Pitfalls and possibilities of radar compressive sensing

Nathan A. Goodman[1] and Lee C. Potter[2,*]

[1]School of Electrical & Computer Engineering, The University of Oklahoma, 110 W. Boyd Street, Norman, Oklahoma 73019, USA

[2]Department of Electrical & Computer Engineering, The Ohio State University, 2015 Neil Avenue, Columbus, Ohio 43210, USA

*Corresponding author: potter.36@osu.edu

Received 4 November 2014; revised 3 January 2015; accepted 3 January 2015;
posted 6 January 2015 (Doc. ID 226150); published 10 February 2015

In this paper, we consider the application of compressive sensing (CS) to radar remote sensing applications. We survey a suite of practical system-level issues related to the compression of radar measurements, and we advocate the consideration of these issues by researchers exploring potential gains of CS in radar applications. We also give abbreviated examples of decades-old radio-frequency (RF) practices that already embody elements of CS for relevant applications. In addition to the cautionary implications of system-level issues and historical precedents, we identify several promising results that RF practitioners may gain from the recent explosion of CS literature.

OCIS codes: (280.0280) Remote sensing and sensors; (280.5600) Radar; (100.3200) Inverse scattering; (100.6640) Superresolution.

http://dx.doi.org/10.1364/AO.54.0000C1

1. Introduction

Compressed sensing is an elegant and appealing collection of results from applied mathematics, statistics, and numerical optimization. The burgeoning literature is now moving into its second decade under the names "compressed" or "compressive" sensing and has garnered attention across many application domains, including radar sensing. Initial compressive sensing (CS) studies for radar application have necessarily been limited in scope, anecdotal, and typically qualitative. For an evolving field seeking transformative impact in engineering application, this paper offers a brief tutorial essay with the aim of identifying practical considerations necessary for a balanced assessment of compressive sensing applied to radar systems.

From the perspective of system performance and practical implementation, a number of issues are deserving of methodical evaluation as CS studies evolve from anecdote to application:

1. Compressive data acquisition suffers from signal-to-noise ratio (SNR) loss; for a radar application with performance limited by thermal noise at the receiver, any SNR loss may be unacceptable.

2. Nonlinear processing in CS inverts an assumed system model; uncalibrated imperfections in the system model effectively contribute additional noise power to the thermal noise physically present in the system. Examples in radar sensing include imperfect calibration of antenna responses across frequency, polarization, and angle; amplifier nonlinearities; distributed clutter; and internal target motion.

3. Radar data are, in many applications, not compressible. For example, in air-to-ground radar imaging, the raw phase history data admit approximately 2:1 lossless compression, and up to 8:1 lossy compression when limiting image degradation to one point on the NIRS scale. Thus, radar data, in many applications, have high entropy.

4. The nonlinear processing intrinsic to CS yields image artifacts that may confound both image analysts and existing automated exploitation algorithms; in contrast, users understand and accept the structured and predictable artifacts of linear processing. Moreover, the nonlinear processing intrinsic to CS may yield output statistics that are

1559-128X/15/0800C1-13$15.00/0

difficult to predict or estimate, thereby degrading performance in detection and tracking applications.

5. Physical parameters relevant to radio-frequency (RF) sensing, such as delay, Doppler frequency, and angle of arrival, are continuous parameters. CS theory, on the other hand, requires a known fixed grid of parameter values. The mismatch between grid points and actual parameter values destroys sparsity.

6. Putative CS gains must be fairly benchmarked to existing practice and measured in terms of reduced size, weight, and power or cost (SWaP-C). Hardware requirements of proposed data acquisition should be considered and compared to existing systems. Seven decades of RF engineering practice have yielded system designs that may already exploit signal parsimony to improve SWaP-C. Two generic examples are sparse arrays, which provide a large aperture with reduced hardware costs, and stretch processing of linear frequency-modulated (LFM) waveforms, which reduces analog-to-digital conversion bandwidth far below Nyquist rates. Additionally, various RF practices have for decades already embodied the nonuniform subsampling, low-coherence measurement operators, and ℓ_1 regularization or greedy optimization algorithms that characterize much of the CS literature.

On the one hand, very few published studies provide quantifiable system-level benefits of CS for radar in terms of SWaP-C, area under receiver operating characteristic curve, parameter estimation error variance, classification rate, etc. Instead, studies typically demonstrate qualitative assessment of signal or image reconstruction and omit comparison to existing RF engineering practice. On the other hand, there have been isolated demonstrations that seem promising. Some quantities of interest, such as aircraft positions at altitude in a ground-to-air surveillance application or temporal changes in a scene, are quite compressible. Nonuniform sparse sampling of an aperture, combined with nonlinear processing, can provide low side-lobes, compared to matched filtering, and, with sufficient SNR, parameter accuracy can exceed the diffraction limit.

In the following sections, we briefly discuss the six cautionary issues listed above, with an eye toward potential impacts on future systems. In addition to the cautionary discussion of issues, we also highlight several promising results and indirect benefits of CS research for radar systems. This brief essay cannot offer a comprehensive survey of the large literature; instead, we seek to enumerate issues and aim only to be representative in characterizing recent literature. Related survey articles have appeared elsewhere [1,2,3,4,5].

The paper is organized as follows. In Section 2 we briefly define CS for the purposes of this paper. Section 3 provides a nominal radar measurement model and relates it to a standard CS formulation. In Section 4 we elaborate on the six system issues enumerated above; for the case in which parameter estimation is the goal, we employ the familiar Cramèr–Rao lower bound to highlight the role of SNR in a system-level evaluation of CS. Section 5 briefly points to promising results and benefits derived from the CS literature. Conclusions are summarized in Section 6.

2. Compressive Sensing

For the purposes of this paper what do we mean by "CS?" The literature is extremely large and continues to grow, but here we invoke four defining components: sparsity, incoherent linear measurements, convex optimization, and provably stable recovery. Consider the recovery of a vector $f \in \mathbb{C}^N$ from $M < N$ noise-corrupted linear measurements, $y = \Phi f + n$. The problem is ill-posed due to the nontrivial nullspace of the measurement operator Φ. However, suppose f is *a priori* assumed to be parsimoniously represented in some basis, i.e., $f = \Psi x$. The vector x is said to be *s-sparse* if it has s or fewer nonzero entries. In practice, x may be only approximately s-sparse; so, for any x let x_s denote the best s-sparse approximation, i.e., the vector consisting of the s largest magnitude entries from x. CS provides guarantees for stable recovery of an approximately sparse x, for suitable measurement operator Φ. Define $A = \Phi\Psi$. The M-by-N matrix A with unit-length columns and ith column denoted a_i has *coherence* μ_A given by

$$\mu_A = \max_{i \neq j} |a_i^H a_j|. \tag{1}$$

Given a matrix A with low coherence, for recovery of x we consider the constrained ℓ_1 minimization, a *convex optimization* known as basis pursuit denoising (BPDN) [6]:

$$\hat{x} = \arg\min_{\bar{x}} \|\bar{x}\|_1 \text{ subject to } \|y - A\bar{x}\|^2 \leq \eta. \tag{2}$$

The ℓ_1 norm appearing in Eq. (2) can be viewed as a *convex relaxation* of the ℓ_0 counting norm, thereby converting the NP-hard optimization of finding the sparsest solution into a tractable convex optimization. The recovery of x is *provably stable*.

Theorem 1 *[7, Thm. 2.1] Let $y = Ax + n$, $\|n\|_2 \leq \epsilon$, $2s < 1 + \mu_A^{-1}$, and $0 < \epsilon \leq \eta$. Then,*

$$\|x - \hat{x}\|_2 \leq C_0(\epsilon + \eta) + C_1 \|x - x_s\|_1, \tag{3}$$

with positive constants C_0, C_1 depending on μ_A and s.

Thus, the reconstruction error energy is bounded in Eq. (3) by two terms: the first depends on noise energy and the second on the sparse approximation error. The convex optimization may equivalently be written in its dual form, yielding the least absolute shrinkage and selection operator (LASSO) [8], which is easily recognized as an ℓ_1-regularized least-squares estimator:

$$\min_{\bar{x}} \|y - A\bar{x}\|_2^2 + \lambda\|\bar{x}\|_1. \quad (4)$$

Thus, CS gives a provable guarantee that convex optimization faithfully recovers an approximately s-sparse signal from about $s \log N$ incoherent (i.e., small μ_A) measurements. For design of Φ yielding small μ_A, the literature generally relies on pseudo-random sampling, for example by selecting rows of an orthogonal matrix (e.g., Hadamard or Fourier) uniformly and independently, or by circular convolution with a random impulse response. And, the optimizations in Eq. (2) or Eq. (4) may equivalently be accomplished by greedy algorithms. The ability to recover a length-N, but sparse, vector from many fewer than N measurements gives rise to the moniker "compressive sensing."

