

On the Compressive Sensing Systems (Part I)

Prepared by:

Mingbo Niu, Il-Min Kim - Department of Electrical and Computer Engineering, Queen's University

Francois Chan - Department of Electrical and Computer Engineering, RMC

Contractor name and address:

Francois Chan, RMC

13 General Crerar Crescent, Kingston, ON, K7K 7B4

Contract number: 2009-0302-SLA

Contract Scientific Authority's name, title, and phone number:

Dr. Sreeraman Rajan, Senior Defence Scientist 613-991-4138

The scientific or technical validity of this Contract Report is entirely the responsibility of the Contractor and the contents do not necessarily have the approval or endorsement of the Department of National Defence of Canada.

Contract Report

DRDC-RDDC-2015-C025

February 2015

ON THE COMPRESSIVE SENSING SYSTEMS (PART I)

Prepared by ¹Mingbo Niu, ¹Il-Min Kim, and ²Francois Chan

1. Department of Electrical and Computer Engineering, Queen's University

2. Department of Electrical and Computer Engineering, RMC

Contractor name and address:

Francois Chan, RMC,

13 General Crerar Crescent, Kingston, ON, K7K 7B4

Contract number: 2009-0302-SLA

Contract Scientific Authority's name, title, and phone number:

Dr. Sreeraman Rajan, Senior Defence Scientist

613-991-4138

The scientific or technical validity of this Contract Report is entirely the responsibility of the Contractor and the contents do not necessarily have the approval or endorsement of the Department of National Defence of Canada.

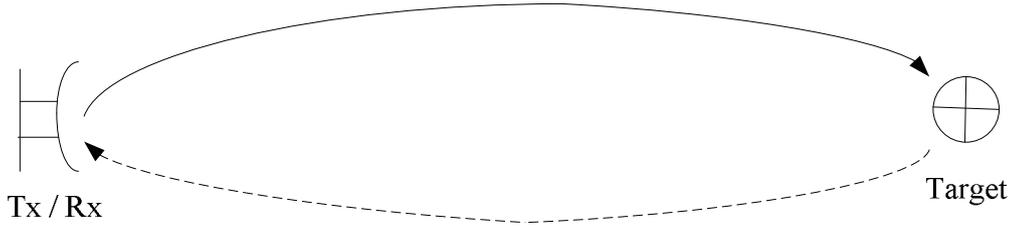


Fig. 1. A traditional active radar system with one target.

I. INTRODUCTION

Compressive sensing (CS) has received considerable attention. CS has shown great potential in diverse fields, for example, high-resolution image processing, wireless communication, and radar systems [1]–[6]. The central goal of CS is to reconstruct a signal using very few measurements [7]. This report starts with a brief review of target detection, communication and CS theory for different radar and wireless communication signals. And then the report focuses on different CS applications and perfect signal reconstruction conditions.

II. ADAPTIVE COMPRESSED SENSING RADAR SETTINGS

A. Preliminary on Traditional Radar

We begin with an introduction to a traditional radar system. A simplified radar system is given in Fig. 1. Such a system is one-dimensional, monostatic, single-pulse, and active radar. Suppose a target located at range x is traveling with constant velocity v and has reflection coefficient s_{xv} . After the transmission of signal $f(t)$, the receiver observes the reflected signal as

$$r(t) = s_{xv}f(t - \tau_x) \exp(2\pi j\omega_{xv}t) \quad (1)$$

where τ_x is the round trip time of the signal, $j = \sqrt{-1}$, and ω_{xv} is the Doppler shift. The target information (x, v) can be readily obtained from (τ_x, ω_{xv}) , the pair of (*time delay*, *Doppler shift*), of $f(\cdot)$. Here, an ideal case is considered without noise.

Using a matched filter, the reflected signal r is correlated with a time-frequency shifted version of the transmitted signal through

$$|A_{rf}(\tau, \omega)| = \left| \int_R r(t) f^*(t - \tau) \exp(2\pi j \omega t) \right| = |s_{xv} A_f(\tau - \tau_x, \omega - \omega_v)| \quad (2)$$

where $A_{rf}(\cdot, \cdot)$ is so called the cross-ambiguity function of two functions f and r , and the superscript $*$ denotes the conjugate operation. From (2), we can see the ambiguity surface of f centered at the target's location and scaled by its reflection coefficient $|s_{xv}|$. It is straightforward to extend (2) to include multiple targets. However, when two or more targets are too close to each other and have overlapping ambiguity functions, there are some uncertainties on the exact locations of targets. Therefore, the range-velocity resolution between targets of classical radar is limited by the radar uncertainty principle.

B. Fundamentals on CS and CS-Based Radar (CSR)

Under appropriate conditions, CSR can *beat* the traditional radar. We now consider K targets with unknown range-velocities and corresponding reflection coefficients in an $N' \times N$ grid plane. If the number of targets $K \ll N'N$ and the transmitted signals are incoherent, we can use the CS technique to recover such a sparse target scene.

A CSR has the following features:

- 1) Eliminating the need of matched filter at the receiver;
- 2) Requiring low sampling bandwidth for low information rate (other than the high Nyquist rate);
- 3) Providing higher resolution compared to the traditional radar.

Considering a target scene in an $L \times M$ range-Doppler plane in Fig. 2, we define the time delay and Doppler shift matrices, respectively, as

$$\mathbf{T}_{L \times N}^l = \begin{pmatrix} \mathbf{0}_{l \times N} \\ \mathbf{I}_{N \times N} \\ \mathbf{0}_{(L-N-l) \times N} \end{pmatrix} \quad (3)$$

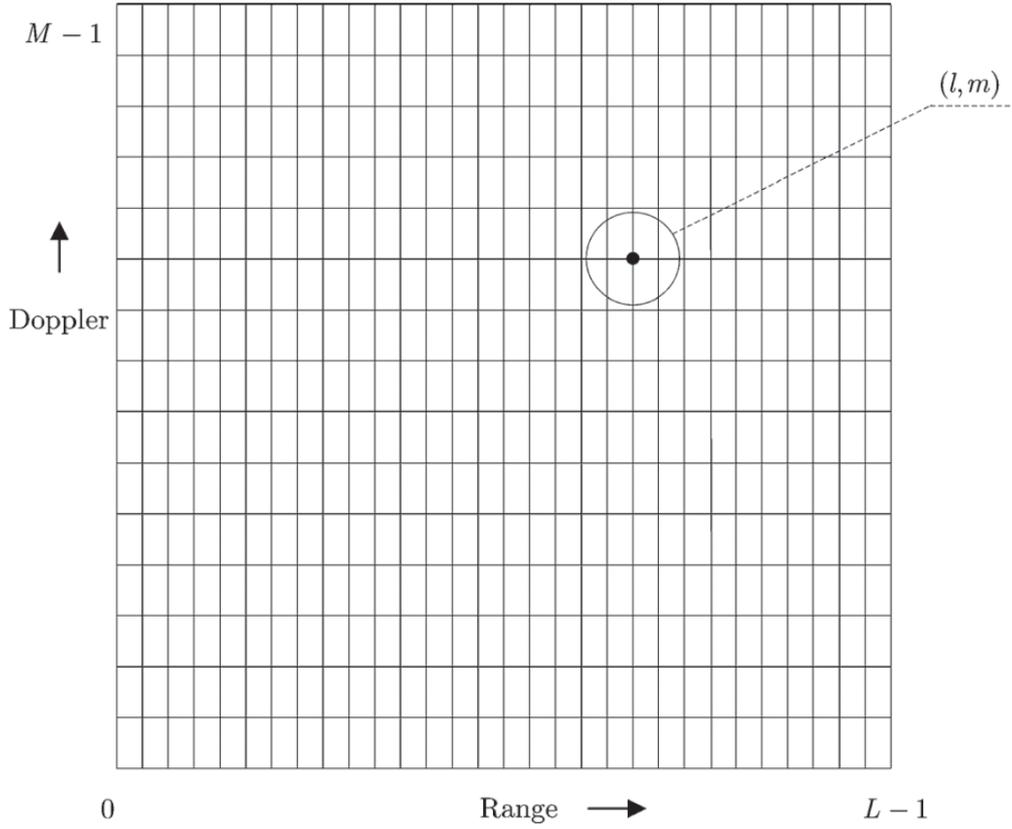


Fig. 2. The range-Doppler plane of the target scene that CSR measures is discretized into an $L \times M$ grid. The target located on the l th range and the m th Doppler bin is labeled.

and

$$\mathbf{F}_{L \times L}^m = \begin{pmatrix} \omega_M^0 & 0 & \cdots & 0 \\ 0 & \omega_M^1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \omega_M^{L-1} \end{pmatrix}^m \quad (4)$$

where L and M denote the numbers of range and Doppler bins that CSR measures, respectively, N is the length of transmission waveform vector, and $\omega_M = \exp(2\pi j/M)$ is the root of unity.

