Cognitive Random Stepped Frequency Radar
with Sparse Recovery

Tianyao Huang, Yimin Liu, Huadong Meng, Xiqin Wang

Abstract

Random stepped frequency (RSF) radar, which transmits random-frequency pulses, can suppress the
range ambiguity, improve convert detection, and possess excellent electronic counter-countermeasures
(ECCM) ability [1]. In this paper, we apply a sparse recovery method to estimate the range and Doppler
of targets. We also propose a cognitive mechanism for RSF radar to further enhance the performance
of the sparse recovery method. The carrier frequencies of transmitted pulses are adaptively designed
in response to the observed circumstance. We investigate the criterion to design carrier frequencies,
and efficient methods are then devised. Simulation results demonstrate that the adaptive frequency-
design mechanism significantly improves the performance of target reconstruction in comparison with
the non-adaptive mechanism.

Index Terms

Random stepped frequency radar, adaptive waveform design, sparse recovery, Subspace Pursuit,
compressed sensing

I. INTRODUCTION

Since stepped frequency (SF) waveforms can synthesize very wide frequency bands with a
narrow bandwidth receiver, they are widely used in radars to generate high-range-resolution
profiles (HRRPs), Synthetic Aperture Radar (SAR) imaging, Inverse SAR (ISAR) imaging, etc.
[2–4]. In SF radar, the carrier frequencies of pulse trains are linearly varied with a constant
frequency step, which produces a ridge in the range-Doppler ambiguity function [1]. When the
transmitted frequencies are changed randomly, rather than linearly, the ridge ambiguity function
can be enhanced to a thumbtack function [5].

In random stepped frequency (RSF) radar, the carrier frequencies are randomly chosen from
a given bandwidth [1]; see Fig. 1 for the comparison of radar waveforms between SF and RSF
radar. Compared with linear SF radar, RSF radar further improves the range-Doppler resolution, suppresses the range ambiguity, and decouples the range and the Doppler \[6\]. This technique is attractive for its merits on electronic counter-countermeasures (ECCM) \[5\] and significantly reduces interference between adjacent radar systems \[1\]. The RSF waveforms were implemented in a wide-angle SAR to mitigate aliasing artifacts \[7, 8\] and were used in an ISAR to suppress the Doppler ambiguity \[9\]. In this paper, we focus on estimating the ranges and Doppler of multiple targets with RSF radar.

Sparse recovery and compressed/compressive sensing (CS) have received significant attention in radar signal processing \[10–13\]. By exploiting the sparsity, the theory of CS promises to exactly recover a sparse vector of length $N$ with high probability from much fewer than $N$ measurements \[12\]. In applications of RSF radars, the number of targets in the same coarse range bin is usually small, which forms a sparse scenario. The CS methods are applicable to detect the targets and recover the ranges, velocities and scattering intensities of the targets.

In order to enhance the performance with sparse modeling, the idea of cognitive radar is introduced to make use of the priori information of the target scenario. Cognitive radar was
first proposed in [14] and has attracted increasing research interest for a number of years [15–17]. In these literatures, cognitive radar is defined as a radar that can adaptively vary the transmission waveform according to the environmental information obtained by the radar. By exploiting the circumstance information, cognitive radar provides a significant improvement on radar performance. In the context of RSF radar, we adaptively design frequencies of transmitted pulses according to the observed target scene to further improve the performance of compressed sensing and pursue more accurate reconstruction of targets. The framework of cognitive RSF radar is demonstrated in Fig. 2.

![Fig. 2. The framework of cognitive RSF radar [18].](image)

Similar ideas, applying the cognitive concept to compressed sensing radar, are also found in [10, 18, 19], which significantly improve in performance over the corresponding radars without cognitive mechanism. Sen et al. [19] adaptively design the amplitudes of the transmitted subcarriers in an orthogonal-frequency-division multiplexing (OFDM) radar. Gogineni et al. [10] develop an adaptive energy-allocation mechanism for different transmitting antennas of multiple-input multiple-output (MIMO) radar system. Zhang et al. [18] optimize the sensing matrix and the phases of continuous phase-coded waveforms. However in this paper, we target optimizing the carrier frequencies of transmitted signals in RSF radar. The optimization parameters differ from those in [10, 18, 19], and these waveform-design methods are not directly applicable in RSF radar. Gogineni et al. [20] and Han et al. [21] consider the carrier frequencies design problem with a sparse model for a frequency-hopping MIMO radar. In [20, 21], the carrier frequencies are designed to reduce the block coherence measure of the sensing matrix. The block coherence is regardless of the target scenario. While in our paper, the carrier frequencies are adaptively optimized to fit the target scenario. The priori information of the targets is exploited to improve the reconstruction performance.
Aiming at reducing the reconstruction errors of RSF radar, the criterion, minimizing the Cramer-Rao bound (CRB) (or Cramer-Rao lower bound, CRLB) of sparse recovery is applied to design the carrier frequencies. The performance of RSF radar can be affected by the target scenario if the target returns interfere each other. The CRB depends on the target scenario and the carrier frequencies. The CRB can be seen as a measure of the interference as shown later in Subsection [IV-B]. Minimizing the CRB by designing the carrier frequencies reduces the interferences between target returns and thus enhances the recovery performance. Considering different potential applications of the cognitive scheme, we devise several efficient algorithms to calculate the optimal transmitting frequencies of radar pulses. For computational convenience, an approximation to the CRB criterion is also proposed for RSF radar. The consistency between two criterions is analyzed.

The rest of the paper is organized as follows. In Section [II] the echo signal model of RSF radar is introduced. In Section [III] we apply a compressed sensing algorithm to reconstruct the target scene and introduce some research results on the lower bound of sparse recovery errors. Then, in Section [IV] we present an adaptive waveform design approach to reduce the lower bound. The merits of the proposed mechanism are demonstrated in Section [V] with some simulation results. Section [VI] is devoted to a brief conclusion.

II. Radar Echo Signal Model

In this section, we describe the signal model of RSF radar. In each coherent processing interval (CPI), $N$ monotone pulses are transmitted with a constant pulse repetition interval $T$. The duration of each pulse is $T_p$. The synthetic bandwidth of the baseband is $B$. The frequency of the $n$th pulse is $f_n \in [f_c, f_c + B]$, $n = 0, 1, \ldots, N-1$, where $f_c$ is the central carrier frequency. Random frequency is set as $f_n = f_c + d_n \Delta f$, where $\Delta f$ is the frequency step size, and $d_n$ is a random integer between 0 and the floor integer $\lfloor B/\Delta f \rfloor$. To avoid ‘ghost image’ phenomenon, the frequency step size $\Delta f$ should be less than $1/T_p$ [22]. This is further discussed in the next-to-last paragraph of this section. Actually, the frequency step size can be rather small benefiting from the development of Direct Digital Synthesizer (DDS) technique. For example, for a bandwidth $B = 400$ MHz, a frequency resolution of $\Delta f \approx 0.23$ Hz can be achieved using the AD9910 [23].