The estimator in Eq. (4) admits a Bayesian interpretation as the maximum posterior probability (MAP) estimate of x. In particular, the objective function is proportional to the negative log likelihood of the data for white Gaussian noise and a Laplace prior on x. Other, non-Laplacian Bayesian approaches to sparse signal recovery have been proposed (e.g., [9,10]), and a sum of regularizers has been widely adopted, including the so-called "cosparse" or "analysis" model in which $\|\Psi' x\|_p$ is adopted to encourage sparse coefficients for the (possibly nonsquare) analysis matrix Ψ'. Further, numerical techniques prompted by CS literature have invigorated the development of Bayesian estimation algorithms, for both MAP and minimum mean-squared error (MMSE), which adopt sparse priors [11,12].

3. Radar Measurement Model and Consequences

A. Projection Operator

In contrast to the standard measurement model for CS, the best starting point for a RF CS measurement model is $y = \Phi(f + n)$, where noise is added *prior* to the action of the measurement operator, Φ. The physical rationale for this seemingly minor modeling adjustment and the significant consequences are discussed in the sequel. Consider that RF sensing involves capture of a signal that propagates in time and three-dimensional (3D) space, and denote this signal as $f(\mathbf{r}, t)$. The signal has finite spatial and spectral bandwidths that set Nyquist sampling requirements in space and time. The signal is measured by projecting it onto a collection of M measurement kernels $\phi_m(\mathbf{r}, t)$ that also have support in time and 3D space. We start with the assumption that the measurement kernels operate on a noisy version of the signal, $f(\mathbf{r}, t) + n(\mathbf{r}, t)$, and will revisit the assumption later. The expression governing the mth measurement is then

$$y_m = \int \phi_m(\mathbf{r}, t)[f(\mathbf{r}, t) + n(\mathbf{r}, t)]\mathrm{d}\mathbf{r}\mathrm{d}t. \quad (5)$$

This general expression is capable of describing the sampling operation for a variety of hardware architectures. For example, traditional temporal sampling of a RF receiver can be described with a 3D effective antenna aperture and a properly shifted temporal sampling function. In this case, the measurement kernel would be $\phi_m(\mathbf{r}, t) = u(\mathbf{r})p(t - t_m)$, where $u(\mathbf{r})$ represents the antenna aperture function, $p(t)$ is the temporal sampling function defined by filters in the receive chain and the physical process of analog-to-digital conversion, and t_m is the mth sampling time.

The antenna aperture function enforces the antenna element pattern. If the antenna aperture is small relative to the reciprocal of the signal's spatial bandwidth, then the element pattern has little effect and the antenna aperture function can be approximated as a delta function at the antenna phase center location. Likewise, the temporal sampling function enforces receiver components such as anti-aliasing filters. If the receiver bandwidth is larger than the signal's temporal bandwidth, then the temporal sampling function can also be approximated as a delta function. Under these approximations, $\phi_m(\mathbf{r}, t) = \delta(\mathbf{r} - \mathbf{r}_0)\delta(t - t_m)$, where $\mathbf{r}_0$ is the antenna location, and the mth measurement is

$$\begin{aligned} y_m &= \int \delta(\mathbf{r} - \mathbf{r}_0)\delta(t - t_m)[f(\mathbf{r}, t) + n(\mathbf{r}, t)]\mathrm{d}\mathbf{r}\mathrm{d}t \\ &= f(\mathbf{r}_0, t_m) + n(\mathbf{r}_0, t_m). \end{aligned} \quad (6)$$

By appropriately designing a set of sampling kernels, Eq. (5) can be used to represent phased array antennas, multichannel receiver systems, slow-time/fast-time sampling of a radar waveform, and any other desired RF receiver system. Here, "slow time" refers to the time scale of multiple pulses, whereas "fast time" refers to time shifts within a single pulse interval.

B. Measurement Compression

As described above, compressive measurement kernels should have low coherence with the basis in which the signal has a sparse approximation. Intuitively, this requirement states that the measurement kernels should not be localized in the same basis in which the signal is sparse; otherwise, it is likely that the localized measurement kernel will completely miss the localized signal. Propagating plane waves are not localized in spatial position; therefore, it is safe to undersample the spatial field with physically localized antennas in 3D space (which is fortunate because it is difficult to implement arbitrary aperture functions over physically large antennas). Spatial-domain RF compressive sampling, therefore, has typically been envisioned via thinned or sparse arrays, which is a form of compression via measurement thinning or *downselection*. Figure 1 shows several examples of radar measurement compression via a thinned-measurement approach.

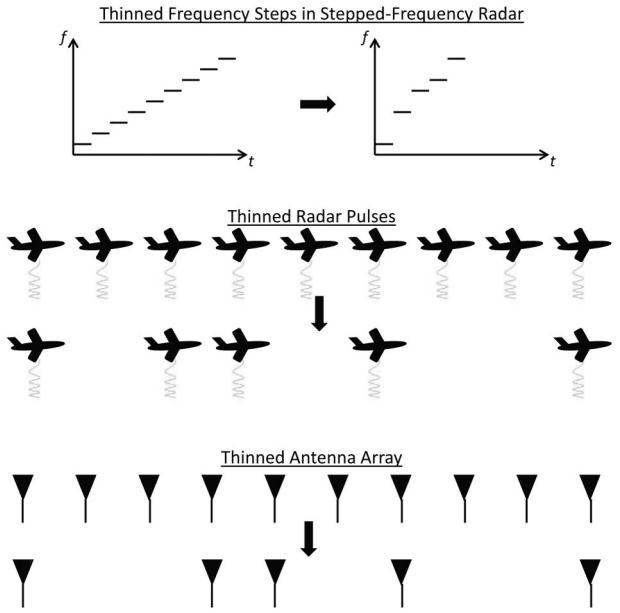

Fig. 1. Illustration of various thinning-based compression schemes for radar and array processing.

On the other hand, time-domain signals of interest sometimes exist for short time durations. Therefore, simply reducing the temporal sampling rate (or thinning the temporal measurements) might not be an effective form of compression. If the elements of the sparse basis are *modulated waveforms* (as in, for example, phase-coded or LFM radar pulses), then their time spread might enable compression via sample rate reduction. But if a simple sample rate reduction means that the sampling interval is similar to or longer than the signal's time duration, then straightforward fast-time undersampling is not appropriate and the analog time-domain signal must be modulated through analog multiplication with a wideband measurement kernel prior to sampling [13], as shown in Fig. 2. The consequence of this behavior is that radar slow-time measurements are usually compressed via downselection of radar pulses [14], while fast-time compression might be best achieved through mixing-based compression [15]. Alternatively, "fast-time compression" can also be achieved through a stepped frequency implementation whereby the radar emits a sequential series of K narrowband tones covering the bandwidth of interest (see Fig. 1). As long as target transfer functions are not sparse in the frequency domain (which they are not), downselection-based compression can be achieved by skipping many of the frequency steps

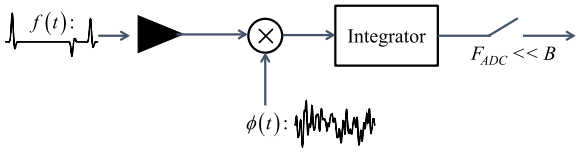

Fig. 2. Illustration of mixing-based compression for sub-Nyquist A/D acquisition.

during the data collection [16,17]. The fundamental tradeoff in stepped frequency systems is data collection time, which can be mitigated by a compressive implementation that skips many of the steps.

In the above discussion, we observe two basic types of measurement compression [18]. When the signal is not localized in space, time, and/or temporal frequency, compression can be achieved by thinning the measurements via sparse array, reduced fast-time sampling rate, thinned pulses, and/or thinned frequency steps (Fig. 1). A second method of compression involves modulating the signal through analog multiplication with a wideband measurement kernel before sampling at a reduced rate (Fig. 2). This *mixing*-based compression is more in line with the typical compression methods employed in optical and spectral imaging applications, where it is possible to use masks, diffraction gratings, and other optical components to implement analog-domain mixing with spatial kernels having a wide space-bandwidth product. In RF applications, the wavelengths involved make it difficult to implement arbitrary, spatially extended measurement kernels that spatially encode the signal over a large physical or synthetic aperture. Hence, spatial compression must be performed via thinned physical or synthetic arrays, which reduces the collected signal energy (and SNR) due to a reduced effective aperture.