A basis dictionary $\mathbf{H} = (|\mathbf{h}_{0,0}\rangle | \mathbf{h}_{0,1}\rangle \cdots | \mathbf{h}_{L-1,M-1}\rangle)$ can be created by concatenating the received signals where $\mathbf{h}_{l,m} = \mathbf{F}^m \cdot \mathbf{T}^l \cdot \mathbf{s}$. Here, \mathbf{s} denotes the transmission signal of length N . \mathbf{H} contains all the possible signal reflected from the target in any grid of the range-Doppler plane and is named

as *sparse representation matrix* in the CS theory.

Define $\theta_{l,m}$ as the complex scattering coefficient in the (l, m) th grid of the target scene, and the corresponding coefficient vector is denoted as

$$\left(\theta_{0,0} \quad \theta_{0,1} \quad \cdots \quad \theta_{L-1,M-1} \right)^T \quad (5)$$

where the superscript T denotes the transpose operation. The task of a CSR is to discover/identify the coefficients in the range-Doppler plane. The received signal reflected from the target scene can be expressed as

$$\mathbf{r} = \sum_{l=0}^{L-1} \sum_{m=0}^{M-1} \theta_{l,m} \mathbf{h}_{l,m} + \mathbf{n} = \mathbf{H}\boldsymbol{\theta} + \mathbf{n} \quad (6)$$

where \mathbf{n} is the additive Gaussian white noise (AWGN) vector of length L . The compressed signal \mathbf{y} can be written as

$$\mathbf{y} = \Phi \mathbf{r} = \underbrace{\Phi \mathbf{H}}_{\Upsilon} \boldsymbol{\theta} + \Phi \mathbf{n} \quad (7)$$

where Φ is the $P \times L$ measurement matrix.

In adaptive CSR, both transmission waveforms \mathbf{s} and measurement matrix Φ are treated as variables which will be optimized and updated based on $\tilde{\boldsymbol{\theta}}$, the estimation of $\boldsymbol{\theta}$. Note that $\Upsilon = \Phi \mathbf{H}$ is the equivalent dictionary for this proposed adaptive CSR. Complex Gram matrix \mathbf{G} is applied to represent the mutual coherence and it is defined as [8]

$$\mathbf{G} = \Upsilon^H \Upsilon \quad (8)$$

where the superscript H denotes the conjugate transpose operation.

Variables \mathbf{s} and Φ are optimized to make the resulting Gram matrix approximate a diagonal matrix (the ideal case), i.e., $\mathbf{G} = \text{Diag}(g_{0,0}, g_{0,1}, \cdots, g_{LM-1,LM-1})$. Algorithms are proposed for the optimization by fixing \mathbf{s} and/or Φ . With the changing of Υ , the target is to make the realistic Gram matrix $\tilde{\mathbf{G}}$ as close to the ideal diagonal \mathbf{G} as possible. This can increase the resolution of CSR significantly. In Fig. 3 on the next page, a work flow chart is given where $\check{\mathbf{G}}$ is the estimation of $\tilde{\mathbf{G}}$.

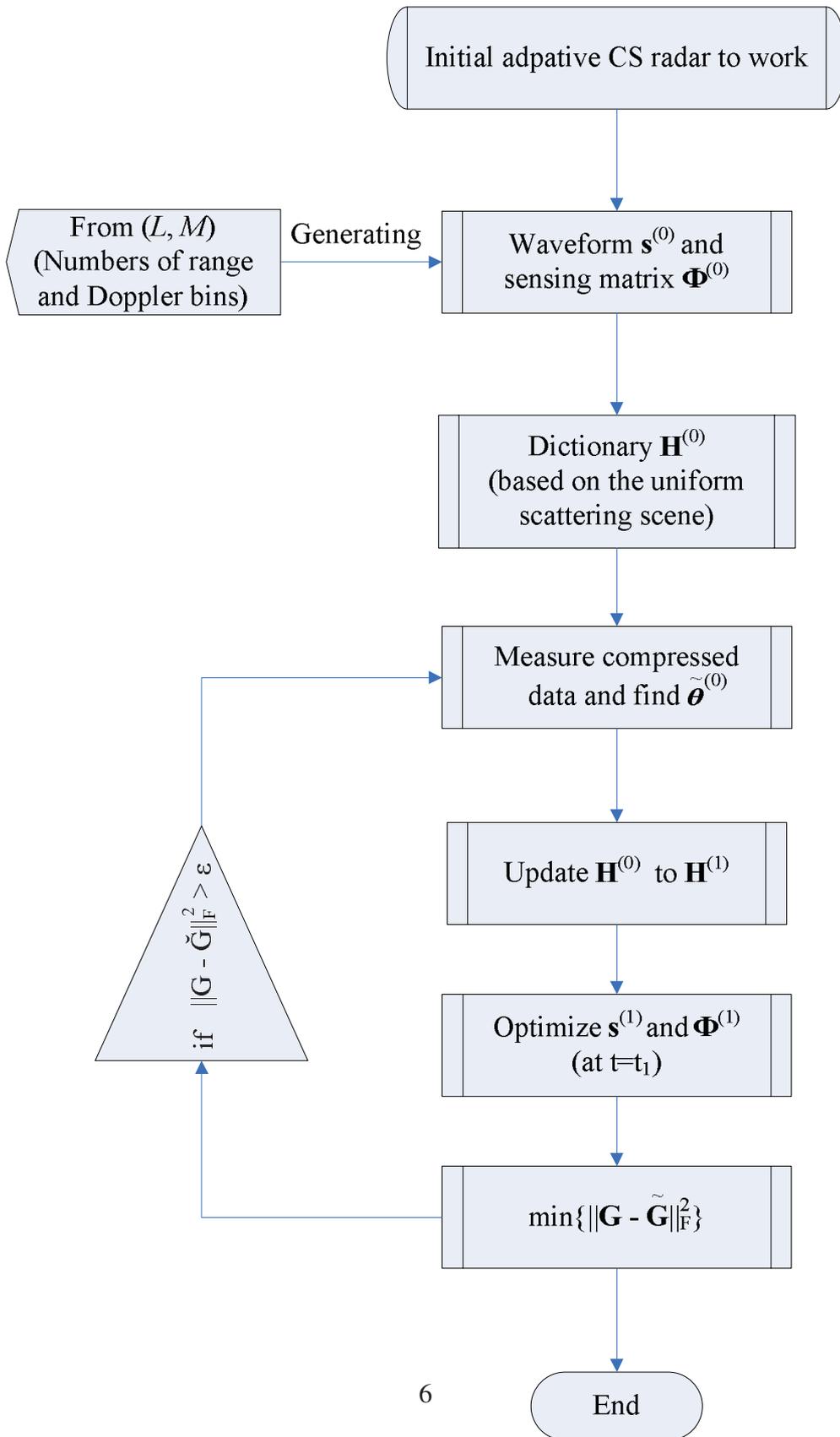


Fig. 3. Work flow chart for this adaptive CSR system.

III. ANALOG-TO-INFORMATION IN CS

In this section, we first generalize CS system for continues-time (CT) signals, and then we present some more possible applications in the field of CS.