In this case, the integer $\lfloor B/\Delta f \rfloor$ is huge. For notational brevity, the carrier frequency is rewritten as $f_n = f_c + c_n B$, where $c_n = d_n \Delta f/B \in [0, 1]$ is called as the $n$th frequency-modulation code.
Since $B/\Delta f$ is huge, $c_n$ is assumed as a continuous, real number. The $n$th transmitted pulse is described as

$$T_x(n, t) = \text{rect}\left(\frac{t - nT}{T_p}\right) e^{j2\pi(f_c + c_n B)(t - nT)}, \quad (1)$$

where $\text{rect}(\cdot)$ is a rectangular function defined as

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

Based on the "stop and hop" assumption, the echo of the $n$th pulse from a scatterer is

$$R_x(n, t) \approx \beta T_x(n, t - \frac{2r(t)}{c}), \quad (3)$$

where $c$ is the wave propagation speed. $\beta$ and $r(t)$ are the scattering intensity and the range of the target at instant $t$ with respect to the radar, respectively. Suppose the target is moving radially at a constant speed $v$, then $r(t) = r(0) + vt$. In this paper, it is simply assumed that tangential or rotational motion, and acceleration (and higher-order terms) of the target are ignorable or have been compensated previously. Refer to [6] and the references therein for details of motion compensation for RSF radar.

The $n$th echo is sampled at fast-time instant $t_s(n, l_r) = nT + l_r/f_s$, where $f_s$ is the sampling rate and $l_r = 1, 2, \ldots, L_r$ $(L_r < T f_s)$ denotes the index of samples. In the case that monotone pulses are transmitted, the sampling rate should be no less than $1/T_p$ such that no return will be missed. It is set as $f_s = 1/T_p$ in this paper. At $t_s(n, l_r)$, echoes from targets located between $r_s(l_r - 1)$ and $r_s(l_r)$ will be sampled, where $r_s(l_r) = l_r/f_s \cdot c/2$ denotes the range corresponding to the $l_r$th sampling time instant; see Fig. 3. The zone $[r_s(l_r - 1), r_s(l_r)]$ is called as a coarse-range bin.

![Fig. 3. Fast-time domain sampling. The sampling period is set as the pulse width $1/f_s = T_p$. As shown in the figure, after transmission of the radar pulse, the echo from the target located at $r_s(l_r - 1)$ arrives previously to echoes from $r > r_s(l_r - 1)$. Echo from targets located inside $[r_s(l_r - 1), r_s(l_r)]$ are sampled at $t_s(n, l_r) = nT + l_r/f_s.$](image-url)
Fig. 4. Samples at fast-time instant $nT + t_s$ comprise a serial of measurements, $n = 0, 1, \ldots, N - 1$.

Samples of successively transmitted pulses from the same coarse-range bin are collected; see Fig. 4. These data form the measurements within a CPI, and are used to generate HRRPs of the coarse-range bin. Samples from different coarse-range bins are processed individually. This paper focuses on HRRP, and without loss of generality only one coarse-range bin is considered. For notational brevity, the index $l_r$ of the coarse-range bin is omitted in the rest of paper. For example, $t_s(n, l_r)$ is simplified as $t_s(n)$. Substitute $t = t_s(n)$ and $r_s = (t_s(n) - nT)c/2$ into (1) and (3). The sampled $n$th echo (3) can be expressed as

$$R_x(n) = \beta e^{j2\pi(f_c + c_nB)(t_s - \frac{2v}{c}r_s - nT)}$$

$$= \beta e^{j4\pi(f_c + c_nB)\frac{r_s - r(t_s)}{c}}$$

$$= \beta e^{-j4\pi(f_c + c_nB)\frac{r(0) - r_s + vnT + 2vr_s/c}{c}}.$$  (4)

Since only one coarse-range bin is considered, the fast-time parameter $t_s$ is omitted in (4) and the rest of paper for simplicity. $R_x(n, t_s)$ is replaced by $R_x(n)$. Denote $R = r(0) - r_s$ as the high-resolution range of the target and ignore the term $2vr_s/c^2$, then

$$R_x(n) \approx \beta e^{-j4\pi(f_c + c_nB)(R + vnT)/c}$$

$$= \beta e^{-j\frac{4\pi f_c R}{c} - j\frac{4\pi B R}{c} c_n - j\frac{4\pi f_c vT}{c} n c'_n},$$  (5)

where $c'_n = 1 + c_nB/f_c$. 

May 11, 2014 DRAFT
When there are \( K \) targets inside a coarse-range bin, the received signals are recast as linear combinations of echoes from different targets and (5) is rewritten as

\[
R_x(n) \approx \sum_{k=1}^{K} \gamma_k \exp (j p_k c_n + j q_k n c'_n),
\]

(6)

where \( \gamma_k = \beta_k \exp (-j 4\pi f_c R_k / c) \), \( p_k = -4\pi B R_k / c \) and \( q_k = -4\pi f_c v_k T / c \) are used for notational brevity. \( \beta_k, R_k \) and \( v_k \) are the scattering magnitude, high-resolution range and velocity of the \( k \)th target, respectively. Since \(|\gamma_k|, p_k \) and \( q_k \) are proportional with \(|\beta_k|, R_k \) and \( v_k \), respectively, in the remainder of this paper they are simply regarded as the scattering coefficient, range and Doppler parameter of the \( k \)th target, respectively.

The HRRPs are synthesized with the observations in the form of (6). Substitute \( c_n = d_n \Delta f / B \) into (6), it can be implied that the unambiguous scope of HRRPs is \( c / (2\Delta f) \) \cite{1}. It should be larger than the scope of a coarse-range bin \( c T_p / 2 \) in order to avoid ghost image \cite{22}, which yields \( \Delta f < 1 / T_p \) as stated in the start of this section. After we obtain all of the sampled data at the same fast-time instant \( t_s \) in a CPI, these data are then used to reconstruct \( \gamma, p \) and \( q \) of all the targets.

In many practical cases, radar signal returns are corrupted by thermal noises and clutters. Noise is discussed in ensuing sections, but we simply assume in this paper that the returns have been filtered for clutter reduction prior to synthesizing HRRPs. Refer to \cite{24} for details of clutter cancelation algorithms for RSF radar. In addition, we limit the scope of this paper to 1-dimensional range profiling only. 2-D imaging (including azimuth dimension) with RSF radar remains for future work. Those readers interested in 2-D imaging with RSF radar are referred to \cite{6} and references therein.

III. SPARSE RECOVERY

A. Sparse Modeling

We discretize the possible range and Doppler values of the targets, i.e., \( p \) and \( q \) in (6), into \( P \) and \( Q \) grid points, respectively. Thus, we have \( PQ \) possible range-Doppler pairs \((p_l, q_l)\), \( l = 1, 2, \ldots, PQ \). Then, we can rewrite (6), the signal that the radar receives, as a combination of echoes from all possible targets,

\[
R_x(n) = \sum_{l=1}^{PQ} \gamma_l \exp (j p_l c_n + j q_l n c'_n),
\]

(7)
where $\gamma_l$ denotes the scattering coefficient of the target presented at $(p_l, q_l)$. If no target exists at $(p_l, q_l)$, $\gamma_l = 0$. We define a vector $\phi_l \in \mathbb{C}^N$, in which the $n$th element is

$$
\phi_l(n) = \exp (j p_l c_n + j q_l n c_n').
$$

Note that the element $\phi_l(n)$ is built from the modulation code $c_n$. Form a matrix $\Phi = [\phi_1, \phi_2, \ldots, \phi_{PQ}] \in \mathbb{C}^{N \times PQ}$ and modulation code sequence $c = [c_0, c_1, \ldots, c_{N-1}]^T \in \mathbb{R}^N$, where $(\cdot)^T$ denotes the transpose of a matrix or a vector. Note that $\Phi(c)$ depends on $c$. Unless specifically stated in the rest of the paper, we use $\Phi$ instead of $\Phi(c)$ for simplicity.