C. Preprojection Noise

At first glance, it may seem strange to insist that noise should be added to the signal before compression via downselection. We can obviously separate the signal and noise terms in Eq. (5) to obtain an equivalent postprojection noise model according to

$$y_m = \int \phi_m(\mathbf{r},t) f(\mathbf{r},t)\mathrm{d}\mathbf{r}\mathrm{d}t + \int \phi_m(\mathbf{r},t) n(\mathbf{r},t)\mathrm{d}\mathbf{r}\mathrm{d}t$$
$$= \int \phi_m(\mathbf{r},t) f(\mathbf{r},t)\mathrm{d}\mathbf{r}\mathrm{d}t + n_m. \quad (7)$$

While Eq. (7) is valid and may appear identical to the usual CS model, using Eq. (7) as a starting point can obscure important system-level factors. Such factors include the impact of measurement kernels on the noise correlation structure and finite transmit power limitations that cannot be increased to compensate for missing antenna aperture, radar pulses, or frequency steps.

When beginning from the model of Eq. (5), it is clear that the transmit power, antenna gain, noise figure of the RF front end, and other system factors must be carefully considered when setting the ratio of $f(\mathbf{r},t)$ to $n(\mathbf{r},t)$ prior to compression. If measurements are subsequently thinned, as in Fig. 1, then signal energy is lost. If compression is obtained through signal mixing, as in Fig. 2, then the same analog operations that are used to encode linearly independent foldings of the wideband signal spectrum into a relatively narrow sampling band are also applied to the full noise spectrum. Thus, mixing-based

compression causes the wideband noise spectrum to alias into the narrow sampling band, which increases noise power through an effect known as noise folding [19,20]. It might even appear that one can improve SNR by scaling the measurement operator in Eq. (7), when in reality the expression in Eq. (5) shows that scaling the measurement operator will cause a proportional scaling in the noise; thus, scaling the compressive measurement operator of a RF system (if physically possible) has no effect on performance. Using Eq. (7) as a starting point makes it too easy to overlook these factors and may lead to improper comparisons of system performance.

Because scaling of the measurement kernel has no effect on system performance, we conveniently assume that each measurement kernel is normalized to unit energy. Further, from the preprojection noise model in Eq. (5) the same noise realization is observed by each of the M measurement kernels. Thus, unlike the case of independent additive noise on each projection, there is no SNR or information gained from a duplicate measurement kernel, and without loss of generality we may assume that the M measurement kernels are orthogonal. Enforcing unit-energy scaling and orthogonality on the measurement kernels, we arrive at an orthonormal constraint on the measurement kernels such that

$$\int \phi_m(\mathbf{r},t)\phi_k^*(\mathbf{r},t)\mathrm{d}\mathbf{r}\mathrm{d}t = \begin{cases} 0 & m \neq k \\ 1 & m = k \end{cases}. \tag{8}$$

Finally, we note that the typical CS model is written in matrix-vector form. The matrix-vector form is obtained by approximating the integral in Eq. (5) with a summation over small intervals in space and time. Thus, the inner product between $\phi_m(\mathbf{r},t)$ and $f(\mathbf{r},t)+n(\mathbf{r},t)$ becomes a matrix multiplication of a row vector containing the discretized sampling kernel and a column vector containing the discretized noisy signal. The mth measurement becomes

$$\begin{aligned} y_m &= \int \phi_m(\mathbf{r},t)[f(\mathbf{r},t)+n(\mathbf{r},t)]\mathrm{d}\mathbf{r}\mathrm{d}t \\ &\approx \phi_m(f+n). \end{aligned} \tag{9}$$

Stacking the discretized versions of all M measurement kernels into the M rows of a matrix Φ, the matrix-vector version of the preprojection noise measurement model is $y = \Phi(f+n)$.

4. System-Level Issues to Consider

In this section, we elaborate on the six system issues enumerated in Section 1. We begin with perhaps the most critical, yet easily overlooked, issue: SNR.

A. Signal-to-Noise Ratio

Regardless of the specific receiver architecture, an essential property of RF systems is that analog-to-digital conversion does not take place until after the propagating wave is transduced to an electrical signal carried on a transmission line. For any receiver operating above zero absolute temperature, the signal carried on the transmission line will be noisy. A low-noise amplifier (LNA) is usually placed early in the receive chain in order to prevent subsequent components from degrading the receiver noise figure, but then any manipulations of the signal in subsequent components are also applied to the noise at the output of the LNA. This has important implications discussed below regarding SNR for compressive RF systems.

One of the major functions of RF systems is signal detection, whether it is detection of bits in a communication signal or detecting the presence of targets in a received radar waveform. Furthermore, signal detection performance is strongly affected by SNR, which is maximized by the matched filter. In fact, detection of a known signal can usually be achieved with optimal or near-optimal performance using a single correctly timed sample at the output of a properly matched filter. If the signal structure is known except for some unknown parameters such as arrival time and/or Doppler shift, then it is possible to construct a bank of filters matched to the signal at different parameter values (i.e., range and/or Doppler bins). However, the typical filter bank approach involves a full set of purposefully *coherent* measurement kernels as opposed to an incomplete set of *incoherent* kernels. This difference causes average SNR loss [19,20], which makes compressive RF systems a questionable choice for any application requiring detection of weak signals.

We have considered the general problem of SNR loss for a set of orthonormal measurement kernels [21]. In our analysis, we model unknown signal parameters as random, which leads to treating the signal being acquired as a random process. By expanding the random signal with a Karhunen–Loeve expansion (KLE) [22], we can express average output SNR in terms of the variance of the KLE coefficients.

Consider the matrix-vector version of the RF CS problem, including expansion of the signal as $f = \Psi x$, where the columns of Ψ are discretized versions of the KLE basis functions. We define the native dimensionality N (i.e., the number of samples required to capture the signal at the Nyquist rate) of $f(\mathbf{r},t)$ as the rank of the signal covariance matrix, $N = \mathrm{rank}\{E[xx^T]\}$, which is also the number of KLE coefficients having nonzero variance. Let the preprojection noise be additive, zero-mean white noise such that $E[n] = 0$ and $E[nn^T] = P_n\mathbf{I}$. We consider two cases in particular. First, we consider a random measurement operator and make no assumptions about the statistical distribution on the coefficients of x. Second, we consider a deterministic measurement operator for the case in which all elements of the KLE vector x have the same variance.

Due to the orthonormal constraint we have placed on the measurement kernels, the output noise power is equal to the input noise power. Therefore, average output SNR depends on the fraction of signal energy that is captured by the $M < N$ measurements. The

output signal samples are $y_s = \Phi\Psi x$, so the expected output signal energy is

$$E[y_s^T y_s] = E[(\Phi\Psi x)^T(\Phi\Psi x)] = \mathrm{Tr}\{E[\Psi^T\Phi^T\Phi\Psi x x^T]\} = \mathrm{Tr}\{E[\Psi^T\Phi^T\Phi\Psi]E[xx^T]\}. \quad (10)$$

From the orthonormality of Φ and of the KLE basis, we can then conclude that the diagonal elements of $E[\Psi^T\Phi^T\Phi\Psi]$ are equal to (M/N). Furthermore, the covariance matrix $E[xx^T]$ is diagonal because KLE coefficients are uncorrelated, which leads to

$$E[y_s^T y_s] = \frac{M}{N}\,\mathrm{Tr}\{E[xx^T]\} = \frac{M}{N}E_x. \quad (11)$$

Random projections capture, on average, (M/N)th of the total available signal energy, resulting in an average SNR loss equal to (N/M), which is the system's compression ratio.

The second case of interest involves deterministic measurement kernels with uniform-variance KLE coefficients. In this case, the coefficient covariance matrix is $E[xx^T] = (E_x/N)\mathbf{I}$, and Eq. (10) becomes

$$E[y_s^T y_s] = (E_x/N)\mathrm{Tr}\{\Psi^T\Phi^T\Phi\Psi\} = (M/N)E_x, \quad (12)$$

where we have again used the orthonormality of Φ and Ψ. When average signal energy is evenly distributed over an N-dimensional space, only (M/N)th of the total available signal energy will be captured, regardless of whether the M measurement kernels are randomly generated or specifically designed. Once again, SNR loss is equal to the system's compression ratio.

