf{AA}^*)_{i,j}| < \frac{2}{\pi b}, \quad i \neq j.$$

By choosing $b = \frac{2(m-1)}{n\pi}$ and applying Gershgorin circle theorem, we have $\|\mathbf{AA}^* - n\mathbf{I}_m\|_2 < n$, or equivalently $\|\mathbf{A}\|_2^2 \leq 2n$. $\qquad \square$

Appendix C. Proof of Lemma 3

Proof. Recall Weyl's theorem (a generalized version of Gauss sum) [6], [23].

Proposition 5. *Let $r \in \mathbb{N}$, $r \geq 2$, and the prime number $q > r$. Then*

$$\left| \sum_{k=1}^{q} \exp\left(2\pi i(a_r k^r + a_{r-1} k^{r-1} + \cdots + a_2 k^2 + a_1 k)/q\right) \right| \leq (r-1)\sqrt{q}$$

provided $a_j \in \mathbb{N}$ for all $j = 1, 2, \ldots, r$ and $(a_1, a_2, \ldots, a_r) \neq (0, 0, \ldots, 0) \mod q$.

Assume also $\tau_k = k\Delta\tau$ and $u_l = (l - n/2)\Delta u$. The inner-product of the p-th and the p'-th columns of $\mathbf{F}$ is then

$$(82) \qquad \langle F_{j,p}, F_{j,p'} \rangle = \sum_{j=1}^{m} \exp\left[-2\pi i \left(\frac{a_2}{m}j^2 + \frac{a_1}{m}j\right)\right], \quad p \neq p'$$

where

$$a_2 = \alpha_2 m \Delta t^2 \Delta\tau(k - k')$$
$$a_1 = -\alpha_2 m \Delta t \Delta\tau^2(k^2 - k'^2) + \alpha_1 m \Delta t \Delta\tau(k - k') + m \Delta t \Delta u(l - l').$$

To ensure that $a_2 \in \mathbb{N}$, we set

$$\alpha_2 = \frac{1}{m \Delta\tau \Delta t^2} \quad \text{(or its integer multiple)}$$

and, similarly, for $a_1 \in \mathbb{N}$ we set

$$\alpha_1 = \frac{1}{m \Delta t \Delta\tau} \quad \text{(or its integer multiple)}$$
$$\Delta u = \frac{1}{m \Delta t} \quad \text{(or its integer multiple)}.$$

Furthermore setting

$$m = n, \quad \Delta t = \Delta\tau = \frac{1}{n \Delta u}$$

we obtain

$$a_2 = k - k', \quad a_1 = -(k^2 - k'^2) + (k - k') + (l - l')$$

which do not have m as a common factor since $|k - k'| \le n - 1$. The sensing matrix $\mathbf{F}$ in (34) becomes

$$(83) \qquad F_{j,p} = \exp\left[-2\pi i \left(k\frac{j^2}{n} + (-k^2 + k + l)\frac{j}{n}\right)\right].$$

If we set $\alpha_1 = 0$, the resulting sequence is called Alltop sequence [19] of bandwidth $B = n/\Delta\tau$ and the resulting sensing matrix becomes

$$(84) \qquad F_{j,p} = \exp\left[-2\pi i \left(k\frac{j^2}{n} + (-k^2 + l)\frac{j}{n}\right)\right].$$

To apply Weyl's theorem, we need only to check that $(a_1, a_2) \ne (0,0) \mod n$. Since $|k - k'|, |l - l'| \le n - 1$, $a_2 = 0 \mod n$ if and only if $k = k'$, which implies $a_1 = 0$ if and only if $l = l'$. This can not happen if $p \ne p'$. By Weyl's theorem,

$$\langle F_{j,p}, F_{j,p'} \rangle \le \sqrt{n}.$$

implying

$$\mu(\mathbf{F}) = \max_{p \ne p'} \frac{|\langle F_p, F_{p'} \rangle|}{\|F_p\|\|F_{p'}\|} \le \frac{1}{\sqrt{n}}.$$