Generate $x = [\gamma_1, \gamma_2, \ldots, \gamma_{PQ}]^T \in \mathbb{C}^{PQ}$, which represents the scattering coefficients. Assume the echoes are corrupted by additive Gaussian white noise; thus, the received echoes (7) can be written in a matrix form as

$$
y = \Phi x + w,
$$

where $y \in \mathbb{C}^N$ represents the corrupted echoes and $w$ is a noise vector with a complex normal distribution $\mathcal{CN}(0, \sigma^2 I_N)$. $\sigma^2$ denotes the variance of the noise and $I_N$ denotes an identity matrix with dimension of $N$. $\Phi$ is often referred to as a dictionary matrix in literatures on compressed sensing. $x$ is an unknown vector to be recovered. $x$ is $K$-sparse, which means that there are only $K$ nonzero or $K$ prominent elements in $x$. Given $y$ and $\Phi$, sparse recovery solves $x$ with constraints on the sparseness of $x$. When $x$ is recovered, the high-resolution ranges and Doppler of the targets can be inferred from the support set $\Lambda = \text{supp}(x)$, where $\text{supp}(\cdot)$ denotes the set that consists of the indices of nonzero elements in the vector.

### B. Sparse Recovery

Sparse recovery algorithms estimate $x$ in (9) by exploiting its sparsity. In this paper, Subspace Pursuit (SP) [25] is adopted as the sparse recovery algorithm. The SP algorithm is a kind of greedy approach [25, 26], and possesses a provable reconstruction capability comparable to that of Basis Pursuit [27] and the Dantzig Selector [28] approaches and a low computational complexity similar to that of the Orthogonal Matching Pursuit (OMP) [29] and Regularized Orthogonal Matching Pursuit (ROMP) [30] approaches. More precisely, SP is a greedy approach to solve the $\ell_0$ minimization problem

$$
\min \|x\|_0, \text{ subject to } \|y - \Phi x\|_2 < \eta.
$$

(10)
where $\| \cdot \|_0$ and $\| \cdot \|_2$ denote the $\ell_0$ and Euclidean ($\ell_2$) norm of a vector, respectively. $\eta$ represents the power of noise.

We recall the main steps of the SP algorithm [25] in Algorithm I, where $x_\Lambda/\Phi_\Lambda$ denotes a sub-vector/matrix that consists of entries/columns indexed in the set $\Lambda$. $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo inverse, i.e., $A^\dagger = (A^H A)^{-1} A^H$, where $(\cdot)^H$ is the Hermitian transpose.

**ALGORITHM I**

**THE STANDARD SP ALGORITHM**

1) Input $K, y, \Phi$. Set the support set $\Lambda^{(0)} = \emptyset$, the residual error $r^{(0)} = y$, and the iteration counter $i = 0$.

2) Calculate the correlations $p = \Phi^H r^{(i)}$.

3) Merge the set $\tilde{\Lambda} = \Lambda^{(i)} \cup \{K \text{ indices corresponding to the largest magnitude entries in } p\}$.

4) Set $x_p = \Phi^\dagger_{\tilde{\Lambda}} y$, and update the set $\Lambda^{(i+1)} = \{K \text{ indices corresponding to the largest entries in } x_p\}$.

5) Update the residual error $r^{(i+1)} = y - \Phi_{\Lambda^{(i+1)}} \Phi^\dagger_{\Lambda^{(i+1)}} y$.

6) Increase $i$. Return to Step 2 until stop criterion, e.g., $\Lambda^{(i)} = \Lambda^{(i-1)}$, is satisfied.

7) Output $\hat{x}$, where $\hat{x}_{\Lambda^{(i)}} = \Phi^\dagger_{\Lambda^{(i)}} y$ and the rest entries are all zeros.

We cite here some brief lower bound analysis on sparse recovery errors, and in Section IV we target adaptively reducing the lower bound via designing radar waveforms. If the lower bound can be achieved by some sparse recovery methods, reduction in the bound yields decrease in errors of these recovery methods. Denote a solution to the sparse recovery problem in (10) as $\hat{x}$; thus, the recovery error can be described as $\|x - \hat{x}\|_2^2$. The Cramer-Rao bound (CRB) or Cramer-Rao lower bound (CRLB) is a well-known tool in estimation theory that expresses a lower bound on the variance of any unbiased estimator of an unknown deterministic parameter [31]. For estimating the sparse vector $x$ of $\|x\|_0 = K$, Ben-Haim etc. [32] derive the constrained CRB as

$$E\left[\|x - \hat{x}\|_2^2\right] \geq \sigma^2 \text{tr}\left((\Phi_{\Lambda^*}^H \Phi_{\Lambda^*})^{-1}\right), \quad \|x\|_0 = K,$$

(11)

where $E[\cdot]$ denotes the expectation of a random variable, $\text{tr}(\cdot)$ denotes the trace of a matrix and $\sigma^2$ is the variance of noise in (9). $\Lambda^*$ is the true support set. Ben-Haim etc. also state that the constrained CRB can be attained when a large number of independent measurements are
available via the Maximum-Likelihood approach \[32\]

\[
\hat{x} = \arg \min_x \|y - \Phi x\|_2^2, \quad \text{subject to } \|x\|_0 \leq K.
\] (12)

In [28], Candes presents an oracle sparse estimator, in which a Genie provides the true support set \(\Lambda^*\). Denote \(\hat{x}_{\text{oracle}}\) as the oracle estimate of \(x\). The mean squared error (MSE) on the estimation of \(x\) is

\[
E \left[ \|x - \hat{x}_{\text{oracle}}\|_2^2 \right] = \sigma^2 \text{tr} \left( (\Phi^H_{\Lambda^*} \Phi_{\Lambda^*})^{-1} \right).
\] (13)

For any unbiased estimator \(\hat{x}\) of \(x\), the variance of the error \(E \left[ \|x - \hat{x}\|_2^2 \right] \geq E \left[ \|x - \hat{x}_{\text{oracle}}\|_2^2 \right]\) \[33\], which is the same as the constrained CRB in (11).

The achievability of the CRB for noisy sparse recovery has been reported in [28, 33, 34]. The Dantzig Selector [28], which is based on linear programming, achieves the error in (13) up to a factor of \(\log(PQ)\). Note that in the Dantzig Selector, a priori knowledge of \(K\) is not necessary. Babadi et al. [33] establish a joint typicality estimator that asymptotically achieves the CRB without any information about the support set \(\Lambda^*\) as the number of measurements \(N \to \infty\) for \(\Phi\), a random Gaussian matrix of which elements are drawn i.i.d. (independent identically distributed) from \(\mathcal{N}(0, 1)\). Niazadeh et al. [34] generalize the conditions for the problem of the asymptotic achievability of CRB. They relax the Gaussianity constraint on \(\Phi\) assuming that \(\Phi\) is randomly generated according to a distribution that satisfies some sort of concentration of measures inequality [34].

IV. OPTIMAL CODE DESIGN

A. Optimization Criterion

In cognitive RSF radar, we use a priori information about the target scene to adaptively design the modulation code sequence \(c\) to better recover the targets. The metric for the recovery performance is mean square errors, \(E \left[ \|x - \hat{x}\|_2^2 \right]\), which is commonly used in radar application. MSE is a good way to capture the systems performance. The estimation error of every target is equally indicated in MSE and the MSE is small only when all targets are accurately estimated. To reduce the MSE, we choose the strategy minimizing the lower bound on the recovery error \(E \left[ \|x - \hat{x}\|_2^2 \right]\) in (11)

\[
c_{\text{opt}} = \arg \min_c \text{tr} \left( (\Phi^H_{\Lambda^*} \Phi_{\Lambda^*})^{-1} \right) = \arg \min_c \text{tr} \left( (A^H A)^{-1} \right),
\] (14)
where $A = \frac{1}{\sqrt{N}} \Phi_A \in \mathbb{C}^{N \times |\Lambda|}$, and $|\Lambda|$ denotes the cardinality of $\Lambda$. Note that the sub-dictionary $\Phi_A$ and $A$ are functions of code sequence $c$; see (8). Since the correct support set $\Lambda^*$ is actually unknown, we use the previous estimate $\Lambda = \text{supp}(\hat{x})$ instead.