As an application example, consider the detection of a radar target at a particular range but with unknown Doppler shift. Let a noncompressive reference system transmit a periodic train of N identical pulses. In contrast, consider a compressive implementation in which a random subset of only $M < N$ pulses is transmitted. Allowing Ψ to be an $N \times N$ manifold of steering vectors for different Doppler shifts (in radar and array processing, a steering vector contains the expected phase shifts due to discrete observations of a sinusoid, such as the rotating Doppler frequency component between multiple pulses or the phase rotation due to signal propagation between antenna elements; a manifold of steering vectors is a matrix in which each column is a steering vector matched to a different frequency), the received slow-time signal is Ψx, where x is a 1-sparse vector with nonzero value in the location corresponding to the target's Doppler shift (assuming the target, if present, is on the grid). The noncompressed noisy observations are then $\Psi x + n$. The sensing matrix Φ for downselecting the number of pulses is an $M \times N$ matrix having all zeros in each row except for a single '1' in a randomly selected column. In order to maintain the orthonormal kernel assumption, each '1' must appear in a different column, but some columns will have all zeros (i.e., they are not retained in the compression). Clearly, the rows of the sensing matrix are orthonormal, and it is clear that the preprojection noise model can reduce to the postprojection noise model for this type of compression kernel. However, by starting with the preprojection noise model, it is also clear that the compressive radar system collects reduced energy from the target because only $M < N$ pulses are collected. It is not possible to increase the power of each transmitted pulse unless one is willing to also increase the power of the noncompressive comparison system. Therefore, SNR is degraded by a factor of (N/M), and any SWaP-C benefits obtained by thinning the measurements must be weighed against the corresponding loss in detection performance. For any RF system that has signal detection as one of its fundamental tasks (and in radar, detecting weak signals in noise is a common task), this SNR loss is likely to be an unacceptable compromise.

B. Resolution Enhancement at High SNR

Intentional RF signal compression introduces SNR loss that severely impacts detection performance. But in some cases, a RF system may be resolution-limited due to SWaP-C constraints on aperture weight, the number of receiver channels, data collection time, or other factors. In these cases, compressive measurement schemes may enable improved precision, which we colloquially refer to as "native resolution."

Consider the Cramèr–Rao bound (CRB) on the time delay estimation variance for a signal $f(t-\tau)$ in white Gaussian noise where τ is unknown. The CRB for this problem [23] satisfies

$$\mathrm{CRB}(\tau) \propto (\mathcal{E}_f B^2)^{-1}, \quad (13)$$

where $\mathcal{E}_f$ denotes SNR and B is the bandwidth of $f(t)$. Therefore, we observe that the CRB for time delay estimation improves linearly with SNR and quadratically with signal bandwidth. Next, consider a case in which the receiver sampling rate is fixed and cannot be increased due to SWaP-C or other constraints. If it is possible to increase the bandwidth of $f(t)$ (it is generally easier to generate higher bandwidths than it is to accurately sample them), then there is potential estimation performance improvement even if the full bandwidth increase cannot be sampled. For example, suppose that the receiver sampling rate matches the original signal bandwidth of B, but that the signal bandwidth is now increased by a factor of K. The receiver is now compressive by a factor of K, which implies a factor of K loss in SNR. Replacing the CRB expression with these new values yields

$$\mathrm{CRB}(\tau) \propto \left[\frac{\mathcal{E}_f}{K}(KB)^2\right]^{-1} = K^{-1}(\mathcal{E}_f B^2)^{-1}. \quad (14)$$

Even though the increased signal bandwidth has forced the need for compressive sampling, the fact that the CRB depends on B^2 means that the overall CRB has improved by a factor equal to the bandwidth increase [24].

This general behavior will hold for other nonlinear estimation problems as well. If compressive implementations enable increased observation time, observation aperture, or bandwidth, then the system's native resolution is increased. And if the system can overcome SNR losses inherent in compressive implementations, then improvement in native resolution can be exploited for better parameter estimation performance. Extrapolating this reasoning to imaging applications, improved native resolution can lead to better imaging performance, but arguably only in the high-SNR asymptotic regime.

Although the CRB line of reasoning is a fundamental one, whether the potential performance improvement is actually realizable and affordable depends on many system-level factors. Obviously, expected SNR is an important factor. The CRB is only a relevant metric in the asymptotic regime where there is sufficient SNR for an estimator to achieve or nearly achieve CRB-level performance. The SNR loss of RF CS systems means that the input SNR (prior to compression) needed to reach the asymptotic region is now higher than for the noncompressive system, but once the asymptotic region is achieved, the estimation error variance can be lower. Therefore, compressive implementations intended to overcome native resolution limitations inherent in finite SWaP-C may only be valid for high-SNR applications requiring very strict imaging and estimation performance. Furthermore, as CS approaches attempt to enhance native resolution at high SNR, dependence on an accurate system model becomes more and more severe. At some point, the calibration accuracy needed to realize the intended performance gains may require such cost and effort that the SWaP-C benefits of a compressive implementation are lost. Compressive implementations may require additional or better analog hardware, such as accurate analog multipliers or wideband amplifiers. The cost, weight, and power of these additional components must be factored into the overall system evaluation in order to determine whether a compressive approach is beneficial. Finally, CRB analyses connote traditional estimation methods such as maximum likelihood estimation, while quantitative evaluations of error variance or mean-squared error have been rare for sparse reconstruction methods. For compressive approaches to see increased application, it is insufficient to qualitatively justify reduced error variance or to demonstrate anecdotal images having higher native resolution. Instead, full system evaluations must be compared and weighed against quantitative assessments of performance indicators that matter to RF system users.

C. Calibration

Nonlinear processing in CS relies on a forward sensor system model through the matrix $A = \Phi\Psi$. The sensing matrix Φ mathematically describes the forward sensing model implemented by the sensing hardware, including the timing and spatial location of samples, and relative gain and phase perturbations across multiple receiver channels. Unfortunately, nonlinearity of real components such as mixers, analog multipliers, amplifiers, and A/D converters cannot be captured in the linear sensing model. The matrix Ψ mathematically describes the waveforms of interest, which might be generated by a separate system as in a communications or intelligence application, or by the same system that possesses the compressive receiver as in a monostatic radar application.

While the performance of any sensor depends on knowledge of its sensing model, this is especially true for model-based techniques that rely on accurate signal manifolds to adaptively cancel interference, accurately estimate parameters, or enforce constraints such as sparsity. Similar to the basis mismatch problem described below, if imperfectly calibrated hardware causes distortion of a signal element away from what the assumed model describes, the solution will lose its sparsity. The scenario worsens with increased subsampling levels as disparities between assumed and actual sensing models become amplified. While some calibration errors can be modeled as additive perturbations on the matrix A [25], nonlinear hardware effects cause unique problems. Unlike the basis mismatch problem, where an off-the-grid signal ceases to be sparse but is still a linear combination of elements in the signal dictionary, nonlinear hardware perturbations cannot be characterized by the assumed linear projection model. Third-order harmonics, signal amplitude compression, and other nonlinear effects are not captured by $y = \Phi\Psi x$ and cannot be modeled by additive noise.

Some theoretical work on calibration errors has been performed, e.g., [25], but calibration and nonlinear hardware errors have not been investigated for their impact on sensor exploitation performance. Indeed, as we have mentioned, quantitative performance and system trade studies are lacking for ideal sensor assumptions, so it is no surprise that they are also lacking for analyses that include system model errors. The ultimate impact of hardware errors will depend on the specific application, compression method, and noise models. For example, we have described above how CS implementations might admit resolution enhancement at high SNR. This enhancement, however, is contingent on precise knowledge of the system model. Whether employing MMSE processing or sparsity-based recovery methods, these algorithms require an accurate signal model. Furthermore, these algorithms are informed with prior knowledge of the measurement accuracy, either through signal and noise statistics or through adjustable error tolerance parameters. At high SNR, it is

implied that all measurements are very accurate, but calibration errors may mean that the measurements are less accurate than believed. Thus, requirements for system model accuracy become more and more strict as SNR increases, yet the high-SNR regime is exactly where RF CS may provide some of its most important benefits.

Nonlinear errors and signal-dependent additive noise will always be present in RF CS systems, and examples of sparse reconstruction with real data have been presented. However, the question of whether RF CS hardware can be characterized in a linear system model with sufficient accuracy to support quantitative performance gains is still an open question.