$\square$

Appendix D. Proof of Lemma 4

Proof. The mutual coherence has the form

$$
(85) \qquad \max_{\mathbf{p} \neq \mathbf{p}'} \frac{1}{m_1 m_2} \left| \sum_{j=1}^{m_2} \sum_{k=1}^{m_1} e^{-2\pi i \ell \nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')} \right|.
$$

Consider the first summation over $k = 1, ..., p$. Let

$$
P_k = \cos\left(2\pi\ell\nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')\right), \quad Q_k = \sin\left(2\pi\ell\nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')\right)
$$

and

$$
S_{m_1} = \sum_{k=1}^{m_1} P_k, \quad T_{m_1} = \sum_{k=1}^{m_1} Q_k.
$$

Then the summation can be bounded by

$$
(86) \quad \left| \sum_{k=1}^{m_1} e^{-2\pi i \ell \nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')} \right| \leq \sqrt{|S_{m_1} - \mathbb{E}S_{m_1}|^2 + |T_{m_1} - \mathbb{E}T_{m_1}|^2} + \sqrt{|\mathbb{E}S_{m_1}|^2 + |\mathbb{E}T_{m_1}|^2}
$$

We apply the Hoeffding inequality to both S_{m_1} and T_{m_1}. With

$$
t = K/\sqrt{m_1}, \quad K > 0
$$

we obtain

$$
(87) \qquad \mathbb{P}\left[m_1^{-1} |S_{m_1} - \mathbb{E}S_{m_1}| \geq K/\sqrt{m_1} \right] \leq 2e^{-K^2/2}
$$

$$
(88) \qquad \mathbb{P}\left[m_1^{-1} |T_{m_1} - \mathbb{E}T_{m_1}| \geq K/\sqrt{m_1} \right] \leq 2e^{-K^2/2}.
$$

Note that the quantities S_{m_1}, T_{m_1} depend on $\mathbf{p} - \mathbf{p}'$ but they possess the symmetry: $S_{m_1}(\mathbf{p} - \mathbf{p}') = S_{m_1}(\mathbf{p}' - \mathbf{p}), T_{m_1}(\mathbf{p} - \mathbf{p}') = -T_{m_1}(\mathbf{p}' - \mathbf{p})$. Furthermore, a moment of reflection reveals that thanks to the square symmetry of the lattice there are at most $n - 1$ different values $|S_{m_1}|$ and $|T_{m_1}|$ among the $n(n-1)/2$ pairs of $(\mathbf{p}, \mathbf{p}')$.

We use (87)-(88) and the union bound to obtain

$$
\mathbb{P}\left[\max_{\mathbf{p} \neq \mathbf{p}'} m_1^{-1} |S_{m_1} - \mathbb{E}S_{m_1}| \geq K/\sqrt{m_1} \right] \leq 2(n-1) \cdot e^{-K^2/2}
$$

$$
\mathbb{P}\left[\max_{\mathbf{p} \neq \mathbf{p}'} m_1^{-1} |T_{m_1} - \mathbb{E}T_{m_1}| \geq K/\sqrt{m_1} \right] \leq 2(n-1) \cdot e^{-K^2/2}
$$

where the factor $4n$ is due to the structure of square lattice. Hence, by (86)

$$
(89) \quad \mathbb{P}\left[\max_{\mathbf{p} \neq \mathbf{p}'} m_1^{-1} \left| \sum_{k=1}^{m_1} e^{-2\pi i \ell \nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')} - \mathbb{E}\left[\sum_{k=1}^{m_1} e^{-2\pi i \ell \nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')} \right] \right| < \sqrt{\frac{2}{m_1}} K \right] > (1 - 2(n-1)e^{-K^2/2})^2.
$$

By (69) the right hand side of (89) is greater than $(1 - \delta)^2$.