From the perspective of the estimation theory, CRB is the inverse of the Fisher information [31], which can be seen as a measure of efficiency of the sensing system $y = \Phi x + w$. The sensing system can be more informative if $\Phi$ is designed to lower the CRB. From the perspective of radar system, the carrier frequencies are designed to adapt to the target scenario. Minimizing the CRB reduces interference between target returns and contributes to the reconstruction of targets. This is further discussed in the next-to-last paragraph in Subsection IV-B. Simulation results in Subsection V-A demonstrate the effect of the CRB criterion on enhancing the reconstruction performance.

In Subsection IV-B, we propose an approximation of the objective function in (14) for computability. For different potential applications of the cognitive mechanism, we develop two types of efficient algorithms to calculate the optimal modulation codes in Subsection IV-C and IV-D, respectively. In Subsection IV-C we develop a steepest descent method for batch-oriented codes optimization, in which a batch of codes are designed in each optimization operation and the previous measurements are not used again. Subsection IV-D describes a sequential code design method that designs only one code in one operation; the previous measurements are reused. The batch-oriented and sequential operation modes are demonstrated in Fig. 5 and Fig. 6, respectively.

Fig. 5. Batch-oriented code design. In each code-optimization operation, a batch of codes, which constitutes a CPI, are designed based on the estimation results with last CPI data. The previous measurements are not used again.
B. Objective Function Approximation

To apply steepest descent method to solve the optimal problem (14), we need to calculate the gradient vector of \( \text{tr} \left( (A^H A)^{-1} \right) \) with respect to \( c \). This task is quite difficult because the matrix inverse is involved. In this subsection, the original objective function in (14) is approximated by

\[
\mathbf{c}_{opt} = \arg \min_{\mathbf{c}} \text{tr} \left( A^H A A^H A \right),
\]

such that it is easier to calculate the gradient vector because of elimination of the matrix inverse operation.

\( \text{tr} \left( A^H A A^H A \right) \) is adopted because it has a performance similar to \( \text{tr} \left( (A^H A)^{-1} \right) \) with respect to \( c \) in the scenario of RSF radar. Actually, both objective functions aim at forcing the eigenvalues of \( A^H A \) to approach 1, which is explained in the rest of this subsection.

In the original optimal problem (14),

\[
\text{tr} \left( (A^H A)^{-1} \right) = \sum_{i=1}^{|\Lambda|} \frac{1}{\lambda_i} \geq \frac{|\Lambda|^2}{\sum_{i=1}^{|\Lambda|} \lambda_i},
\]

where \( \lambda_i \) is the \( i \)th eigenvalue of \( A^H A \), i.e., \( A^H A = U \Sigma U^H \). \( \Sigma \in \mathbb{C}^{|\Lambda| \times |\Lambda|} \) is the characteristic matrix of \( A^H A \), which is a unitary matrix \( U U^H = U^H U = I_{|\Lambda|} \). \( \Sigma = \text{diag} \left( [\lambda_1, \lambda_2, \ldots, \lambda_{|\Lambda|}] \right) \in \mathbb{R}^{|\Lambda| \times |\Lambda|} \), where \( \text{diag} \) denotes a diagonal matrix with diagonal elements as indicated in the given vector. The inequality in (16) holds because the harmonic mean \( |\Lambda|/\sum_{i=1}^{|\Lambda|} \frac{1}{\lambda_i} \) is less than or equal to the arithmetic mean \( \sum_{i=1}^{|\Lambda|} \lambda_i/|\Lambda| \). Two means are equal if and only if there is a code sequence...
such that
\[ \lambda_1 = \lambda_2 = \cdots = \lambda_{|\Lambda|}. \] (17)

Referring to (8) and (14), since the \( l \)th diagonal element of \( A^H A \) equals 1, i.e.,
\[
(A^H A)_{ll} = \sum_{n=0}^{N-1} \frac{1}{\sqrt{N}} \exp(-j p_l c_n - j q_l n c'_n) 
\cdot \frac{1}{\sqrt{N}} \exp(j p_l c_n + j q_l n c'_n) = 1,
\] (18)
it always holds in the scenario of RSF radar that
\[
\sum_{i=1}^{|\Lambda|} \lambda_i = \text{tr} (A^H A) = \sum_{i=1}^{|\Lambda|} (A^H A)_{ll} = | \Lambda |;
\] (19)
thus, the equality in (16) holds if and only if
\[ \lambda_1 = \lambda_2 = \cdots = \lambda_{|\Lambda|} = \frac{\sum_{i=1}^{|\Lambda|} \lambda_i}{|\Lambda|} = 1, \] (20)
i.e., \( \Sigma = I_{|\Lambda|} \). As a result, \( A \) is a semi-unitary matrix, i.e.,
\[ A^H A = U \Sigma U^H = U U^H = I_{|\Lambda|}. \] (21)

As discussed above, minimizing the original objective function \( \text{tr} \left( (A^H A)^{-1} \right) = \sum_{i=1}^{|\Lambda|} \frac{1}{\lambda_i} \) provides a code sequence \( c \) that leads to \( \Sigma = I_{|\Lambda|} \) if it exists. When there is no such code sequence that results in a semi-unitary \( A \), \( \sum_{i=1}^{|\Lambda|} \frac{1}{\lambda_i} \) approaches the minimum if all of the eigenvalues \( \lambda_i \) are close to 1 because the eigenvalues obey the constraint \( \sum_{i=1}^{|\Lambda|} \lambda_i = | \Lambda |. \) We see that the original optimization seeks all eigenvalues close to 1.

Note that it is difficult to calculate the gradient of the objective function, minimizing \( \sum_{i=1}^{|\Lambda|} \frac{1}{\lambda_i} \) or \( \text{tr} \left( (A^H A)^{-1} \right) \), because of the matrix inverse involved. Since the original objective in (14) substantially aims at that all eigenvalues approach 1, we can approximately apply least squares
to force the eigenvalues close to 1; thus, the initial objective function is replaced with

$$c_{opt} = \arg \min_c \sum_{i=1}^{\vert \Lambda \vert} (\lambda_i - 1)^2$$

$$= \arg \min_c \sum_{i=1}^{\vert \Lambda \vert} \lambda_i^2 - \sum_{i=1}^{\vert \Lambda \vert} (2\lambda_i - 1)$$

$$= \arg \min_c \sum_{i=1}^{\vert \Lambda \vert} \lambda_i^2 - \vert \Lambda \vert$$

$$= \arg \min_c \text{tr} \left( A^H A A^H A \right) ,$$

where $\sum_{i=1}^{\vert \Lambda \vert} \lambda_i = \vert \Lambda \vert$ is substituted in the third line. We obtain the new objective function $\text{tr} \left( A^H A A^H A \right)$ in (15) instead of the former function $\text{tr} \left( (A^H A)^{-1} \right)$ in (14). It is easier to obtain the analytic gradient of the objective function in (15); see the ensuing subsection. Two objective functions in (14) and (15) have similar variation trends with respect to $c$, which is demonstrated by numerical examples in Section V.