D. Entropy of Radar Data

Sparsity of an unknown signal—the ability to represent the signal using only a few terms from a fixed incoherent dictionary—is the central premise of CS. However, radar data are, in many applications, not compressible. For example, in air-to-ground radar imaging, raw phase history data typically admit approximately 2:1 lossless compression, and only 8:1 lossy compression. Factors contributing to high entropy scenes include speckle and the scene roughness at the scale of radar wavelengths. The coherent interference of reflectors within a resolution cell gives rise to a random scattered field, causing the speckle phenomenon; the same is true in optics for diffuse objects [26]. Yet, in specific applications, the radar signals may admit sparse representation. Examples include echoes in ground-to-air surveillance, where the goal is to detect reflectors against a zero-return background and high-frequency scattering from man-made objects in a low-clutter environment [27].

E. Nonlinear Processing

CS employs nonlinear processing to reconstruct signals and/or form images from sub-Nyquist data samples. Without the regularizing effect of sparsity-constrained nonlinear processing, the resulting signals or images are typically degraded by high side-lobe levels. However, sparsity-constrained nonlinear processing also introduces artifacts. Although sparse reconstruction may make signal or image quality more visually appealing, the error statistics behave in ways that are poorly handled by automatic exploitation algorithms. Likewise, human image analysts may respond poorly to the artifacts due to nonlinear processing, while understanding and accepting the structured and predictable artifacts of linear processing.

Very little work has yet evaluated exploitation performance with compressive [28,29] or sparsity-constrained [30,31] systems. Here we comment on constant false alarm rate (CFAR) detection. A typical radar system collects data that include interference with unknown power levels and correlation structure. Once data are processed to produce a map of received energy versus parameters such as range and Doppler, thresholds must be derived from the data and applied to each resolution cell. The goal in deriving a threshold is to control the false alarm rate, and hence to control the processing load of post-detection stages. Techniques for CFAR threshholding include cell average methods [32,33], order statistic methods [34], and even normalizations on space-time adaptive filters [35,36] that theoretically produce known output statistics in the target-absent case. However, a common theme in these methods is that they generally operate on processed data (e.g., range-Doppler cells or adaptive filter outputs) that have well-behaved distributions such as Rayleigh or exponential. These distributions are reasonable when it comes to setting thresholds that achieve a desired probability of false alarm. ℓ_1 regularization, on the other hand, produces processed data with sparse distributions (indeed, this is by design), and the relationship between the sparse output statistics, thresholds, and false alarm rates is generally unknown and unpredictable. This scenario wreaks havoc on existing detection thresholding algorithms [28]. While order statistics in CFAR methods allow for some outliers, none of the existing methods are suitable for the sparse densities produced by ℓ_1 regularization.

The CFAR methods that have thus far been explored for compressive RF systems attempt to soften the hard thresholding effects of ℓ_1 regularization [29]. In [28], iterative signal reconstruction generates both thresholded and unthresholded signal reconstructions at each step. The thresholded version of the signal is unsuitable for CFAR detection, while the unthresholded signal is sensitive to algorithm parameters and can have high side-lobes. CFAR algorithms for compressive radar systems are nascent and require maturation before they can be applied in practice.

In addition to detection, classification is a frequent inference goal for RF sensing, and template matching remains a standard processing paradigm. As with detection, poorly understood output statistics degrade the performance of distance metrics in matching. Moreover, scene-dependent parameter selection, such as λ in Eq. (4), has significant impact on both signal and noise properties in reconstructed signatures. Template storage and search are typically system bottlenecks due to an exponentially large variety of target variations [37,38]; therefore, the system design precludes expansion of a template library to include variations in algorithm parameters.

F. Basis Mismatch

Physical parameters relevant to RF sensing, such as delay, Doppler frequency, and angle of arrival, are continuous parameters. The nominal CS theory, on the other hand, requires a known fixed grid of parameter values. The mismatch between grid points and actual parameter values destroys sparsity. The basis mismatch issue has been noted by many authors (e.g., [39,3,4,5]), and the severity of the issue grows

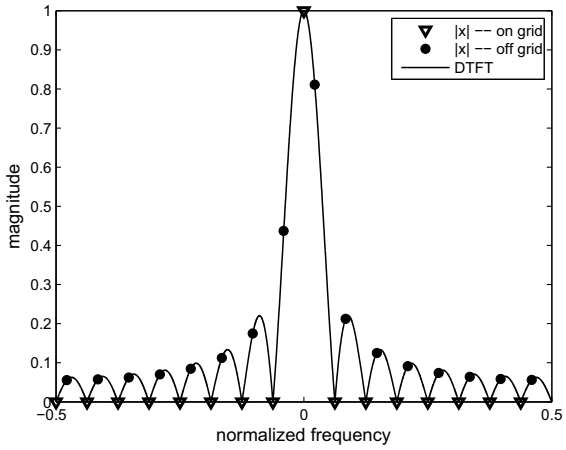

Fig. 3. Illustration of basis mismatch in the DFT basis.

with less sparse signals (larger s) and higher levels of subsampling (larger N/M).

A simple illustration of the basis mismatch "off-the-grid" effect is illustrated in Fig. 3 for a single complex sinusoid with $N = 16$ and a DFT basis. The graph shows the amplitudes of the DFT coefficients in two cases. For the triangular markers, the tone coincides with a grid point and results in a highly sparse exact representation, $s = 1$. In contrast, for the circle markers the tone lies at normalized frequency 0.35/16 between grid points and results in $s = 0.5N$ to achieve even a modest 15 dB SNR in the sparse approximation.

Recent CS literature has partially addressed the off-the-grid issue, for the special case of a superposition of tones. For example, convex minimization of an atomic norm [40], structured low-rank matrix completion [41], and a modified basis pursuit [42] provide near-minimax estimation in the Fourier basis for determining delays or angles of arrival. Proofs rely on the properties of complex exponentials and the associated continuous Fourier dictionary. The results show that convex programming and certain greedy algorithms can provide good reconstructions using oversampled, and hence highly correlated, Fourier dictionaries. However, the results require well-separated tones, and hence *do not* offer the promise of super-resolution. Empirically, however, a modest super-resolution is observed at high SNR, resulting in effective resolution somewhere between the Rayleigh limit and the Cramèr–Rao lower bound [43,41,44].

In addition to these provable performance guarantees, heuristic algorithms have also been proposed for addressing the off-the-grid effect [43,45,46]. For example, in [46] orthogonal matched pursuit is modified to use nonlinear least-squares (NLLS) [47]; at each iteration the maximum correlation between a dictionary element and the residual initializes the NLLS nonconvex local optimization. The technique is experimentally demonstrated in an optical wideband converter whereby a RF signal modulates an optical field, which then passes through a spatial light modulator before transduction by a photodiode.

G. Benchmarking and Precedent

When benchmarking CS results, we note that many elements of recent CS literature have long appeared in RF practice; the absent piece prior to the 1990s [1,48] and vigorously investigated since 2004 was the mathematical foundation of provable performance guarantees. Here, we highlight only a few of the many relevant historical precedents. First, for fast-time sub-Nyquist sampling, so-called "stretch" processing of linear FM chirp waveforms has permitted receiver sampling at rates determined by the scene extent, rather than the signal bandwidth [49,50]. Subsampling by a factor of 2 to 10 is commonplace. Patented hardware implementation appeared under the name "compressive receiver" [51]; stretch processing has been combined with sub-Nyquist analog-to-digital conversion for provable and experimentally demonstrated recovery of sparse signals using a sum-of-chirps waveform [52].

Second, random sampling has for six decades appeared in practice as staggered pulse repetition intervals [50] or in subsampled antenna arrays [53,54,55,56]; the random arrays provide a large effective aperture with reduced hardware costs and tolerable grating lobes. Third, ℓ_1 regularization of linear inverse problems for side-lobe deconvolution has long history in seismic processing [57] and in imaging [58], for example, in minimization of the ℓ_1 norm of gradients [59,60] and for sparse representation in physics-based dictionaries [61]. Recently, more general sparsity-inducing nonquadratic regularization [62,2,63,64,65] has been applied to radar imaging; system-level assessment has shown classification performance improvements on par with a two-times increase in sensor bandwidth [31]. Similarly, a joint sparse signal model [66] has been used for side-lobe suppression in polarimetric radar imaging [67].