A. Traditional Form of CS

We describe the traditional form of a CS system in discrete-time (DT) domain in this subsection. CS deals with the problem of acquiring an $N \times 1$ discrete time signal vector \mathbf{f} . To proceed, \mathbf{f} should be K -sparse in some $N \times N$ sparsity basis matrix Ψ in which each column is a basis vector ψ_i , i.e., $\Psi = [\psi_1, \dots, \psi_i, \dots, \psi_N]$. If we represent

$$\mathbf{f} = \Psi \mathbf{x} \quad (9)$$

where \mathbf{f} is K -sparse, meaning that only $K \ll N$ expansion coefficients in \mathbf{x} are non-zero. With Φ denoting an $M \times N$ sensing matrix, we have

$$\mathbf{y} = \Phi \mathbf{f} \quad (10)$$

which represents the N -dimensional input signal encoded into an M -dimensional set of measurements and $M \ll N$. Using (9) and (10), \mathbf{y} can be rewritten as

$$\mathbf{y} = \Phi \mathbf{f} = \underbrace{\Phi \Psi}_{\mathbf{U} = \Phi \Psi} \mathbf{x} \quad (11)$$

where \mathbf{y} is an $M \times 1$ vector, Φ is an $M \times N$ matrix, Ψ is a $N \times N$ matrix, and \mathbf{x} is a $N \times 1$ vector.

In the present form, CS is only applicable to DT signals. Below we extend the framework to CT signals. In the literature, there exist a few kinds such conversion architectures. We will present an architecture of such conversion from CT signal $f(t)$ to DT signal \mathbf{f} in the next section.

B. Analog-to-information: CS from CT to DT

We now generalized CS theory from the DT to CT domain. To develop CT domain CS, one will need to define analog signal models for sparse signals and then construct an analog system which is CS-compatible.

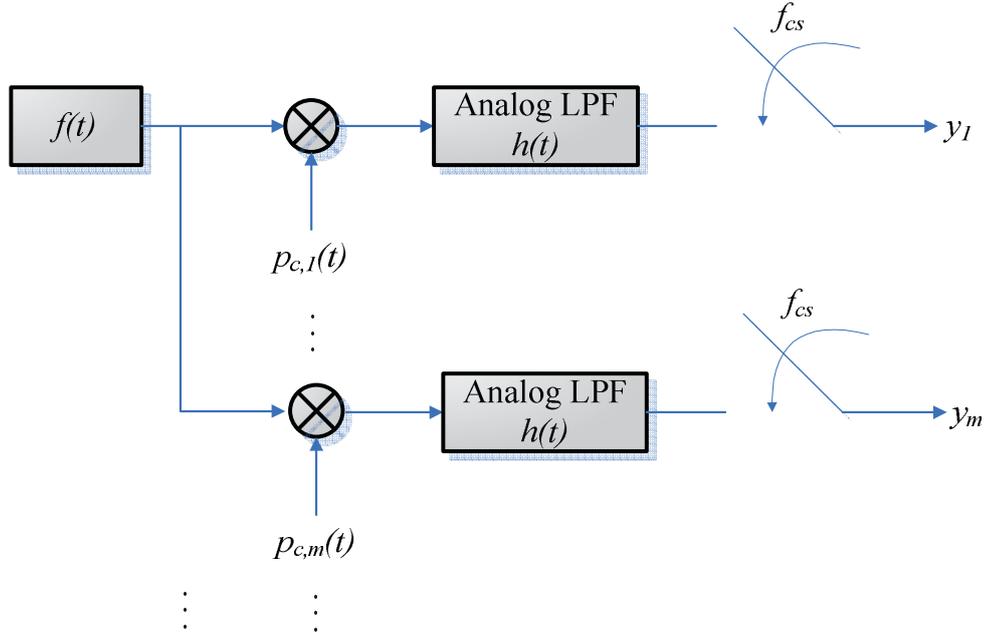


Fig. 4. An analog-to-information CS architecture.

Assuming the analog signal $f(t)$ can be represented through a finite number of parameters x_i in some continuous basis $\psi_i(t)$, $i = 1, \dots, N$, then $f(t)$ can be decomposed as

$$f(t) = \sum_{i=1}^N x_i \psi_i(t), \quad x_i \in \mathbf{R}. \quad (12)$$

In the case that there are a small number of nonzero entries in $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$, again we can say that the signal $f(t)$ is sparse.

The next question: how one can convert CT signal $f(t)$ to DT measurements y_m ($m = 1, \dots, M$)?

Figure 4 gives an architecture to solve above issue. In Fig. 4, $P_{c,m}(t)$ denotes the m th distinct chipping sequence which is a pseudo-random sequence of ± 1 s. The alternating frequency of the chipping sequence must be between values at or faster than the Nyquist frequency of the input signal. That is, f_{cs} is a much lower (compared to the Nyquist rate $f_s = 2f_{\max}$) sampling rate (e.g., $f_{cs} = f_s/N$).

From Fig. 4, y_m can be obtained as a result of modulation, convolution and sampling (at rate of

f_m) operations. Therefore, we can write y_m in the following equation

$$y_m = \int_{-\infty}^{\infty} f(\tau) p_c(\tau) h(t - \tau) d\tau \Big|_{t=f_{cs}}. \quad (13)$$

Substituting the analog signal $f(t)$ into (12), we have

$$y_m = \sum_{i=1}^N x_i \int_{-\infty}^{\infty} \psi_i(\tau) p_c(\tau) h(f_{cs} - \tau) d\tau. \quad (14)$$

We define the $M \times N$ matrix $\mathbf{U} = \mathbf{\Phi}\mathbf{\Psi}$. The matrix \mathbf{U} can be determined from the expression of y_m . It is clear that we can separate out the expression for each element $u_{m,i} \in \mathbf{U}$ for the m th row and i th column as

$$u_{m,i} = \int_{-\infty}^{\infty} \psi_i(\tau) p_{c,m}(\tau) h(f_{cs} - \tau) d\tau. \quad (15)$$

Through the analog-to-information process, rather than measuring and sampling $f(t)$ directly, we acquire $M \ll N$ linear projections by $\mathbf{y} = \mathbf{\Phi}\mathbf{f} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x}$. We comment that the minimal number M of measurements required to recover the signal is bounded by

$$M \geq C \mu^2(\mathbf{\Phi}, \mathbf{\Psi}) K \log_2 N \quad (16)$$

where C denotes a small known constant (empirically ~ 2 to 2.5), N is the dimension of the signal to be reconstructed, and $\mu(\cdot, \cdot)$ represents the measurement of the coherence between $\mathbf{\Phi}$ and $\mathbf{\Psi}$ and is defined as

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \max_{1 \leq i \leq N, 1 \leq k \leq M} |\langle \phi_k, \psi_i \rangle|. \quad (17)$$

Here, $\mu^2(\mathbf{\Phi}, \mathbf{\Psi})$ can range from 1 to N . The smaller $\mu(\mathbf{\Phi}, \mathbf{\Psi})$ is, the fewer number of measurements is required to reconstruct the signal. Note that K represents the sparsity of the signal, and M is proportional to the required number of measurements.

When \mathbf{y} and \mathbf{U} are known, we have a common and practical approach to find the sparse solution to the following convex optimization problem [3]:

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x} \in \mathbf{R}^n} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{U}\mathbf{x}. \quad (18)$$

where $\|\cdot\|_1$ denotes the l_1 -norm. The recovered signal is $\hat{\mathbf{f}} = \Psi \mathbf{x}^*$ with \mathbf{x}^* being the optimal solution to (18).

We comment that there are many different approaches in finding the solution to the undetermined CS system described by $\mathbf{y} = \mathbf{U}\mathbf{x} = \Phi\Psi\mathbf{x}$. One common algorithm to solve the minimization problem in (18) is basis pursuit (BP). Since this algorithm has been shown to have a tractable complexity [3], [15], it is widely used in practice.

C. Applications of CS

CS theory asserts that one can recover certain signals from far fewer samples or measurements than traditional methods at Nyquist rate. In this section, some possible applications of CS is briefly summarized.

A) **Data compression.** When the sparse basis Ψ is unknown or impractical to implement at data compression stage, one can apply a randomly designed Φ as a universal encoding strategy. Note that Ψ is required only at the recovery stage of \mathbf{f} . This encoding strategy is particularly useful for distributed data source coding such as sensor networks [16], [5].