As mentioned in the preceding paragraphs, both optimizations in (14) and (15) are making the sub-dictionary $A$ approach a semi-unitary matrix $A^H A \rightarrow I_{\vert \Lambda \vert}$, which means the columns of $A$ are optimized to be approximately perpendicular to each other. In a radar system, target returns which are highly correlated can interfere each other and are difficult to distinguish. Such returns can affect the reconstruction performance. In the context of RSF radar, a column in $A$ represents a normalized echo from a target, and interference between returns of two targets can be characterized by correlation between the columns in $A$, i.e. $\vert A_i^H A_j \vert$, $i \neq j$. Higher correlation result indicates that more intensive interference exist between the target returns. Adaptively designed $A$, in which the columns are perpendicular to each other, implies that echoes from different targets are orthogonal, and the interferences among target returns are thus avoided. The orthogonality contributes to better recovery of targets.

In [10] and [19], different criterions are applied for adaptive waveforms design to improve sparse recovery performance. In [10], the criterion is to maximize the minimum target returns. However, it is not applicable to the RSF radar. We assume that all radar pulses are transmitted at an invariant peak power. Thus, the power of the transmissions is not adjustable and the intensities of the target returns are independent of the transmitted waveforms. In [19], the criterion is based
on minimizing an upper bound on $\|x - \hat{x}\|_2^2$. However, the upper bound is not tight enough. Decreasing upper bound does not directly imply decrease in the recovery errors. In this paper, a lower bound on MSE, i.e. CRLB, is chosen as the optimum criterion to design the carrier frequencies. A lower bound defines the best performance that an estimator can obtain. The lower bound equals the MSE if the bound is achievable by some estimators. The achievability of CRLB is discussed in the last paragraph of Section II. Decreasing this bound can result in decrease in MSE. Note that both [10] and [19] concern with the power allocation of the waveforms. The methods involved in these two papers are not directly applicable to RSF radar, because the power of the transmissions in the RSF radar is non-adjustable.

C. Batch-oriented Code Design

We consider the situation in which a batch of codes are simultaneously designed; see Fig. 5. Since the bandwidth of transmitted waveforms is limited, the code sequence $c$ in RSF radar is optimized by solving the constrained minimization problem

$$\min_c \text{tr} \left( A^H A A^H A \right), \text{ subject to } 0 \preceq c \preceq 1. \tag{23}$$

where 0 and 1 denote a vector with all entries 0 and 1, respectively.

Our first step is to rewrite (23) as an unconstrained problem with variable substitutions. Since $f(z) = \arctan(z)$ has a definition domain $z \in (-\infty, \infty)$ and a range $f(z) \in (-\pi/2, \pi/2)$, we can apply the arctan function for variable substitution to eliminate the domain constraints on the code sequence $c$ in (23). Relax the constraints $c_n \in [0, 1]$ as $c_n \in (-\delta, 1+\delta)$, where $\delta$ is a small positive constant, $n = 0, 1, \ldots, N-1$. Introduce a vector $z = [z_0, z_1, \ldots, z_{N-1}]^T \in \mathbb{R}^N$, where

$$z_n = \tan \left( \frac{\pi}{1 + 2\delta} \left( c_n - \frac{1}{2} \right) \right), \tag{24}$$

or equivalently,

$$c_n = \frac{1}{2} + \frac{1 + 2\delta}{\pi} \arctan z_n. \tag{25}$$

Substitute (25) into (23); thus, the matrix $A$ becomes a function of $z$ and (23) is replaced with an unconstrained optimum problem

$$\min_z \text{tr} \left( A^H A A^H A \right), z \in \mathbb{R}^N. \tag{26}$$
The second step is to calculate the gradient $\frac{\partial \text{tr} (A^H A A^H A)}{\partial z}$ such that steepest descent method can be applied to solve (26). The gradient is

$$\frac{\partial \text{tr} (A^H A A^H A)}{\partial z} = 4 \text{Re} \left[ \text{diag} \left( DA^H A A^H \right) \right],$$

(27)

where the $n$th row, $i$th column element of $D \in \mathbb{C}^{N \times |\Lambda|}$ is $D(n, i) = dA(n, i)/dz_n$. $\text{Re}[]$ denotes the real part of a complex number. The derivation is presented in the Appendix.

For the steepest descent method, the computational complexity is determined by the required number of iterations and the computational load in each iteration. The convergence speed is discussed with numerical experiments in Subsection V-C. In each iteration, calculating the steepest descent direction with (27) and searching for the step size along the direction compose the main computational load. Calculating the gradient vector (27) requires $3/2 N |\Lambda|^2 + N |\Lambda|$ times complex-valued multiplication and addition. The locally optimal step size is decided by the values of cost function (26) and the calculation requires $1/2 N |\Lambda|^2 + |\Lambda|$ times complex-values multiplication and addition.

### D. Sequential Code Design

In this subsection, we discuss algorithms that design the frequency-modulation codes sequentially. In each design operation, only one code is optimized; see Fig. 6. Suppose $N$ codes $c \in \mathbb{R}^N$ have been utilized, and measurements $y \in \mathbb{C}^N$ are available. Our goal is to determine the next code $c_{N+1}$ with the previous data and the corresponding estimate results.

With the measurements $y$, assume that we obtain the estimate of the support set as $\Lambda$ and the corresponding sub-dictionary as $A = \Phi_{\Lambda} \in \mathbb{C}^{N \times |\Lambda|}$. When the new code $c_{N+1}$ is applied, a new row $a^H \in \mathbb{C}^{1 \times |\Lambda|}$ is added to the sub-dictionary $A$. Denote $A_{\text{new}} = [A^H a]^H$ as the new sub-dictionary; thus, the objective function for sequential code design is

$$c_{N+1} = \arg \min c \text{tr} \left( \left( A_{\text{new}}^H A_{\text{new}} \right)^{-1} \right),$$

(28)

or approximately,

$$c_{N+1} = \arg \min c \text{tr} \left( A_{\text{new}}^H A_{\text{new}} A_{\text{new}}^H A_{\text{new}} \right).$$

(29)

Let $B = A^H A$ and $F = B^{-1}$. Since

$$\left( A_{\text{new}}^H A_{\text{new}} \right)^{-1} = (B + aa^H)^{-1} = F - \frac{Faa^H F}{1 + a^H Fa},$$

(30)
the optimum problem (28) can be simplified as
\[ c_{N+1} = \arg \min_\mathbf{a} \, \frac{\text{tr} (\mathbf{F}^H \mathbf{F})}{1 + \mathbf{a}^H \mathbf{F} \mathbf{a}} \]
\[ = \arg \max_\mathbf{a} \frac{\mathbf{a}^H \mathbf{F}^2 \mathbf{a}}{1 + \mathbf{a}^H \mathbf{F} \mathbf{a}}. \]

The approximation (29) can be rewritten as
\[ c_{N+1} = \arg \min_\mathbf{B} \, \text{tr} (\mathbf{B}^2 + \mathbf{a} \mathbf{a}^H \mathbf{B} + \mathbf{B} \mathbf{a} \mathbf{a}^H + \mathbf{a} \mathbf{a}^H \mathbf{a} \mathbf{a}^H) \]
\[ = \arg \min_\mathbf{B} \, \text{tr} (\mathbf{B}^2) + 2 \mathbf{a} \mathbf{a}^H \mathbf{B} \mathbf{a} + (\mathbf{a}^H \mathbf{a})^2 \]
\[ = \arg \min_\mathbf{a} \mathbf{a}^H \mathbf{B} \mathbf{a}. \]

where \( \mathbf{a}^H \mathbf{a} = |\Lambda|; \) see (8).