Consider the identity

$$
(90) \quad \frac{1}{m_1} \mathbb{E}\left[\sum_{k=1}^{m_1} e^{-2\pi i \ell \nu_j \widehat{\mathbf{d}}_k \cdot (\mathbf{p} - \mathbf{p}')} \right] = \int_0^{2\pi} e^{-2\pi i \ell \nu_j \widehat{\mathbf{d}} \cdot (\mathbf{p} - \mathbf{p}')} f(\theta) d\theta, \quad \widehat{\mathbf{d}} = (\cos\theta, \sin\theta).
$$

24

Expanding f in the Fourier series

$$f(\theta) = \sum_l a_l e^{il\theta}$$

and denoting the angle of $\mathbf{p}' - \mathbf{p}$ by $\theta_{\mathbf{p}'-\mathbf{p}}$ we can write (90) as

$$\sum_l a_l e^{il\theta_{\mathbf{p}'-\mathbf{p}}} \int_0^{2\pi} e^{il\theta} e^{2\pi i \ell \nu_j |\mathbf{p}-\mathbf{p}'| \cos\theta} d\theta$$

(91)
$$= 2\pi \sum_l a_l e^{il(\theta_{\mathbf{p}'-\mathbf{p}}+\pi/2)} J_l(2\pi\ell\nu_j)$$

where J_l is the Bessel function of order l. The large argument asymptotic for J_l

$$J_l(z) = \sqrt{\frac{2}{\pi z}} \left\{ \cos\left(z - l\pi/2 - \pi/4\right) + \mathcal{O}(|z|^{-1}) \right\}, \quad z \gg 1$$

suggests the decay estimate (71).

Alternatively and more directly, the asymptotic of (90) can be derived by the method of stationary phase (Theorem XI. 14 and XI. 15 of [26]).

Proposition 6. *Let* $g_{\mathbf{p},\mathbf{p}'}(\theta) = \widehat{\mathbf{d}} \cdot (\mathbf{p}-\mathbf{p}')/|\mathbf{p}-\mathbf{p}'|$ *which is in* $C^\infty([-\pi,\pi]), \forall \mathbf{p}, \mathbf{p}' \in \mathcal{L}$.
(i) Suppose $\frac{d}{d\theta} g_{\mathbf{p},\mathbf{p}'}(\theta) \neq 0, \forall \theta \in [a,b], \forall \mathbf{p}, \mathbf{p}' \in \mathcal{L}$. *Then for all* $f \in C_0^h([a,b])$

(92)
$$\left| \int e^{-2\pi i \ell \nu_j |\mathbf{p}-\mathbf{p}'| g_{\mathbf{p},\mathbf{p}'}(\theta)} f(\theta) d\theta \right| \leq c_h (1 + \omega|\mathbf{p}-\mathbf{p}'|)^{-h} \|f\|_{h,\infty}$$

for some constant c_h *independent of* f. *Moreover, since* $\{g_{\mathbf{p},\mathbf{p}'} : \mathbf{p}, \mathbf{p}' \in \mathcal{L}\}$ *is a compact subset of* $C^{h+1}([a,b])$, *the constant* c_h *can be chosen uniformly for all* $\mathbf{p}, \mathbf{p}' \in \mathcal{L}$.

(ii) Suppose $\frac{d}{d\theta} g_{\mathbf{p},\mathbf{p}'}(\theta)$ *vanishes at* $\theta_* \in (a,b)$. *Since* $\frac{d^2}{d\theta^2} g_{\mathbf{p},\mathbf{p}'}(\theta_*) \neq 0$, *there exists a constant* $c_t, t > 1/2$ *such that*

(93)
$$\left| \int e^{-2\pi i \ell \nu_j |\mathbf{p}-\mathbf{p}'| g_{\mathbf{p},\mathbf{p}'}(\theta)} f(\theta) d\theta \right| \leq c_t (1 + \omega|\mathbf{p}-\mathbf{p}'|)^{-1/2} \|f\|_{t,\infty}$$

where the constant c_t *is independent of* $\mathbf{p}, \mathbf{p}' \in \mathcal{L}$.

Combining the estimates for (90) using (89) and the identity

$$UV = (U - \bar{U})(V - \bar{V}) + \bar{U}(V - \bar{V}) + \bar{V}(U - \bar{U}) + \bar{U}\bar{V}$$

we obtain (70) with probability greater than $1 - 2\delta m_2$.

$\square$

References

[1] W. Alltop, "Complex sequences with low periodic correlation," *IEEE Trans. Inform. Theory* **26(3)** (1980), 350354.