B) **Channel coding.** When the data is sparsity, CS theory can be used for channel coding, e.g, error-correcting code (ECC). Such a coding system model can be described as [6]

$$\mathbf{y}_c = \mathbf{A}\mathbf{f}_c + \mathbf{e} \quad (19)$$

where \mathbf{f}_c is an input vector, \mathbf{y}_c is corrupted measurement, \mathbf{A} is a (coding) matrix, and \mathbf{e} denotes some unknown error. Based on the sparsity property, with proper designing of \mathbf{A} , one may decode \mathbf{f}_c from measurement \mathbf{y}_c .

C) **Inverse problem.** Related discussion is omitted in this report. More details can be found in [9] and [10].

D) **Data acquisition.** When high rate or huge amount of N DT samples is difficult, sub-Nyquist rate samples are desired in designing physical sampling devices [11].

E) **A/I**. It can be considered as a specific application in data acquisition. As described in Section III-B, **A/I** is to measure and compress high bandwidth CT signal (K -sparse) to DT signal (with a small number of measurements).

IV. THROUGH WALL IMAGING RADAR

A. Preliminary on Through-Wall Imaging

Urban radar has become a keen research topic in recent years because of the urgent needs of urban warfare, counterterrorism, calamity rescue, etc. Through-wall imaging (TWI) radar systems can provide high range and cross-range solutions by using ultra-wideband (UWB) signals and antenna arrays with large apertures, respectively.

To begin, we present a brief description of a conventional TWI radar system. We first pose some assumptions:

1. A target consists of (at least) one point of scatters, and the reflection coefficient of each scatter is a (complex) constant;
2. In the range of resolution, human bodies are replaced by point targets.

To model the TWI system, a stepped-frequency scheme is considered. Suppose the stepped signal $\in [f_L, f_H]$ has N narrow band signals with carrier frequencies

$$f_n = f_L + n\Delta f, \quad n = 1, \dots, N \quad (20)$$

where f_L is the starting frequency, f_H is the highest frequency, and $\Delta f = (f_H - f_L)/(N - 1)$ is the frequency step.

Figure 5 shows the scene with the wall for the point target $P(x_p, y_p)$ and the m -th antenna pair (T_x, R_x) . Suppose that the wall thickness d and the relative permittivity of the wall are known. Under the two assumptions and the stepped-frequency scheme, the received signal of frequency f_n at the receiver position of the m th antenna pair can be represented through complex envelope

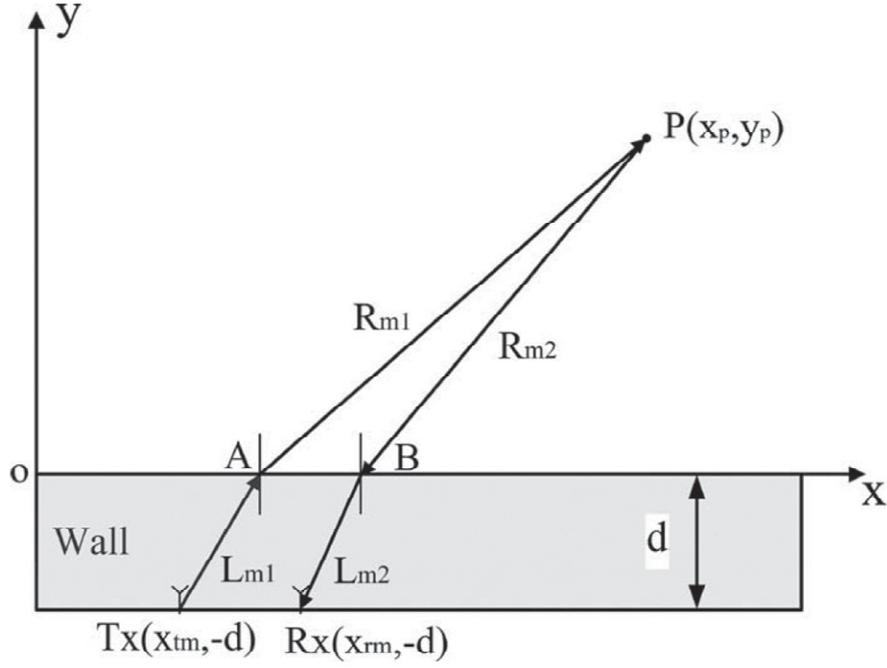


Fig. 5. A geometry scene with the wall for one target.

(after normalizing by the transmitted signal) as

$$z[m, n] = \sum_{p=0}^{P-1} \sigma_p \exp(-j2\pi f_n \tau_{pm}) \quad (21)$$

where σ_p is the reflection coefficient of the p th target, P is the number of targets in the scene, m and n represent the spatial index for $m = 0, \dots, M - 1$ and frequency index for $n = 0, \dots, N - 1$, respectively, and τ_{pm} is the round-trip propagation delay between the m -th antenna pair and the p -th target. The round-trip propagation delay can be calculated through

$$\tau_{pm} = \frac{R_{m,1} + R_{m,2}}{c} + \frac{L_{m,1} + L_{m,2}}{v} \quad (22)$$

where $R_{m,1}$ and $R_{m,2}$ are the respective distances between the p th target and the refraction points A and B, c denotes the velocity of light, and $v = c/\sqrt{\epsilon_r}$ is the velocity inside the wall with ϵ_r denoting the relative permittivity of the wall.

We comment that the coordinates of the refraction points A and B can be obtained as

$$X_{m,1} = x_{tm} + \frac{d(x_p - x_{tm})}{\sqrt{\varepsilon_r}(y_p + d)} \quad (23)$$

and

$$X_{m,2} = x_{rm} + \frac{d(x_p - x_{rm})}{\sqrt{\varepsilon_r}(y_p + d)}. \quad (24)$$

In the conventional TWI, the image can be recovered through delay-and-sum beamforming (DS-BF). Thus, an image pixel $s[k, l]$ located at (x_k, y_l) is recovered by the following DSBF process:

$$s[k, l] = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} z[m, n] \exp(j2\pi f_n \tau_{pm}) \quad (25)$$

where $k = 0, 1, \dots, K - 1$ and $l = 0, 1, \dots, L - 1$. In (25), K and L are the numbers of pixels in the image. Multiple-signal classification subspace method can be used to facilitate the image reconstruction [4].

B. Compressed Sensing in TWI Radar

In TWI radar, if the number of targets is very small compared to the number of pixels in the image, i.e., $P \ll KL$, we may use the CS technique to recover such a sparse target scene.

a) *Formulating a dictionary Ψ for TWI Radar Data:* First of all, we define two long vector in the $X - Y$ plane and frequency-spatial plane as

$$\mathbf{s} = [s_{11}, s_{21}, \dots, s_{K1}, s_{12}, \dots, s_{K2}, \dots, s_{KL}]_{KL \times 1}^T \quad (26)$$

and

$$\mathbf{z} = [z_{11}, z_{21}, \dots, z_{M1}, z_{12}, \dots, z_{M2}, \dots, z_{MN}]_{MN \times 1}^T \quad (27)$$

where KL determines the resolution, \mathbf{s} is a weighted indicator vector, i.e., the kl index is nonzero if a target is located at position $[k, l]$. Here, z_{mn} represents the data set $z[m, n]$.