Sequential code design with (31) or (32) is computationally efficient. Both (31) and (32) have simple and analytical expressions, and they are single-parameter (i.e., \( c_{N+1} \)) optimization problems. Exhaustive search or other numerical methods \([35]\) can be applied to solve them. Since there is no matrix-inverse operation involved in (32), the approximated objective function in (29) is more computationally efficient than the original function in (28).

V. Simulation Results

Numerical results demonstrate the merits of the proposed cognitive scheme. We choose an RSF radar where the central frequency \( f_c = 10 \) GHz, the synthetic bandwidth \( B = 40 \) MHz, pulse duration \( T_p = 0.1 \) µS. We assume that \( \Delta f \to 0; \) thus, the codes are continuous parameters. The impacts of the discretization of the codes are discussed in Subsection V-E. The width of a coarse range bin is \( cT_p/2 \). The resolution of HRRP is \( c/(2B) \). Then there are \( M = (cT_p/2)/(c/2B) = 4 \) high-range-resolution cells in a coarse range bin. Unless specifically noted, a coherent processing interval (CPI) consists of \( N = 20 \) radar pulses. We set the number of possible high-resolution range cells \( P = M \), and the number of possible Doppler cells \( Q = N \). The noise are Gaussian white noise and we define the normalized signal to noise ratio as \( \text{SNR}_i = |\gamma_i|^2/(N\sigma^2) \) with respect to \( i \)th target. Since we focus on applying cognitive idea to RSF radar in this paper, we simply assume that the sparsity level \( K \) is known a priori in the SP method. However, in some practical situations, \( K \) is unknown. For those cases, we could first estimate the number of targets by combining SP with some well-known order-selection criteria, e.g., the Akaike’s information criterion (AIC) \([36]\), the Kullback-Leibler information criterion (KIC) \([37]\), and the minimum
description length (MDL) criterion [38]. In all of the following examples, SP iterates no more than 50 times.

A. Performance of Optimization Criterion

We discuss whether a code sequence with a lower constrained CRB leads to a smaller MSE. There are $K = 4$ targets with uniform scattering coefficients $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 1$. The range and motion parameters of the targets are set such that the parameters of the high-resolution range are $p_1 = 3\Delta p$, $p_2 = 2\Delta p$, $p_3 = 3\Delta p$, $p_4 = 0$, and Doppler parameters are $q_1 = 7\Delta q$, $q_2 = 13\Delta q$, $q_3 = 14\Delta q$, $q_4 = 15\Delta q$; see (6) for the definitions of $p$ and $q$. The normalized SNR of each target is SNR= 20 dB. We randomly generate 100 code sequences, $c_1, c_2, \ldots, c_{100}$. All of the codes are independently and uniformly distributed between 0 and 1. For each sequence $c_i$, we calculate the normalized, constrained CRB $LB(c_i) = \text{tr}((A^HA)^{-1})$ in (14) and estimate the sparse vector $x$ with SP in 1000 dependent Monte-Carlo trials. Then, we calculate MSE($c_i$), the MSE of the estimates with code sequence $c_i$, and plot MSE($c_i$) versus LB($c_i$) in Fig. 7.

![Fig. 7. MSE of x estimates versus the normalized, constrained CRB LB(e). Each blue circle in the figure represents a code sequence e.](image)

As shown in Fig. 7, a code with lower constrained CRB most likely leads to a lower MSE, so lowering the constrained CRB could be a valid strategy to reduce the recovery error. As discussed in Section IV, the matrix $A$ with lower CRB indicates a more informative sensing system, or indicates that interferences among target returns are relieved from the perspective of radar system.
B. Approximation of the Original Objective Function

In Subsection IV-B we approximate the original objective $\text{tr } ((A^H A)^{-1})$ in (14) with $\text{tr } (A^H A A^H A)$ in (15) for efficient computation of the gradient vector. In this subsection, numerical experiments are presented to demonstrate that the approximation is reasonable in the scenario of RSF radar. We randomly generate 2000 code sequences $c^i$ with an uniform distribution between 0 and 1, $i = 1, 2, \ldots, 2000$. The target scene is the same as that described in Subsection V-A. Then, we compare values of two objective functions $L_B(c^i) = \text{tr } ((A^H A)^{-1})$ and $L_B^2(c^i) = \text{tr } (A^H A A^H A)$ with respect to $c^i$. The relationship of the two functions is shown in Fig. 8. We could see that the approximation $L_B^2(c^i)$ has a trend similar to the original function $L_B(c^i)$. The original objective trends to descend when $L_B^2(c^i)$ is reduced. Note that the lower limits of both functions are $K = 4$.

![Fig. 8. LB(c^i) = tr ((A^H A)^{-1}) versus LB^2(c^i) = tr (A^H A A^H A). The plot contains 2000 black dots, each of which represents a code sequence c^i.](image)

C. Convergence of the Steepest Descent Method

In this subsection, we discuss the convergence of the proposed code design method in Subsection IV-C. The target scheme is the same as that in Subsection V-A. The initial code sequences are randomly created, each of which obeys an uniform distribution between 0 and 1. Then, these code sequences are optimized under the objective function (26) with the steepest descent method, which iterates no more than 100 times. The small constant $\delta$ in (24) equals 0.1. We calculate the
mean of values of the objective function \( \text{LB}_2(c) = \text{tr} \left( \mathbf{A}^H \mathbf{A} \mathbf{A}^H \mathbf{A} \right) \) versus the iteration counter \( l \) in the steepest descent method. The plot are presented in Fig. 9. The results show that the devised algorithm rapidly reduces values of \( \text{LB}_2(c) \) and closely converges to \( K = 4 \), which is the lower limit of \( \text{tr} \left( \mathbf{A}^H \mathbf{A} \mathbf{A}^H \mathbf{A} \right) \). The time complexity of the algorithm is also tested on a personal computer with Intel® Core™ 2 Duo CPU 3 GHz, 4 GB RAM. The codes are run by MATLAB® 2012. It takes 8.5 seconds to perform 100 Monte-Carlo trials with 100 iterations in each trail.

![Plot](image)

Fig. 9. The mean of \( \text{tr} \left( \mathbf{A}^H \mathbf{A} \mathbf{A}^H \mathbf{A} \right) \) versus the iteration counter \( l \) of the steepest descent method. The curve is obtained via averaging the results of 100 Monte-Carlo trials.

**D. Performance of the Batch-oriented Cognitive Scheme**

In this subsection, simulations are executed to test the reconstruction performance with the cognitive mechanism that implements the steepest descent method in Subsection IV-C. We consider two successive CPIs to demonstrate the merits of the cognitive scheme. In the first CPI, a predefined pseudo-random code sequence \( c^{(0)} \) is applied, and we obtain the estimate \( \hat{x}^{(0)} \) and the corresponding support set \( \Lambda^{(0)} = \text{supp} \left( \hat{x}^{(0)} \right) \). In the second CPI, we simulate the predefined mode and the adaptive mode, respectively, and compare the corresponding results. In the predefined mode, \( c^{(0)} \) is used again. In the latter mode, optimal code sequences are used, which are calculated via the steepest descend method with \( \Lambda^{(0)} \). The target scene is the same as that in Subsection V-A. The variance of the Gaussian white noise is varied such that
E. Performance of the Sequential Cognitive Scheme

In this subsection, we consider the cognitive mechanism that uses the sequential code design algorithms in Subsection IV-D. The target scene is the same as that in Subsection V-A. The initial $N = 20$ modulation codes $c^{(0)}$ are randomly drawn from an uniform distribution over $[0, 1]$. We test the adaptive modes and the random mode. In the adaptive modes, we sequentially designed the ensuing 20 codes with the methods in Subsection IV-D, while in the random mode the ensuing codes are drawn randomly as comparison. Once a new code is determined, the recovery errors are calculated. As discussed in Subsection IV-D, two methods are applicable for the adaptive modes. We denote the code design processes with (31) and (32) as 'Adaptive Mode 1' and 'Adaptive Mode 2', respectively.