Fourth, greedy algorithmic approaches [68] and iteratively reweighted least-squares [69,70] have been employed in optics and RF for the deconvolution of side-lobes in image recovery from sparse and irregularly sampled Fourier data. Maximum entropy [71], as well, has been used as a regularizing criterion in reconstruction of sparse objects from incomplete data. Other sparsity-driven nonlinear processing techniques, including harmonic retrieval and spectrum estimation, are surveyed in [30,72,73]. Fifth, the CS measure of coherence, μ_A, defined in Eq. (1) is fully prescribed by the classical radar *ambiguity function* [74]. RF engineers seek data acquisition designs that yield an ambiguity function with a sharp central peak and low side-lobe levels. The ambiguity function is defined for the continuously valued radar parameters, while coherence μ_A is given by samples of the ambiguity function at locations corresponding to a sampled grid of parameter values; thus, low-coherence measurement designs have a

long history in radar. This connection is detailed in [4].

The themes of sparse reconstruction algorithms and low-coherence measurement operators via pseudo-randomized data acquisition are long-standing concepts in radar and optical processing; the provable recovery guarantees provided by CS theory serve to illuminate and invite the power of purposefully combining these two themes.

H. Implementation Costs

In the case of sub-Nyquist analog-to-digital sampling, additional hardware required must be factored into the overall system evaluation. Sub-Nyquist sampling requires analog multiplication of two wideband signals—the signal being acquired and the measurement kernel—to produce an even wider bandwidth output. This wideband analog multiplication convolves the spectrum of the acquired signal in a known way such that its full spectrum is partially folded into a narrow sampling band prior to reduced-rate analog-to-digital conversion. Microwave mixers are typically designed for optimum operation when the signal at the local oscillator port is a pure tone with very precise power level; therefore, they may not be the best choice for performing the analog multiplication. Of course, the required fidelity in the analog multiplication will be driven by the user's system requirements, but very few studies have mapped analog multiplier accuracy to signal reconstruction accuracy [75,76,77] and ultimately to exploitation performance.

Sub-Nyquist sampling also requires that all analog components support the full signal bandwidth that will ultimately be undersampled. The LNA that appears early in the receiver architecture not only must have the high-gain and low-noise figures that identify it as an LNA, but also must be wideband such that the signal is not distorted prior to compression. If the signal is to be split after the LNA in order to implement multiple compression kernels in parallel [78], then the LNA is critical for setting SNR prior to dividing the signal; even an *ideal* power divider has a noise figure equal to the number of output branches. Given the extra hardware requirements, one must be careful to assess whether a compressive implementation is a better *system* choice than, say, a channelized architecture that uses multiple narrower bands to cover the full bandwidth.

That said, there are some indicators that CS-based sub-Nyquist sampling could be an acceptable system trade. Because the analog multiplier design has not focused on this application, the designs could probably be improved if there were sufficient need. Furthermore, the wideband analog signal must be accurately generated, so sub-Nyquist A/D conversion would not make sense if it cost more to generate the kernel than it would to sample faster. But in a recent evaluation of available commercial components, a 12-bit, 2-GSPS digital-to-analog converter (DAC) needed for generating wideband kernels could be obtained for about $50 and had a power consumption of approximately 1 W. In contrast, a 12-bit, 2-GSPS ADC cost approximately $1,500 and consumed about 3.5 W of power. There are obviously many other system factors to consider, but these types of analyses indicate the potential for compressive A/D conversion to be a reasonable choice.

For compression via thinned array or thinned slow-time pulses, there are no real system costs associated with not taking a measurement, which is why implementations such as thinned arrays have been around long before the recent CS trend. However, there clearly are increased computational costs because ℓ_1 regularization is more computationally intensive than beamforming or correlation-based filter banks. As computational power continues to increase, the trend will favor architectures that save on sampling hardware while shifting the burden to the processing.

As a community, if we want CS techniques to find application in real systems, we must do a better job of considering these types of SWaP-C impacts on the receiver architecture, such that we can weigh system impacts and computational loads against possible performance gains.

5. Promising Benefits

Although the ℓ_1 minimization, greedy algorithms, and randomized sampling present in the CS literature each have a long history in RF sensing, we identify six emerging benefits—both direct and indirect—of CS literature on engineering practice.

First, CS provides impetus for practitioners to seek a provably accurate *convex relaxation* of a nonconvex optimization task when recovering a signal or image from incomplete measurements. An example of this benefit appears in phase retrieval [79,80], where recent CS literature has prompted a new look at a classical problem. In [81], a convex formulation is shown to provide stable phase recovery, and a fast algorithm is given in [82]. Phase retrieval also marks the extension of CS theory from a linear to a nonlinear measurement model. In a similar vein, the classical harmonic retrieval problem [83,84,73] has been revisited with the aid of CS-inspired mathematical constructs. Convex minimization of an atomic norm [40] and structured low-rank matrix completion [41] provide near-minimax estimation in the Fourier basis for determining delays or angles of arrival. The techniques provably offer precision beyond the diffraction limit and empirically may perform well for closely spaced tones; additionally, the techniques seamlessly address the difficult issue of model order selection. The convex formulation provides provable finite-sample performance guarantees.

Second, while radar data in many applications may not be compressible, the *difference of data channels* may indeed have low entropy. This compressive prior is exploited, for example, in radar change detection [85], where a Bayesian formulation is paired with message passing computation to maintain or

even improve detection performance versus traditional techniques, with fewer radar pulses across an aperture. The system-level savings permit multiplexed use of radar resources. Similarly, the sparsity of reflectors in a 3D resolution cell is exploited in multipass radar interferometry [86,87,44] to implement position estimation with precision exceeding the Rayleigh resolution.

Third, CS theory invites the *codesign of data acquisition and data processing* to minimize hardware costs. Such a design is found, for example, in a wideband RF system designed and fabricated for estimation of radar pulse parameters without directly acquiring Nyquist-rate samples [88]. The hardware design includes a four-channel random modulation preintegrator; data are processed using a heuristic chain of estimation and detection steps, with fine parameter grids and cubic interpolation to address basis mismatch. However, wideband noise folds in the analog signal path, as discussed in Section 4.A; the laboratory demonstration does not report power and performance comparisons with existing electronic intelligence solutions. The codesign theme is also present in waveform design and structured illumination [89]; an active RF system allows some flexibility to influence the properties of the Φ matrix.

A fourth emerging benefit of CS literature for radar practice is the proof-of-performance guarantees in *multi-input, multi-output* (MIMO) radar [90,91]. The possible gains of MIMO operation have been a subject of interest and controversy [92]. For R receive and T transmit antennas, a debiased LASSO estimator was shown to provide a log R gain in both resolution and detection threshold versus matched filter processing of random arrays [90], and resolution scales as $1/(RT)$, versus $1/(R+T)$ for uniform arrays.

Fifth, several preliminary works offer optimism for *self-calibrating* systems in which both the unknown signal parameters and the data acquisition model are jointly estimated [2,93,94,95,96]. These joint estimation techniques are characterized by *a priori* assumptions of parsimonious structure in both the signal and the calibration.

Sixth, CS theory and the associated interdisciplinary publications provide several *indirect benefits* to RF engineers. Foremost is the availability of convex programming methods for nonsmooth (e.g., ℓ_1) optimization problems. The toolbox of freely available software is large and includes split-Bregman [97] and approximate message passing [11,12] techniques, among many others. As a simple illustration of the rapid pace of algorithm development, consider that a recent four-year interval witnessed roughly three orders of magnitude acceleration of robust principal component analysis (PCA) algorithms [93]. The algorithmic advances have enabled large-scale problems, such as 3D video in cardiac magnetic resonance imaging, and have mitigated the computational overhead of nonlinear processing. Arguably, the elegance and prominence of early CS literature was a catalytic force for the interdisciplinary efforts producing this wealth of convex optimization tools. Further, the mathematical proofs at the heart of CS challenge algorithm developers to rigorously establish performance guarantees for proposed procedures. Simulation results or claims of asymptotic statistical efficiency are no longer the gold standards.

6. Conclusions

Is CS relevant for RF sensing? We have seen that the answer is a qualified "yes" for the simple reason that CS processing techniques have been present in RF practice for decades. For specific applications in which sufficiently high SNR can be obtained from irregularly sampled apertures, CS processing has a long history, and the recent CS literature provides provable performance guarantees and greatly improved numerical optimization software. The speed and scalability of the optimization software broaden the applicability of regularized signal recovery techniques to new RF approaches; examples discussed include scenarios in which the difference of data channels may exhibit low entropy and the joint recovery of both a signal and hardware calibration parameters.