Thus, \mathbf{z} can be expressed as

$$\mathbf{z} = \Psi \mathbf{s} \quad (28)$$

where Ψ is a $MN \times KL$ matrix with $\Psi = [\psi_1, \psi_2, \dots, \psi_i, \dots, \psi_{KL}]$. The u th column of ψ_i is

$$(\psi_i)_u = \exp(-j2\pi f_a \tau_{i,u}) \quad (29)$$

where

$$\tau_{i,u} = \frac{\sqrt{(x_q - x_{b1})^2 + y_e^2} + \sqrt{(x_q - x_{b2})^2 + y_e^2}}{c} + \frac{\sqrt{\varepsilon_r}(\sqrt{(x_{b1} - x_{tb})^2 + d^2} + \sqrt{(x_{b2} - x_{rb})^2 + d^2})}{c} \quad (30)$$

where $a = i \bmod N$, $b = \lfloor i/M \rfloor$, $q = j \bmod K$, and $e = \lfloor j/L \rfloor$. Here, $\lfloor \cdot \rfloor$ denotes the floor function, $f_a = f_L + a\Delta f$, and X_{b1} and X_{b2} can be found via (23) and (24). Note that $(x_{tb}, -d)$ and $(x_{rb}, 0)$ are coordinates of the b th antenna pair. Therefore, once the size of detect range and radar parameters are specified, the dictionary Ψ can be constructed through (29).

b) Designing of Basis Φ : In CS based TWI radar system, Ψ is the conversion matrix, while Φ is the data acquisition matrix. Rather than sampling $z[m, n]$ (or the $MN \times 1$ vector \mathbf{z}), we measure the linear projection of \mathbf{z} onto a set of basis Φ as

$$\mathbf{y}_{J \times 1} = \Phi \mathbf{z} = \Phi_{J \times MN} \Psi_{MN \times KL} \mathbf{s}_{KL \times 1}. \quad (31)$$

A proper selection of Φ is needed to minimize the mutual coherence between Φ and Ψ . There are two possible selections of Φ are presented.

Principle of Selection I: Each row has only one nonzero element, being equals to one, in the MN elements. An example of such a selection is shown in Fig. 2. In this case, both spatial and frequency points are randomly selected, and this may lead to relative complex implementation in hardware of the TWI radar.

Principle of Selection II: As shown in Fig. 3, one first randomly picks R points out of M antenna pairs, then randomly selects T out of N frequency bins for all picked R positions of radar antennas. This approach is easier to implement in the TWI radar.

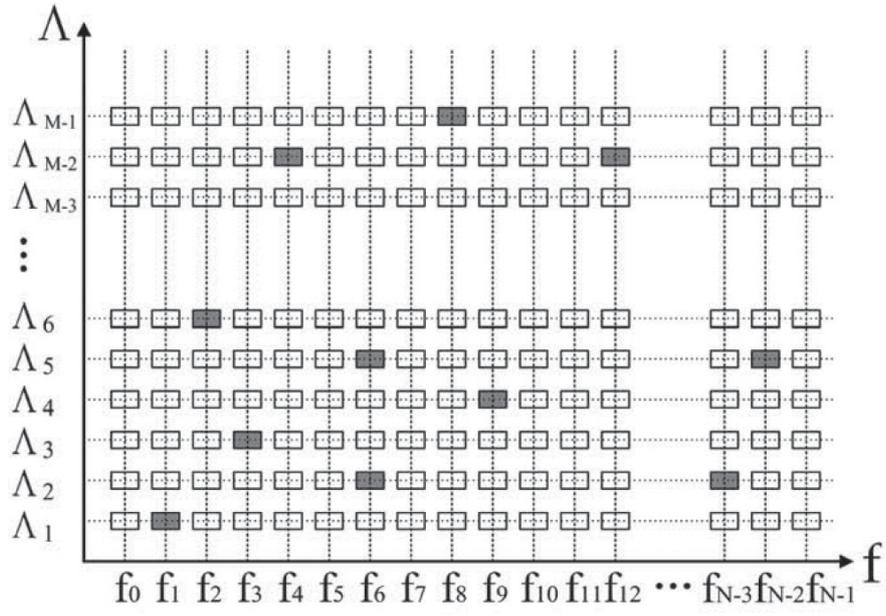


Fig. 6. Data acquisition algorithm for the TWI radar based on CS (Selection I).

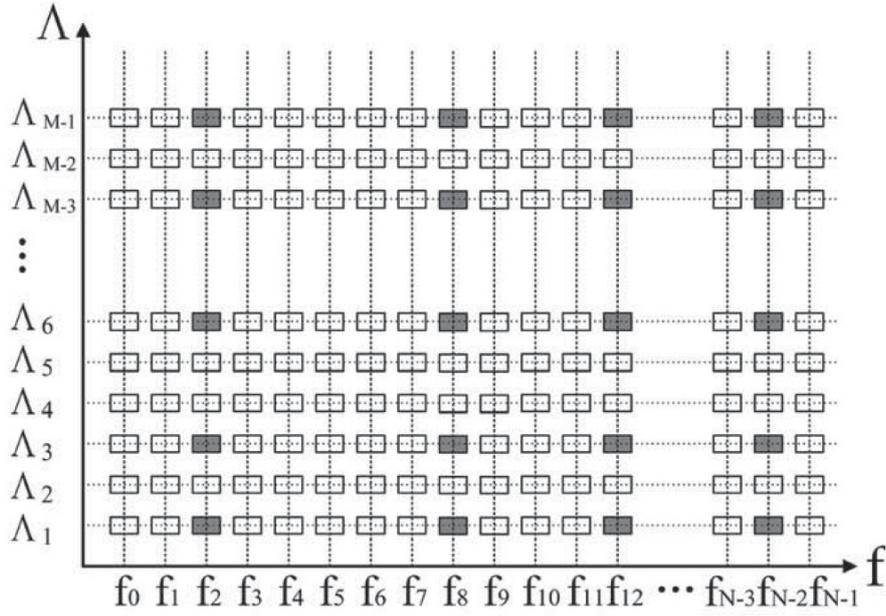


Fig. 7. Data acquisition algorithm for the TWI radar based on CS (Selection II).

C. TWI Detection with CS

Under the aforementioned assumptions, the target space indicator \mathbf{s} vector can be recovered from $J \geq C\mu^2(\Phi, \Psi)(\log(KL/P))P$ measurements. Using the convex optimization approach, we can find \mathbf{s} by solving an l_1 -norm minimization problem

$$\mathbf{s} = \operatorname{argmin} \|\mathbf{s}\|_{l_1} \text{ subject to } \mathbf{y} = \Phi\Psi\mathbf{s}. \quad (32)$$

Since the dictionary Ψ , measurement \mathbf{y} , and the indicator \mathbf{s} are complex, to facilitate the processing of (32), one can make some conversions as follows.

Define $\Theta = \Phi\Psi$. \mathbf{y} can be rewritten as

$$\mathbf{y} = \Theta\mathbf{s} = \Re\{\mathbf{y}\} + j\Im\{\mathbf{y}\} \quad (33)$$

where $\Re\{\mathbf{y}\} = \Re\{\Theta\}\Re\{\mathbf{s}\} - \Im\{\Theta\}\Im\{\mathbf{s}\}$ and $\Im\{\mathbf{y}\} = \Re\{\Theta\}\Im\{\mathbf{s}\} + \Im\{\Theta\}\Re\{\mathbf{s}\}$.

Thus, eqn. (32) becomes

$$\hat{\mathbf{I}} = \operatorname{argmin} \|\tilde{\mathbf{I}}\|_{l_1} \text{ subject to } \tilde{\mathbf{y}} = \Gamma\tilde{\mathbf{I}} \quad (34)$$

where $\tilde{\mathbf{y}} = [\Re\{\mathbf{y}\}^T \Im\{\mathbf{y}\}^T]^T_{2J \times 1}$, $\tilde{\mathbf{I}} = [\Re\{\mathbf{s}\}^T \Im\{\mathbf{s}\}^T]^T_{2KL \times 1}$ and

$$\Gamma = \begin{pmatrix} \Re\{\Theta\} & -\Im\{\Theta\} \\ \Im\{\Theta\} & \Re\{\Theta\} \end{pmatrix} \quad (35)$$

which is a $2J \times 2KL$ matrix. We comment that as soon as $\hat{\mathbf{I}}$ is reconstructed, one can readily obtain \mathbf{s} .

D. Discussion

Based on our study, we find that the topics on a CSR system can be summarized as

- 1) Application and setting;
- 2) Formulation in the context of CS: involving “reduced measurement + reduced DSP complexity” or “reduced DSP complexity” only;

- 3) Design of the matrix Φ ;
- 4) Reconstruction of signal or target position (via convex optimization);
- 5) Impacts from noise and inaccurate (modified) matrix $\tilde{\Psi}$.