The reconstruction errors and fractions of exactly recovered support sets are depicted in Fig. 11, where the noise variances are set as $\sigma^2 = 0$ dB or $\sigma^2 = 5$ dB. Two adaptive modes have similar recovery errors, which are much smaller than those of the random mode. Fractions
of exactly recovered support sets of both adaptive modes are close to each other, which are much higher than the random mode. The time complexity of these three modes are tested on a personal computer with Intel® Core™ 2 Duo CPU 3 GHz, 4 GB RAM. The codes are run by MATLAB® 2012. For 1000 Monte-Carlo trials, the consumed time of 'Adaptive Mode 1' and 'Adaptive Mode 2' are 37.0 and 34.5 seconds, respectively, which are very close to 33.7 seconds, the counterpart of 'Random Mode'. This indicates that the code optimum processes consume minor computational efforts. Note that Adaptive Mode 2 has lower computation load than Adaptive Mode 1, and it may be more suitable for some real-time applications.

![Graphs](image_url)

(a) MSEs, $\sigma^2 = 0$ dB  
(b) Fractions of exactly recovered support sets, $\sigma^2 = 0$ dB  
(c) MSEs, $\sigma^2 = 5$ dB  
(d) Fractions of exactly recovered support sets, $\sigma^2 = 5$ dB

Fig. 11. The MSEs and the fractions of exactly recovered support sets with random and adaptive code sequences versus number of measurements. In (a) and (b), the noise variance $\sigma^2 = 0$ dB, and in (c) and (d) $\sigma^2 = 5$ dB. In 'Random Mode', the codes are randomly generated. In 'Adaptive Mode 1' and 'Adaptive Mode 2', the codes are sequentially optimized by (31) and (32), respectively. Each circle or asterisk in the plots is obtained from 1000 independent Monte-Carlo trials.
Then we consider the influence of the frequency step size $\Delta f$. The performances of the three modes are evaluated with respect to different $\Delta f$. There are 4 targets, of which the ranges and velocities are randomly generated. The scattering intensities of the targets are all one. The initial $N = 20$ codes are randomly generated. The MSEs are calculated after the ensuing 5 pulses are transmitted. The noise variance $\sigma^2 = 0$ dB. The results are shown in Fig. 12, and the adaptive modes outperform the nonadaptive mode for all tested $\Delta f$.

![Fig. 12](image)

Fig. 12. The MSEs with random and adaptive code sequences versus the frequency step size $\Delta f$. The results are obtained from 20000 independent Monte-Carlo trials.

We also discuss impacts of the number of targets $K$ on the performance of the cognitive mechanism. The target scene and the initial $N = 20$ codes are randomly generated. The scattering intensities of the targets are all one. We calculate the recovery errors after the ensuing 5 codes are designed. The noise variance $\sigma^2 = -5$ dB. The results are shown in Fig. 13, and we can see that the adaptive modes produce lower recovery errors for all $K$ that appear in Fig. 13.

VI. CONCLUSION

In this paper, we develop a novel notion of cognitive random stepped frequency radar that employs compressed sensing algorithm to reconstruct the target scene. We propose a new criterion and several effective methods to adaptively optimize the carrier frequencies of the radar pulses. With the information about the observed targets, the proposed cognitive mechanism significantly eliminates the interference between target returns and improves the sensing performance of radar.
Numerical simulations are performed to demonstrate that the devised cognitive system reduces the recovery errors of the targets.

APPENDIX

We derive the gradient of $f(z) = \text{tr}(A^H A A^H A)$ with respect to $z$. Denote $a_{ml}$ and $b_{ml}$ as the $m$th row, $l$th column element of $A$ and $B = A^H A \in \mathbb{C}^{\|A\| \times |A|}$, respectively. We have that

$$\frac{dc_n}{dz_n} = \frac{1 + 2\delta}{\pi(z_n^2 + 1)};$$

$$\frac{\partial a_{ml}}{\partial c_n} = \begin{cases} 0, m \neq n, \\ a_{nl}(j\rho_l + j\gamma_mB/f_c), m = n, \end{cases}$$

$$\frac{\partial b_{ml}}{\partial c_n} = \frac{\partial}{\partial c_n} \sum_i a^*_{im}a_{il} = \frac{\partial a^*_{nm}a_{nl}}{\partial c_n} = \left(\frac{\partial a_{nm}}{\partial c_n}\right)^*a_{nl} + a^*_{nm}\frac{\partial a_{nl}}{\partial c_n},$$

$$\frac{\partial \text{tr}(B^H B)}{\partial c_n} = \sum_{m,l} \frac{\partial b^*_{ml}b_{ml}}{\partial c_n} = 2\text{Re} \left[ \sum_{m,l} b^*_{ml}\frac{\partial b_{ml}}{\partial c_n} \right].$$
With the substitution of (33 - 36), we get
\[
\frac{\partial f(z)}{\partial z_n} = \frac{\partial \text{tr} \left( \mathbf{B}^H \mathbf{B} \right)}{\partial c_n} \cdot \frac{dc_n}{dz_n}
\]
\[
= 2 \text{Re} \left[ \sum_{m,l} d_{mn} b_{ml} a_{nl}^* + a_{nm} b_{nl} a_{ml}^* \right],
\]
where
\[
d_{nl} = \frac{da_{nl}}{dz_n} = \frac{dc_n}{dz_n} \frac{dc_n}{dz_n} = a_{nl} \left( j p_l + j q_m B \right) \frac{1 + 2 \delta}{\pi (z_n^2 + 1)}
\]
denotes the \( n \)th row, \( l \)th column element of matrix \( D \). Rewrite (37) in matrix form,
\[
\frac{\partial f(z)}{\partial z} = 2 \text{Re} \left[ \text{diag} \left( \mathbf{DBA}^H + \mathbf{ABD}^H \right) \right] = 4 \text{Re} \left[ \text{diag} \left( \mathbf{DBA}^H \right) \right].
\]

**ACKNOWLEDGMENT**

This work was supported in part by the National Natural Science Foundation of China (No. 60901057 and No. 61201356) and in part by the National Basic Research Program of China (973 Program, No. 2010CB731901).

The authors would like to thank Prof. Arye Nehorai for his insightful comments on various versions of this manuscript, anonymous reviewers for their important suggestions, and also Mr. Yuanxin Li who contributed some suggestions on writing Subsection IV-B.

**REFERENCES**

[1] S. R. J. Axelsson, “Analysis of Random Step Frequency Radar and Comparison With Experiments,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 4, pp. 890–904, 2007.

[2] D. Wehner, *High-Resolution Radar*. Artech House, Incorporated, 1995.

[3] Y. Hua, F. Baqai, Y. Zhu, and D. Heilbronn, “Imaging of point scatterers from step-frequency ISAR data,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 29, no. 1, pp. 195–205, Jan 1993.