[2] R. Baraniuk and P. Steeghs, "Compressive radar imaging." *IEEE Radar Conf.* (2007), 128-133.

[3] E. J. Candès, "The restricted isometry property and its implications for compressed sensing." *Comptes Rendus Mathematique* **346** (2008), 589-592.

[4] E.J. Candès, Y.C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.,* **31** (2011), pp. 5973.

[5] E. J. Candès and Y. Plan, "Near-ideal model selection by ℓ_1 minimization." *Ann. Stat.* **37** (2009), 2145-2177.

[6] L. Carlitz and S. Uchiyama, "Bounds for exponential sums." *Duke Math J.* **24** (1957), 37-41.

[7] M. Cheney and B. Borden, "Imaging moving targets from scattered waves." *Inverse Probl.* **24** (2008), 035005.

[8] Y. Chi, L. L. Scharf, A. Pezeshki and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing." *IEEE T. Signal Proces.* **59** (2011), 2182-2195.

[9] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction." *IEEE T. Inform. Theory* **55** (2009), 2230-2249.

[10] I. Daubechies, *Ten Lectures on Wavelets.* SIAM, Philadelphia, 1992.

[11] M.F. Duarte and R.G. Baraniuk, "Spectral compressive sensing," preprint 2010.

[12] D. L. Donoho, M. Elad and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise." *IEEE T. Inform. Theory* **52** (2006), 6-18.

[13] J. Ender, "On compressive sensing applied to radar." *Signal Proces.* **90** (2010), 1402-1414.

[14] A. Fannjiang, "Compressive inverse scattering II. Multi-shot SISO measurements with Born scatterers." *Inverse Probl.* **26** (2010), 035009.

[15] A. Fannjiang, " Compressive inverse scattering with TV-MIN and greedy pursuit," *Math. Mech. Complex Syst.* 2012, to appear.

[16] A. Fannjiang and W. Liao, "Mismatch and resolution in compressive imaging," Wavelets and Sparsity XIV, edited by Manos Papadakis, Dimitri Van De Ville, Vivek K. Goyal, *Proc. SPIE* Vol. 8138, 2011.

[17] A. Fannjiang and W. Liao, "Coherence-Pattern guided compressive sensing with unresolved grids," *SIAM J. Imag. Sci.* **5** (2012), 179-202.

[18] M. J. Gerry, L. C. Potter, I. J. Gupta, and A. van der Merwe, "A parametric model for synthetic aperture radar measurements," *IEEE Trans. Antennas Propag.* **47**, pp. 11791188, 1999.

[19] M. Herman and T. Strohmer, "High resolution radar via compressed sensing." *IEEE Trans. Signal Process.* **57** (2009), 2275-2284.

[20] C. V. Jakowatz et al., *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach.* New York: Springer, 1996.

[21] J. B. Keller, "Geometrical theory of diffraction", *J. Opt. Soc. Amer.*, **5** pp. 116130, 1962.

[22] D. C. Munson, Jr., J. D. O'Brien, and W. K. Jenkins, "BA tomographic formulation of spotlight-mode synthetic aperture radar," *Proc. IEEE* **71**, pp. 917-925, 1983.

[23] J. L. Nelson and V. N. Temlyakov, "On the size of incoherent systems." *J. Approx. Theory* **163** (2011), 1238-1245.

[24] L. C. Potter, E. Ertin, J. T. Parker and Müjdat Çetin, "Sparsity and compressed sensing in radar imaging." *Proc. IEEE* **98** (2010), 1006-1020.

[25] H. Rauhut, "Stability results for random sampling of sparse trigonometric polynomials." *IEEE Trans. Inform. Theory* **54** (2008), 5661-5670.

[26] M. Reed and B. Simon, *Methods of Modern Mathematical Physics III. Scattering Theory.* Academic Press, San Diego, 1979.

[27] J. Yang and Y. Zhang, "Alternating direction algorithms for L1 problems in compressive sensing." *CAAM, Rice University* TR 09-37 (2010).

[1]Department of Mathematics, University of California, One Shields Ave., Davis, CA 95616-8633, USA. Corresponding author: fannjiang@math.ucdavis.edu