V. A GENERAL MODEL OF CS USING UWB

A. Preliminary on UWB

UWB is one of the major breakthroughs in area of wireless communications. As UWB signal is highly sparse in the time domain, it is natural to introduce CS in the design of UWB systems.

Typical UWB systems have the following features:

1. The devices have low power consumption [12];
2. Signals with short pulse width (on the order of nanoseconds).

Based on the second feature, one can find that the corresponding Nyquist rate could be on the order of GHz. Analog-to-digital converters (ADCs) working at such high frequencies would be expensive. Two methods are proposed to reduce the sampling frequency. One way is to use an array of samplers with different time offsets to realize a serial sampling; however, the hardware cost will increase with the increasing of the number of samplers. The other way is that, in commercial applications, UWB receivers may ignore the details of UWB signals, i.e., detecting only the signal energy and realizing the on-off keying (OOK) modulation. In this case, timing information from the peak of a UWB signal is lost and can degrade the locationing precision.

B. Mathematic Model of UWB

Considering a wireless multi-path channel, the channel response can be expressed as

$$h(t) = \sum_{l=0}^{L-1} h_l \delta(t - \tau_l) \quad (36)$$

where L is the total number of paths, $\delta(\cdot)$ denotes the impulse function, h_l is the channel gain of the l th path, and τ_l is the delay in path l .

C. Compression of UWB

One considers $N + L - 1$ transmitted sequences and we denote the corresponding symbol by $\mathbf{s} = (s_1, s_2, \dots, s_N, \dots, s_{N+L-1})$. Thus, the sampled output $r[n_k]$ at time n_k can be written as

$$r[n_k] = \sum_{l=0}^{L-1} h_l[n_k] s[n_k - l] \quad (37)$$

where $h_l[n_k]$ is the l th complex channel filter tap at time n_k . Assuming that the channel taps do not vary over N symbol times, the received symbol reduces to

$$r[n_k] = \sum_{l=0}^{L-1} h_l s[n_k - l] \quad (38)$$

where $s[n_k] = 0$ for $n_k < 1$. To analyze the UWB channel, we transform the frequency-selective channel into flat fading channels. As a result, the received signal can be written as

$$\mathbf{r} = \mathbf{H}\mathbf{s} \quad (39)$$

where \mathbf{r} is the received signal vector after sampling and \mathbf{H} denotes a $N \times N + L - 1$ equivalent channel matrix which is given as

$$\mathbf{H} = \begin{pmatrix} h_0 & & \cdots & h_{L-1} & \cdots & h_1 \\ h_1 & h_0 & \cdots & & \cdots & h_2 \\ h_2 & h_1 & h_0 & & \cdots & h_3 \\ \vdots & \vdots & & & \vdots & \vdots \\ & & & h_{L-1} & \cdots & h_1 & h_0 \end{pmatrix}. \quad (40)$$

It is noted that (39) is a standard format of CS. However, the matrix \mathbf{H} may not satisfy the conditions for CS signal reconstruction because it is completely determined by the channel. As a result, we need to add additional compression procedure for the UWB signal. Such a compression can be done at either receiver or transmitter side. Details are given in the following subsections.

c) *Receiver Side Compression:* Receiver side compression for UWB signal can be implemented based on distributed amplifier (DA) [13]. The basic operation is letting the received UWB signal pass through a microchip circuit, along which the signal is added to predetermined weights and then sampled.

DA, also called transverse filter, consists of multiple repeated taps, each containing a section of microchip input and output transmission lines, and then gain cells. Such an architecture for DA is given in Fig. 8. Thus, the output of a DA at time t is

$$y(t) = \sum_{k=1}^N a_k x(t - k\tau) \quad (41)$$

where N is the number of taps in a DA, a_k 's are attenuation coefficients of different gain cells, which can be set randomly, x is the input signal, and τ is the fixed time delay of each section of transmission line. We comment that the attenuation coefficients and time delay can be predetermined from hardware design. DA is suitable for UWB signal as the characteristic impedance of a transmission line changes very little over several GHz, and the time delays can easily achieve a time scale of 50 ps or less without changing the DA structure.

Based on the proposed DA structure, as shown in Fig. 8, the received UWB signal is put into the compressor with M DAs, each of which has N gain cells, and then sampled by M ADCs. For simplicity, here we ignore the noise. Then the output of the i th ADC is given by

$$y_i = \sum_{k=1}^N \phi_{ik} x_k \quad (42)$$

where $x_k = x(xT_s - k\tau)$ with T_s denoting the sampling period of ADC and ϕ_{ik} is the coefficient of the k th gain cell of the i th DA.

Let us define $\mathbf{x} = (x_1, x_2, \dots, x_N)^T$ and $\mathbf{y} = (y_1, y_2, \dots, y_M)^T$. Then we can obtain the linear compression equation as

$$\mathbf{y} = \Phi \mathbf{x} \quad (43)$$

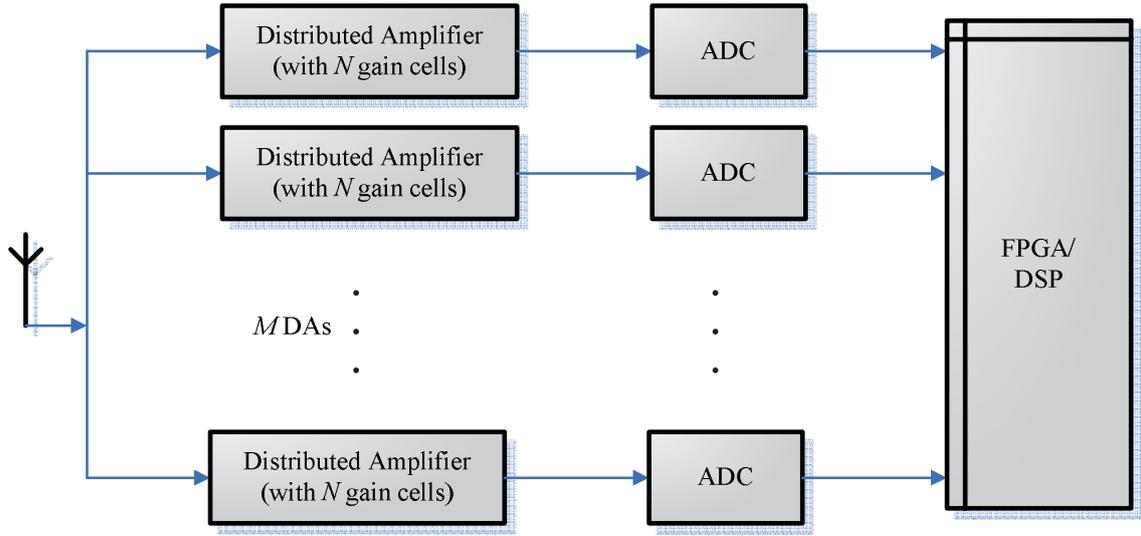


Fig. 8. An illustration of the DA-based receiver structure.

where the measurement matrix Φ is given by

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{M1} & \phi_{M2} & \cdots & \phi_{MN} \end{pmatrix}. \quad (44)$$

With proper design of DA, one can make (43) suitable for CS signal processing.

d) Transmitter Side Compression: Alternatively, we can modify the transmitter to accommodate CS technique. One approach with the functionality of signal mixing is illustrated in Fig. 9.