[4] Y. Liu, H. Meng, G. Li, and X. Wang, “Velocity Estimation and Range Shift Compensation for High Range Resolution Profiling in Stepped-Frequency Radar,” *Geoscience and Remote Sensing Letters, IEEE*, vol. 7, no. 4, pp. 791–795, Oct. 2010.
[5] ——, “Range-velocity estimation of multiple targets in randomised stepped-frequency radar,” *Electronics Letters*, vol. 44, no. 17, pp. 1032–1034, 2008.

[6] T. Huang, Y. Liu, G. Li, and X. Wang, “Randomized Stepped Frequency ISAR Imaging,” in *Radar Conference (RADAR), 2012 IEEE*, may 2012, pp. 553–557.

[7] 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,” *Electronics Letters*, vol. 40, no. 22, pp. 1447–1448, 2004.

[8] ——, “Doppler aliasing reduction in SAR imagery using stepped-frequency waveforms,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 163–175, 2007.

[9] Z. Liu and S. Zhang, “Motion Parameters Estimation for Hopped-frequency Radar,” *Journal of Electronics*, vol. 22, no. 4, pp. 591–596, 2000.

[10] S. Gogineni and A. Nehorai, “Target Estimation Using Sparse Modeling for Distributed MIMO Radar,” *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5315–5325, 2011.

[11] A. Gurbuz, V. Cevher, and J. Mcclellan, “Bearing Estimation via Spatial Sparsity using Compressive Sensing,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 48, no. 2, pp. 1358–1369, april 2012.

[12] Y. Yu, A. Petropulu, and H. Poor, “CSSF MIMO RADAR: Compressive-Sensing and Step-Frequency Based MIMO Radar,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 48, no. 2, pp. 1490–1504, april 2012.

[13] S. Gogineni and A. Nehorai, “Frequency-Hopping Code Design for MIMO Radar Estimation Using Sparse Modeling,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 6, pp. 3022–3035, June 2012.

[14] S. Haykin, “Cognitive radar: a way of the future,” *Signal Processing Magazine, IEEE*, vol. 23, no. 1, pp. 30 – 40, Jan. 2006.

[15] E. Conte, A. De Maio, A. Farina, and G. Foglia, “Design and analysis of a knowledge-aided radar detector for Doppler processing,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 42, no. 3, pp. 1058–1079, July 2006.

[16] M. Inggs, “Passive Coherent Location as Cognitive Radar,” *Aerospace and Electronic Systems Magazine, IEEE*, vol. 25, no. 5, pp. 12 –17, May 2010.

[17] Y. Wei, H. Meng, Y. Liu, and X. Wang, “Radar Phase-Modulated Waveform Design for
Extended Target Detection,” *Tsinghua Science & Technology*, vol. 16, no. 4, pp. 364–370, 2011.

[18] J. Zhang, D. Zhu, and G. Zhang, “Adaptive Compressed Sensing Radar Oriented Toward Cognitive Detection in Dynamic Sparse Target Scene,” *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1718–1729, 2012.

[19] S. Sen, G. Tang, and A. Nehorai, “Multiobjective Optimization of OFDM Radar Waveform for Target Detection,” *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 639–652, 2011.

[20] S. Gogineni and A. Nehorai, “Frequency-hopping code design for colocated MIMO radar using sparse modeling,” in *Proc. 6th Int. Waveform Diversity and Design (WDD) Conf.*, Kauai, Hawaii, USA, Jan. 2012, pp. 54–58.

[21] K. Han and A. Nehorai, “Joint Frequency-hopping Waveform Design for MIMO Radar Estimation Using Game Theory,” in *Radar Conference (RADAR), 2013 IEEE*, 2013.

[22] 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,” *Radar, Sonar Navigation, IET*, vol. 3, no. 5, pp. 512–520, Oct. 2009.

[23] Data sheet of AD9910 (AGSPS, 14-Bit, 3.3V CMOS, Direct Digital Synthesizer). Analog Devices. [Online]. Available: [http://www.analog.com/static/imported-files/data_sheets/AD9910.pdf](http://www.analog.com/static/imported-files/data_sheets/AD9910.pdf)

[24] S. R. Axelsson, “Suppression of Noise Floor and Dominant Reflectors in Random Noise Radar,” in *Radar Symposium, 2006. IRS 2006. International*, May 2006, pp. 1–4.

[25] W. Dai and O. Milenkovic, “Subspace Pursuit for Compressive Sensing Signal Reconstruction,” *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[26] T. Huang, Y. Liu, and X. Wang, “Adaptive Subspace Pursuit and Its Application in Motion Compensation for Step Frequency Radar,” in *1st International Workshop on Compressed Sensing applied to Radar (CoSeRa)*, 2012.

[27] S. Chen, D. Donoho, and M. Saunders, “Atomic decomposition by basis pursuit,” *SIAM review*, pp. 129–159, 2001.

[28] E. Candes and T. Tao, “The Dantzig selector: Statistical estimation when p is much larger than n,” *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

[29] M. A. Davenport and M. B. Wakin, “Analysis of Orthogonal Matching Pursuit Using
the Restricted Isometry Property,” *IEEE Transactions on Information Theory*, vol. 56, pp. 4395–4401, 2010.

[30] D. Needell and R. Vershynin, “Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit,” *Foundations of computational mathematics*, vol. 9, no. 3, pp. 317–334, 2009.

[31] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, ser. Prentice Hall Signal Processing Series. Prentice-Hall PTR, 1993, no. v. 1.

[32] Z. Ben-Haim and Y. Eldar, “The Cramer-Rao Bound for Estimating a Sparse Parameter Vector,” *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3384 –3389, june 2010.

[33] B. Babadi, N. Kalouptsidis, and V. Tarokh, “Asymptotic Achievability of the Cramer-Rao Bound for Noisy Compressive Sampling,” *Signal Processing, IEEE Transactions on*, vol. 57, no. 3, pp. 1233 –1236, march 2009.

[34] R. Niazadeh, M. Babaie-Zadeh, and C. Jutten, “On the Achievability of Cramer–Rao Bound in Noisy Compressed Sensing,” *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 518–526, 2012.

[35] P. Gill, W. Murray, and M. Wright, *Practical optimization*. Academic Press, 1981.

[36] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[37] J. E. Cavanaugh, “A large-sample model selection criterion based on Kullback’s symmetric divergence,” *Statistics & Probability Letters*, vol. 42, no. 4, pp. 333 – 343, 1999.

[38] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, no. 5, pp. 465 – 471, 1978.
Tianyao Huang received the B.S. degree in electronic and information engineering from Harbin Institute of Technology, Harbin, Heilongjiang, China, in 2009.

He is currently pursuing his Ph.D. degree in electronic engineering at Tsinghua University, Beijing, China. From 2012 to 2013, he was a visiting student with Dr. Wei Dai at Imperial College London, London, UK. His research interests are radar signal processing, compressed sensing and electronic system design.

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

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

Huadong Meng (S’01-M’04) received the B.Eng. degree and Ph.D. degree in electronic engineering, both from Tsinghua University, Beijing, China, in 1999 and 2004 respectively.

In 2004, he joined the Faculty of Tsinghua University, where he is currently an Associate Professor in the Department of Electronic Engineering. He is a member of the Technical Committee of the 2013 IET International Radar Conference. His current research interests include statistical signal processing, target tracking, sparse signal processing, cognitive radar, and spectrum sensing in cognitive radio networks.

Xiqin Wang received his 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 from 2000 to 2003.

He is currently a professor with the Department of Electronic Engineering of Tsinghua University, and also chaired the department from 2006-2012. 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.