This paper outlined a set of issues attendant to a balanced and system-level assessment of CS-motivated solutions to RF sensing problems. We believe that as inquiry into CS methods matures, the evolution beyond qualitative, anecdotal results will require the sometimes difficult assessment of system performance metrics including power, cost, system complexity, robustness to system calibration, false alarm rate, and classification rate. Equally challenging can be the benchmarking of performance to relevant existing methods. We have highlighted, through the RF preprojection noise model, the SNR issues that require purposeful consideration in the RF case and differ from some other sensing modalities. Indeed, many RF applications are SNR-limited, and therefore not amenable to the unavoidable SNR loss from CS data acquisition. And, we have noted the difficulties associated with stably estimating the nondifferentiable probability distributions of decision statistics derived from nonlinear CS processing. Broadly, the explosion of CS literature invigorates evaluation of sensor resource management in light of cumulative decades of technological advances in the processing and storage for digital data.

The authors thank Professor James Fienup for identifying several historical precedents: [79,26,59,80,71,58]. Preparation of the manuscript was supported in part by DARPA grants N66001-11-1-4079 and N66001-10-1-4079 and by the Air Force Research Laboratory under FA8650-14-C-1712. The manuscript is based on presentations given by the authors at the OSA Incubator Meeting, "Implications of Compressive Sensing Concepts to Imaging Systems," held April 9–11, 2014, at OSA Headquarters, Washington, D.C.

References

1. Y. Bresler, "*The invention of compressive sensing and recent results: from spectrum-blind sampling and image compression on the fly to new solutions with realistic performance guarantees*" (Plenary, 2012), http://www.eecs.umich.edu/ssp2012/bresler.pdf.
2. M. Çetin, I. Stojanović, N. O. Önhon, K. R. Varshney, S. Samadi, W. C. Karl, and A. S. Willsky, "Sparsity-driven synthetic aperture radar imaging: reconstruction, autofocusing, moving targets, and compressed sensing," IEEE Signal Process. Mag. **31**(4), 27–40 (2014).
3. J. H. G. Ender, "On compressive sensing applied to radar," Signal Process. **90**, 1402–1414 (2010).
4. L. C. Potter, E. Ertin, J. T. Parker, and M. Çetin, "Sparsity and compressed sensing in radar imaging," Proc. IEEE **98**, 1006–1020 (2010).
5. T. Strohmer, "Measure what should be measured: progress and challenges in compressive sensing," IEEE Signal Process. Lett. **19**, 887–893 (2012).
6. S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM J. Sci. Comput. **20**, 33–61 (1998).
7. T. T. Cai, L. Wang, and G. Xu, "Stable recovery of sparse signals and an oracle inequality," IEEE Trans. Inform. Theory **56**, 3516–3522 (2010).
8. R. Tibshirani, "Regression shrinkage and selection via the lasso," J. Roy. Statist. Soc. B **58**, 267–288 (1996).
9. F. B. Lempers, *Posterior Probabilities of Alternative Linear Models* (Rotterdam University, 1971).
10. M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," J. Mach. Learn. Res. **1**, 211–244 (2001).
11. D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," Proc. Natl. Acad. Sci. USA **106**, 18914–18919 (2009).
12. S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2011), pp. 2168–2172.
13. J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond Nyquist: efficient sampling of sparse bandlimited signals," IEEE Trans. Inform. Theory **56**, 520–544 (2010).
14. V. M. Patel, G. R. Easley, Jr., D. M. Healy, and R. Chellappa, "Compressed synthetic aperture radar," IEEE J. Sel. Top. Signal Process. **4**, 244–254 (2010).
15. Y. Gu, N. A. Goodman, and A. Ashok, "Radar target profiling and recognition based on TSI-optimized compressive sensing kernel," IEEE Trans. Signal Process. **62**, 3194–3207 (2014).
16. M. Amin, F. Ahmad, and W. Zhang, "A compressive sensing approach to moving target indication for urban sensing," in *IEEE Radar Conference* (IEEE, 2011), pp. 509–512.
17. A. C. Gurbuz, J. H. McClellan, and W. R. Scott, "A compressive sensing data acquisition and imaging method for stepped frequency GPRs," IEEE Trans. Signal Process. **57**, 2640–2650 (2009).
18. N. A. Goodman, "Measurement structures and constraints in compressive RF systems," in *10th International Conference on Sampling Theory and Applications (SampTA)*, Germany, 2013, pp. 49–52.
19. E. Arias-Castro and Y. C. Eldar, "Noise folding in compressed sensing," IEEE Signal Process. Lett. **18**, 478–481 (2011).
20. M. A. Davenport, J. N. Laska, J. Treichler, and R. G. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: noise folding versus dynamic range," IEEE Trans. Signal Process. **60**, 4628–4642 (2012).
21. B. Pollock and N. A. Goodman, "An examination of the effects of sub-Nyquist sampling on SNR," Proc. SPIE **8365**, 836504 (2012).
22. H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I* (Wiley, 1968).
23. S. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory* (Prentice-Hall, 1993).
24. Y. Gu and N. A. Goodman, "Compressive sensing kernel optimization for time delay estimation," in *IEEE Radar Conference* (IEEE, 2014), pp. 1209–1213.
25. M. A. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressed sensing," IEEE J. Sel. Top. Signal Process. **4**, 342–349 (2010).
26. K. Choi, R. Horisaki, J. Hahn, S. Lim, D. L. Marks, T. J. Schulz, and D. J. Brady, "Compressive holography of diffuse objects," Appl. Opt. **49**, H1–H10 (2010).
27. E. Ertin, C. D. Austin, S. Sharma, R. L. Moses, and L. C. Potter, "Gotcha experience report: three-dimensional SAR imaging with complete circular apertures," Proc. SPIE **6568**, 656802 (2007).
28. L. Anitori, A. Maleki, M. Otten, R. G. Baraniuk, and P. Hoogeboom, "Design and analysis of compressed sensing radar detectors," IEEE Trans. Signal Process. **61**, 813–827 (2013).
29. L. Anitori, M. Otten, W. van Rossum, A. Maleki, and R. Baraniuk, "Compressive CFAR radar detection," in *IEEE Radar Conference* (IEEE, 2012), pp. 320–325.
30. G. R. Benitz, "High-definition vector imaging," Lincoln Lab. J. **10**, 147–170 (1997).
31. M. Çetin, W. C. Karl, and D. A. Castañon, "Feature enhancement and ATR performance using nonquadratic optimization-based SAR imaging," IEEE Trans. Aerosp. Electron. Syst. **39**, 1375–1395 (2003).
32. R. Nitzberg, "Analysis of the arithmetic mean CFAR normalizer for fluctuating targets," IEEE Trans. Aerosp. Electron. Syst. **AES-14**, 44–47 (1978).
33. L. L. Scharf and D. Lytle, "Signal detection in Gaussian noise of unknown level: an invariance application," IEEE Trans. Inform. Theory **17**, 404–411 (1971).
34. H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," IEEE Trans. Aerosp. Electron. Syst. **AES-19**, 608–621 (1983).
35. E. J. Kelly, "An adaptive detection algorithm," IEEE Trans. Aerosp. Electron. Syst. **AES-22**, 115–127 (1986).
36. F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," IEEE Trans. Aerosp. Electron. Syst. **28**, 208–216 (1992).
37. J. M. Irvine, "Some issues with ATRs based on template matching," Proc. SPIE **3584**, 8–16 (1999).
38. J. C. Mossing and T. D. Ross, "Evaluation of SAR ATR algorithm performance sensitivity to MSTAR extended operating conditions," Proc. SPIE **3370**, 554–565 (1998).
39. Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," IEEE Trans. Signal Process. **59**, 2182–2195 (2011).
40. G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," IEEE Trans. Inform. Theory **59**, 7465–7490 (2013).
41. Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," IEEE Trans. Inform. Theory **60**, 6576–6601 (2014).
42. A. Fannjiang and W. Liao, "Mismatch and resolution in compressive imaging," Proc. SPIE **8138**, 81380Y (2011).
43. C. D. Austin, R. L. Moses, J. N. Ash, and E. Ertin, "On the relation between sparse reconstruction and parameter estimation with model order selection," IEEE J. Sel. Top. Signal Process. **4**, 560–570 (2010).
44. X. X. Zhu and R. Bamler, "Superresolving SAR tomography for multidimensional imaging of urban areas: compressive sensing-based tomoSAR inversion," IEEE Signal Process. Mag. **31**(4), 51–58 (2014).
45. D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," IEEE Trans. Signal Process. **53**, 3010–3022 (2005).
46. G. C. Valley, G. A. Sefler, and T. J. Shaw, "Sensing RF signals with the optical wideband converter," Proc. SPIE **8645**, 86450P (2013).
47. P. Stoica and R. Moses, *Spectral Estimation of Signals* (Prentice-Hall, 2005).
48. Y. Bresler and P. Feng, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proceedings of International Conference on Image Processing*, 16–19 September (1996, vol. **1**, pp. 701–704.
49. J. R. Klauder, A. C. Price, S. Darlington, and W. J. Albersheim, "Theory and design of chirp radars," Bell Syst. Tech. J. **39**, 745–808 (1960).