In such a structure, the UWB impulses are generated according to the bit sequence and then pass through a finite impulse response (FIR) filter. The filter output is then sent out through the wireless channel and reaches the receiver. Suppose that the UWB pulse, together with the silent period, is denoted by an N -vector \mathbf{s} in the discrete time domain. Then, the received signal at the receiver is given by

$$\mathbf{y}_0 = \mathbf{f} * \mathbf{h} * \mathbf{s} \quad (45)$$

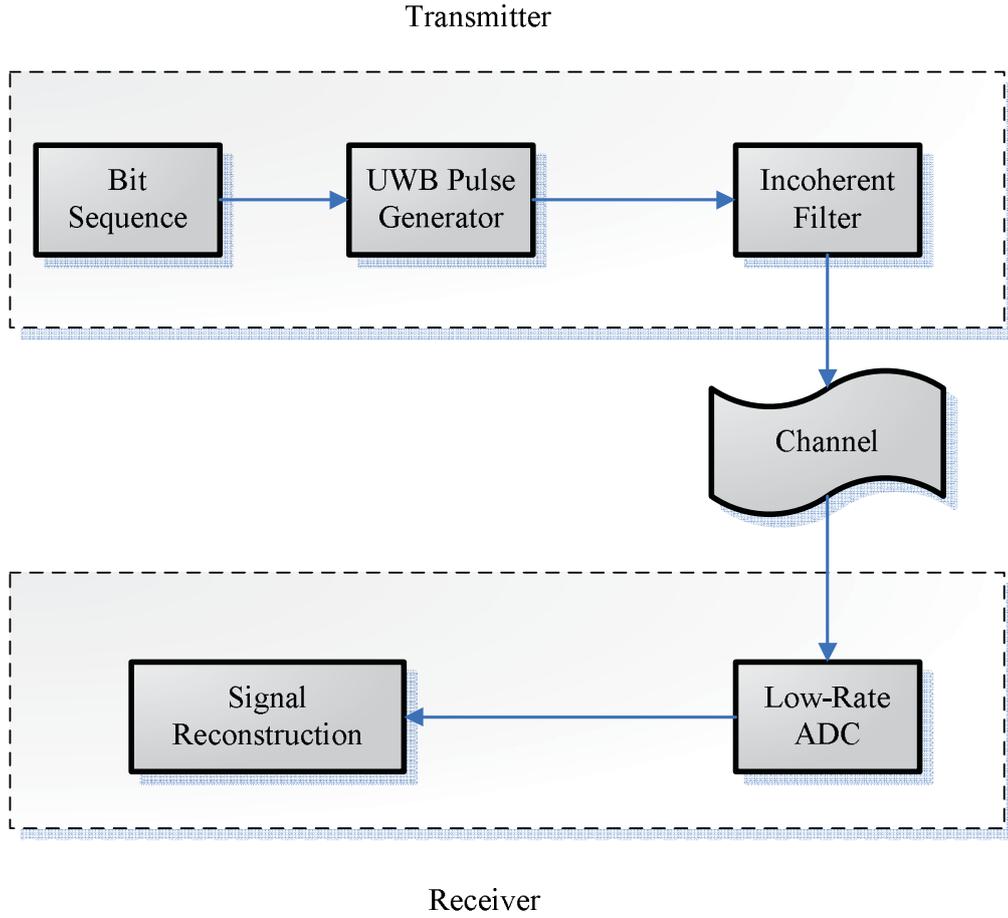


Fig. 9. An illustration of the UWB transmitter structure with signal mixing.

where \mathbf{f} is the discrete time impulse response of the FIR filter at the transmitter and $*$ denotes the convolution operation. Due to the limitation of sampling rate, the vector \mathbf{y}_0 may not be obtained directly. After downsampling to M uniform measurements, the signal after sampling is now given by

$$\mathbf{y} = \Phi_{\mathbf{t}} * \mathbf{x} \quad (46)$$

where we ignore the quantization error, for simplicity of analysis. The measurement matrix for transmitter compression is determined by

$$y(n) = \sum_{j=0}^{B-1} x(n \lfloor N/M \rfloor + j) (\mathbf{f} * \mathbf{h})(B - j) \quad (47)$$

where B is the dispersion of the composite filter $\mathbf{f} * \mathbf{h}$. The matrix Φ_t is a quasi-Toeplitz matrix in which the non-zero elements are located within a strip and each column has B elements. Thus, with designing of larger B , the measurement matrix can be filled with more non-zero elements and this make \mathbf{y} more dispersed in the time domain. More details of this approach can be found in [14].

D. Reconstruction of UWB

In UWB system, the UWB impulse arrive in a block manner. Hence, in the unknown vector \mathbf{r} , the non-zero elements are grouped in blocks, it is impossible to have the non-zero elements distributed in arbitrary locations of \mathbf{x} . Traditional BP or OMP do not consider this prior information. Hence, by using block reconstruction, one can improve the receiver performance.

e) Block Property: One can exploit the block property of UWB signals. For simplicity, we ignore the noise in this discussion. One can define the block sparsity by rewriting the unknown received vector r as

$$\mathbf{x} = \left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{\lceil \frac{N}{d} \rceil} \right). \quad (48)$$

Each subvector \mathbf{x}_z is called a block. Except the last one, each \mathbf{x}_z has d elements. Here, one introduces the block sparsity definition: if at most k blocks have non-zero elements, we say that the vector x is **block k -sparse**.

f) l_1 and l_2 combined optimization, L-OPT: Define $Z = \lceil \frac{N}{d} \rceil$, we denote the mixed l_1/l_2 -norm as

$$\|\mathbf{x}\|_{2,Z} = \sum_{z=1}^Z \|\mathbf{x}_z\|_2 \quad (49)$$

where $\|\cdot\|$ denotes the l_2 norm. The L-OPT minimizes a combination of l_1 - and l_2 -norm of the unknown vector \mathbf{x} through

$$\min_{\mathbf{x}} \sum_{z=1}^Z \|\mathbf{x}_z\|_2 \quad s.t. \quad \mathbf{y} = \Phi \mathbf{r}. \quad (50)$$

We comment that: when $d = 1$, one minimizes l_2 -norm of \mathbf{x} ; when $d = N$, one minimizes l_1 -norm of \mathbf{x} . For a general value of d , one minimizes something between l_1 -norm and l_2 -norm.

g) *Condition of Perfect Reconstruction:* A matrix \mathbf{A} is said to satisfy the block restricted isometric property (RIP) condition if there exist parameters δ_k such that, for every vector \mathbf{c} that is block k -sparse, such that

$$(1 - \delta_k)\|\mathbf{c}\|_2^2 \leq \|\mathbf{A}\mathbf{c}\|_2^2 \leq (1 + \delta_k)\|\mathbf{c}\|_2^2. \quad (51)$$

This condition can be used to evaluate the matrix \mathbf{A} in CS-based UWB system design.

VI. CS PERFECT SIGNAL RECONSTRUCTION

Clearly, two basic conditions are required for a CS system. Theoretical analysis of CS requires signal sparsity. In addition, RIP or incoherence condition on sensing matrix is desired for reliable signal recovery [17].

To facilitate the description of incoherence conditions, we rewrite the CS signal model as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} = \underbrace{\mathbf{\Phi}\mathbf{\Psi}}_{\mathbf{\Theta}}\mathbf{s} = \mathbf{\Theta}\mathbf{s} \quad (52)$$

where $\mathbf{\Phi}$ is the measurement matrix, $\mathbf{\Psi}$ is the basis matrix, $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$ denotes the sensing matrix, and vector \mathbf{s} is K -sparse.

A. Conditions on RIP

Let us define RIP. Suppose that there exists a constant δ such that, for every K -sparse vector \mathbf{v}_K

$$(1 - \delta)\|\mathbf{v}_K\|_2^2 \leq \|\mathbf{\Theta}\mathbf{v}_K\|_2^2 \leq (1 + \delta)\|\mathbf{v}_K\|_2^2. \quad (53)$$

To continue, we define the minimum of all constants δ satisfying (53) as the isometry constant δ_K . In [19], Davenport and Wakin showed (in 2010) that for

$$\delta_K < \frac{1}{3\sqrt{K}} \quad (54)$$

if RIP is satisfied, the orthogonal matching pursuit (OMP) algorithm is guaranteed to recover s exactly from the l_1 -norm minimization problem. Later in 2012, Wang and Shim loosened the inequity condition in (54) to [20]

$$\delta_K < \frac{1}{\sqrt{K} + 1}. \quad (55)$$

B. Conditions on Mutual Coherence

The mutual coherence of a sensing matrix Θ is defined as

$$\mu(\Theta) = \max_{i \neq j} | \langle \theta_i, \theta_j \rangle | \quad (56)$$

where θ_i and θ_j are two properly normalized column vectors of Θ .

Lemma 1

If the following condition is satisfied

$$K < 0.5 \left[1 + \frac{1}{\mu(\Theta)} \right] \quad (57)$$

OMP algorithm is guaranteed to find s exactly from the l_1 -norm minimization problem. Note that the condition in Lemma 1 is a sufficient and necessary condition. Its necessity can be shown by slightly relaxing the condition ($K = 0.5 [1 + 1/\mu(\Theta)]$), which leads to imperfect recovery. Detailed proofs can be found in [21].

Sometimes, perfect recovery may be hard. Thus, one can ask for recovery with high probability. Such a condition is given as follows [23].

Lemma 2

Given $\mu(\Theta) = 1/\sqrt{M}$, suppose a k -sparse vector s follows $K \leq M/(16 \log(M/\varepsilon))$ for some sufficient small ε . For a CS system, basis pursuit (BP) algorithm will recover s with probability greater than $1 - 2\varepsilon^2 - K^{-\vartheta}$ for some $\vartheta \geq 1$ s.t. $\sqrt{\vartheta \log M / \log(M/\varepsilon)} \leq c$ where c is an absolute constant.

C. Study of the Coherence Condition

Suppose we have a CS based high resolution radar system. The received signal can be formulated in the CS matrix format as

$$\mathbf{y} = \Phi \mathbf{x} = \underbrace{\Phi \Psi}_{\Theta} \mathbf{s} = \Theta \mathbf{s} \quad (58)$$

where Φ is a $M_m \times N$ measurement matrix, Ψ is a $N \times ML$ basis matrix, $\Theta = \Phi \Psi$ denotes a $M_m \times ML$ sensing matrix, and vector \mathbf{s} is K -sparse.

The mutual coherence of a sensing matrix Θ is defined as

$$\mu(\Theta) = \max_{i \neq j} | \langle \theta_i, \theta_j \rangle | \quad (59)$$

where θ_i and θ_j are two properly normalized column vectors of Θ . One of our target is to find the CDF of $\mu(\Theta)$ for a given CS system (fixed Ψ).

We have an $M_t \times M_r$ MIMO CS system setup, which can be extended to an electronic surveillance multiple-receiver at a later time. By discretizing the range-Doppler plane into $M \times L$ grids, we are able to write the elements of Ψ as

$$\psi_{uv} = \sum_{l=1}^{M_t} W_l(uT_s) \exp(-j2\pi f_c \tau_{lq}(x_v)) \exp(j2\pi f_{lq}^v(uT_s - \tau_{lq}(x_v))) \quad (60)$$

where $W_l(\cdot)$ is the received signal waveform, $\tau_{lq}(\cdot)$ denotes the time delay of target reflected signal from the l th transmitter to q th receiver for a specific grid, and f_{lq} is the Doppler frequency shift corresponding to the grid.

Based on the CS model, we derive the CDF of mutual coherence. We can rewrite it as the maximum of $|g_{ij}| = |\theta_i^H \theta_j|$.

Case I: If we assume the elements of Φ are independently drawn from the random sequence “ $\{\pm 1\}$ ”, it can be shown that

$$|g_{ij}| = M_m |\Psi_i^H \mathbf{I} \Psi_j|. \quad (61)$$

From (61), we conclude that mutual coherence will not be affected by Φ matrix and depend on Ψ matrix, i.e., the system parameter.

Case II: If we assume the elements of Φ are independently drawn from a zero-mean gaussian distribution, the direct derivation for the target CDF can be difficult as g_{ij} 's are dependent complex random variables (RVs). For simplicity, we consider two elements g_{12} and g_{13} for analytical analysis purpose. Through a pre-whitening process (detailed steps omitted here), we have

$$|\tilde{g}_{12}| = |T_{11}g_{12} + T_{12}g_{13}| \quad (62)$$

and

$$|\tilde{g}_{13}| = |T_{21}g_{12} + T_{22}g_{13}|. \quad (63)$$

Thus, $|\tilde{g}_{12}|$ and $|\tilde{g}_{13}|$ are found to be independent Rician RVs.

If $|\tilde{g}_{12}|$ and $|\tilde{g}_{13}|$ are i. n. i.d. RVs, the CDF and PDF of $\max\{|\tilde{g}_{12}|, |\tilde{g}_{13}|\}$ are respectively given by

$$\prod_{i=1, j=1, 2} F_{\tilde{g}_{ij}}(\tilde{g}) \quad (64)$$

and

$$\sum_{j=1}^2 f_{\tilde{g}_{1j}}(\tilde{g}) \prod_{k=1, k \neq j}^2 F_{\tilde{g}_{1k}}(\tilde{g}). \quad (65)$$

We would like to determine a reasonable tight bounding for the CDF of $\max\{|g_{ij}|\}$, which can be used as a metric to assess the recovery performance of a CS system. We also target to find the minimum number of samples required for a perfect recovery through our studies. Such a parameter can be useful in determining the required power and speed of signal processing unit.

In addition, coherence between Φ and Ψ may be investigated to provide more insights into the low coherence condition for high-resolution CS systems. For the CS system discussed, we will explore the CS system with more realistic conditions, for instance, the additive noise/interference, in the future work.

REFERENCES

- [1] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21-30, Feb. 2008.

- [2] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [3] Y. Zhang, "Theory of compressive sensing via l_1 -minimization: A non-RIP analysis and extensions," *CAAM Technical Report*, Rice University, TR08-11, Houston, TX, 2008.
- [4] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing", *ECE Department Tech. Report TREE-0612*, Rice University, Houston, TX, Nov. 2006.
- [5] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *IEEE Signal Process. Mag.*, vol. 25, pp. 92-101, Feb. 2008.
- [6] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203-4215, Dec. 2005.
- [7] L. C. Potter, E. Ertin, J. T. Parker, M. Cetin, "Sparsity and Compressed Sensing in Radar Imaging," *Proc. IEEE*, vol. 98, pp. 1006-1020, June 2010.
- [8] 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, pp. 1718-1729, Apr. 2012.
- [9] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", *IEEE Trans. Inform. Theory*, vol. 52, pp. 489-509, Feb. 2006.
- [10] M. Lustig, D.L. Donoho, and J.M. Pauly, "Rapid MR imaging with compressed sensing and randomly under-sampled 3DFT trajectories", in *Proc. 14th Ann. Meeting ISMRM*, Seattle, WA, May 2006.
- [11] A. P. Chandrakasan and V. M. Stojanovic, "Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors," *IEEE J. Solid-State Circuits*, vol. 47, pp. 744-756, Mar. 2012.
- [12] J. Zhang, P. Orlik, Z. Sahinoglu, A. F. Molisch, and P. Kinney, "UWB systems for wireless sensor networks," *Proc. IEEE*, vol. 97, pp. 313-331, Feb. 2009.
- [13] D. Yang, H. Li, G. D. Peterson, and A. Fathy, "Compressed sensing based UWB receiver: Hardware compressing and FPGA reconstruction", in *Proc. Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MD, USA, Mar. 2009.
- [14] P. Zhang, H. Zhen, R. C. Qiu, and B. M. Sadler, "A compressed sensing based ultra-wideband communication system," in *Proc. IEEE Intl' Conf. Commun. (ICC)*, Dresden, Germany, June 2009.
- [15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [16] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing", *ECE Department Tech. Report TREE-0612*, Rice University, Houston, TX, Nov. 2006.
- [17] C. Hegde, M. F. Duarte, and V. Cevher, "Compressive sensing recovery of spike trains using a structured sparsity model," *Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, Saint Malo, France, Apr. 2009.
- [18] L. Anitori, A. Maleki, M. Otten, R. G. Baraniuk, P. Hooeboom, "Design and Analysis of Compressed Sensing Radar Detectors," *IEEE Trans. Signal Process.*, vol. 61, pp. 813-827, Feb. 2013.
- [19] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4395-4401, Sept. 2010.
- [20] J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, pp. 4973-4976, Sept. 2012.

- [21] J. Wang and B. Shim. "A simple proof of the mutual incoherence condition for orthogonal matching pursuit" submitted for publication.
- [22] W. Min, "High resolution radar imaging based on compressed sensing and adaptive Lp norm algorithm," *IEEE CIE International Conference on Radar*, pp. 206-209, Oct. 2011.
- [23] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, pp. 2275-2284, June 2009.