50. M. Skolnik, *Radar Handbook* (McGraw Hill, 2008).
51. C. E. Stromswold, J. T. Apostolos, R. P. Boland, and W. J. Albersheim, "Compressive receiver," U.S. patent 4,305,159 (8 December 1981).
52. E. Ertin, L. C. Potter, and R. L. Moses, "Sparse target recovery performance of multi frequency chirp waveforms," in *Proceedings of the 19th European Signal Processing Conference* (EURASIP, 2011), pp. 446–450.
53. R. F. Harrington, "Sidelobe reduction by nonuniform element spacing," IRE Trans. Antennas Propag. **AP-9**, 187–192 (1961).
54. R. M. Leahy and B. D. Jeffs, "On the design of maximally sparse beamforming arrays," IEEE Trans. Antennas Propag. **39**, 1178–1187 (1991).
55. Y. T. Lo, "A mathematical theory of antenna arrays with randomly spaced elements," IEEE Trans. Antennas Propag. **12**, 257–268 (1964).
56. H. Unz, "Linear arrays with arbitrarily distributed elements," IRE Trans. Antennas Propag. **AP-8**, 222–223 (1960).
57. J. F. Claerbout and F. Muir, "Robust modeling with erratic data," Geophysics **38**, 826–844 (1973).
58. R. Mammone, "Image restoration using linear programming," in *Image Recovery Theory and Application*, H. Stark, ed. (Academic, 1987), Chap. 4.
59. D. C. Dobson and F. Santosa, "An image-enhancement technique for electrical impedance tomography," Inverse Probl. **10**, 317–334 (1994).
60. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Phys. D **60**, 259–268 (1992).
61. J. W. Burns, N. S. Subotic, and D. Pandelis, "Adaptive decomposition in electromagnetics," in *Proceedings of the Antennas and Propagation Society International Symposium* (IEEE, 1997), pp. 1984–1987.
62. M. Çetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," IEEE Trans. Image Process. **10**, 623–631 (2001).
63. R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," Inverse Probl. **24**, 035020 (2008).
64. G. Harikumar and Y. Bresler, "A new algorithm for computing sparse solutions to linear inverse problems," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 7–10 May 1996, vol. **3**, pp. 1331–1334.
65. T. J. Kragh and A. A. Kharbouch, "Monotonic iterative algorithms for SAR image restoration," in *Proceedings of the IEEE International Conference on Image Processing* (IEEE, 2006), pp. 645–648.
66. K. Lee, Y. Bresler, and M. Junge, "Subspace methods for joint sparse recovery," IEEE Trans. Inform. Theory **58**, 3613–3641 (2012).
67. N. Ramakrishnan, E. Ertin, and R. L. Moses, "Enhancement of coupled multichannel images using sparsity constraints," IEEE Trans. Image Process. **19**, 2115–2126 (2010).
68. J. Högbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," Astron. Astrophys. Suppl. Ser. **15**, 417–426 (1974).
69. S. D. Cabrera, B. C. Flores, E. Rodriguez, and G. Thomas, "Two-dimensional extrapolation and spectral estimation from arbitrary sampling configurations for SAR/ISAR imaging," in *Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers* (IEEE, 1994), pp. 145–150.
70. I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," IEEE Trans. Signal Process. **45**, 600–616 (1997).
71. S. F. Gull and G. J. Daniell, "Image reconstruction from incomplete and noisy data," Nature **272**, 686–690 (1978).
72. S. R. DeGraaf, "SAR imaging via modern 2-D spectral estimation methods," IEEE Trans. Image Process. **7**, 729–761 (1998).
73. B. D. Rao and K. S. Arun, "Model based processing of signals: a state space approach," Proc. IEEE **80**, 283–309 (1992).
74. P. M. Woodward, "Radar ambiguity analysis," Tech. Rep. 731 (Royal Radar Establishment, Ministry of Technology, 1967).
75. Z. Charbiwala, P. Martin, and M. B. Srivastava, "Capmux: a scalable analog front end for low power compressed sensing," in *International Green Computing Conference (IGCC)* (IEEE, 2012), pp. 1–10.
76. J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *IEEE International Symposium on Circuits and Systems* (IEEE, 2007), pp. 1959–1962.
77. T. Ragheb, S. Kirolos, J. Laska, A. Gilbert, M. Strauss, R. Baraniuk, and Y. Massoud, "Implementation models for analog-to-information conversion via random sampling," in *50th Midwest Symposium on Circuits and Systems* (IEEE, 2007), pp. 325–328.
78. M. Mishali and Y. C. Eldar, "From theory to practice: sub-Nyquist sampling of sparse wideband analog signals," IEEE J. Sel. Top. Signal Process. **4**, 375–391 (2010).
79. W. H. Bragg, "X-rays and crystalline structure," Science **40**, 795–802 (1914).
80. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21**, 2758–2769 (1982).
81. E. J. Candes, T. Strohmer, and V. Voroninski, "PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming," Commun. Pure Appl. Math. **66**, 1241–1274 (2013).
82. P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," IEEE Trans. Signal Process., doi: 10.1109/TSP.2014.2386294 (to be published).
83. S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao, "State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem," J. Opt. Soc. Am. **73**, 1799–1811 (1983).
84. R. Prony, "Essai éxperimental et analytique," J. de l'École Polytechnique Floréal et Plairial **1**(22), 24–76 (1795).
85. J. N. Ash, "Joint imaging and change detection for robust exploitation in interrupted SAR environments," Proc. SPIE **8746**, J1–J9 (2013).
86. C. D. Austin, E. Ertin, and R. L. Moses, "Sparse signal methods for 3-d radar imaging," IEEE J. Sel. Top. Signal Process. **5**, 408–423 (2011).
87. E. Ertin, R. L. Moses, and L. C. Potter, "Interferometric methods for three-dimensional target reconstruction with multipass circular SAR," IET Radar Sonar Nav. **4**, 464–473 (2010).
88. J. Yoo, C. Turnes, E. B. Nakamura, C. K. Le, S. Becker, E. A. Sovero, M. B. Wakin, M. C. Grant, J. Romberg, A. Emami-Neyestanak, and E. Candes, "A compressed sensing parameter extraction platform for radar pulse signal acquisition," IEEE J. Emerg. Sel. Top. Circuits Syst. **2**, 626–638 (2012).
89. E. Ertin, "Frequency diverse waveforms for compressive radar sensing," in *IEEE International Waveform Diversity and Design Conference (WDD)* (IEEE, 2010), pp. 216–219.
90. T. Strohmer and B. Friedlander, "Analysis of sparse MIMO radar," Appl. Comput. Harmon. Anal. **37**, 361–388 (2014).
91. T. Strohmer and H. Wang, "Accurate imaging of moving targets via random sensor arrays and Kerdock codes," Inverse Probl. **29**, 085001 (2013).
92. F. Daum and J. Huang, "MIMO radar: snake oil or good idea?" IEEE Aerosp. Electron. Syst. Mag. **24**(5), 8–12 (2009).
93. J. Parker, P. Schniter, and V. Cevher, "Bilinear generalized approximate message passing—Part II: applications," IEEE Trans. Signal Process. **62**, 5854–5867 (2014).
94. R. G. Raj, R. Lipps, and A. M. Bottoms, "Sparsity-based image reconstruction techniques for ISAR imaging," in *IEEE Radar Conference* (IEEE, 2014), pp. 974–979.
95. N. Ramakrishnan, E. Ertin, and R. L. Moses, "Gossip-based algorithm for joint signature estimation and node calibration in sensor networks," IEEE J. Sel. Top. Signal Process. **5**, 665–673 (2011).
96. F. Sroubek and P. Milanfar, "Robust multichannel blind deconvolution via fast alternating minimization," IEEE Trans. Image Process. **21**, 1687–1700 (2012).
97. T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," SIAM J. Imaging Sci. **2**, 323–343 (2009).