SPARSE SIGNAL DETECTION FROM INCOHERENT PROJECTIONS

Marco F. Duarte, Mark A. Davenport, Michael B. Wakin, and Richard G. Baraniuk

Rice University

Department of Electrical and Computer Engineering

ABSTRACT

The recently introduced theory of Compressed Sensing (CS) enables the reconstruction or approximation of sparse or compressible signals from a small set of incoherent projections; often the number of projections can be much smaller than the number of Nyquist rate samples. In this paper, we show that the CS framework is *information scalable* to a wide range of statistical inference tasks. In particular, we demonstrate how CS principles can solve *signal detection* problems given incoherent measurements without ever reconstructing the signals involved. We specifically study the case of signal detection in strong inference and noise and propose an Incoherent Detection and Estimation Algorithm (IDEA) based on Matching Pursuit. The number of measurements and computations necessary for successful detection using IDEA is significantly lower than that necessary for successful reconstruction. Simulations show that IDEA is very resilient to strong interference, additive noise, and measurement quantization. When combined with random measurements, IDEA is applicable to a wide range of different signal classes.

1. INTRODUCTION

Over the past decades the amount of data generated by sensing systems has grown from a trickle to a torrent. This has stimulated much research in the fields of compression and coding, which enable compact storage and rapid transmission of large amounts of information. Compression is possible because often we have considerable a priori information about the signals of interest. For example, many signals are known to have a *sparse* representation in some transform basis (Fourier, DCT, wavelets, etc.) and can be expressed or approximated using a linear combination of only a small set of basis vectors.

The traditional approach to compressing a sparse signal is to compute its transform coefficients and then store or transmit the few large coefficients and their locations. This is an inherently wasteful process (in terms of both sampling rate and computational complexity), since it forces the sensor to acquire and process the entire signal even though an exact representation is not ultimately required. For instance, in many signal processing applications (including most communications and many radar systems), signals are acquired only for the purpose of making a detection or classification decision.

A new framework for simultaneous sensing and compression has developed recently under the rubric of *Compressed Sensing* (CS). CS enables a potentially large reduction in the sampling and computation costs at a sensor. CS builds on the work of Candès, Romberg, and Tao [1] and Donoho [2], who showed that a signal having a sparse representation in one basis can be reconstructed from a small set of projections onto a second, measurement basis that is *incoherent* with the first.[1] Interestingly, *random projections* are a *univer-

---

[1]Roughly speaking, *incoherence* means that no element of one basis has a sparse representation in terms of the other basis.

sal measurement basis in the sense that they are incoherent with any other fixed basis with high probability. The CS measurement process is nonadaptive; the reconstruction process is nonlinear. A variety of reconstruction algorithms have been proposed [1–3].

While the CS literature has focused almost exclusively on problems in signal reconstruction, approximation, and estimation in noise [1–7], our aim in this paper is to show that the CS framework is *information scalable* to a much wider range of statistical inference tasks. Tasks such as detection do not require a reconstruction of the signal, but only require estimates of the relevant *sufficient statistics* for the problem at hand. Our key finding is that it is possible to directly extract these statistics from a small number of random projections without ever reconstructing the signal. The two upshots are that (*i*) significantly fewer measurements are required for signal detection than for signal reconstruction (3× fewer in the example in Figure 1), and (*ii*) the computational complexity of detection is much reduced compared to reconstruction (4× lower in the example in Figure 1). Both of these bode well for many applications. As in reconstruction, random measurements are *universal*, in the sense that with high probability the sufficient statistics can be extracted from them regardless of the signal structure.

This paper is organized as follows. Section 2 provides background on CS. Section 3 states our detection problem and proposes a greedy algorithm for CS detection. Section 4 presents a case study involving wideband signal detection in narrowband interference. Section 5 gives ideas for extensions to classification, and Section 6 concludes with directions for future work.

2. COMPRESSED SENSING BACKGROUND

Let $x \in \mathbb{R}^N$ be a signal and let the matrix $\mathbf{\Psi} := [\psi_1, \psi_2, \ldots, \psi_Z]$ have columns that form a *dictionary* of vectors in $\mathbb{R}^N$. (This dictionary could be a basis or a redundant frame.) When we say that x is *K-sparse*, we mean that it is well approximated by a linear combination of K vectors from $\mathbf{\Psi}$; that is, $x \approx \sum_{i=1}^{K} \theta_{n_i} \psi_{n_i}$ with $K \ll N$.

2.1. Incoherent measurements

Consider a signal x that is K-sparse in $\mathbf{\Psi}$. Consider also an $M \times N$ *measurement* matrix $\mathbf{\Phi}$, $M \ll N$, where the rows of $\mathbf{\Phi}$ are incoherent with the columns of $\mathbf{\Psi}$. For example, let $\mathbf{\Phi}$ contain i.i.d. Gaussian entries; such a matrix is incoherent with any fixed $\mathbf{\Psi}$ with high probability (universality). Compute the *measurements* $y = \mathbf{\Phi}x$ and note that $y \in \mathbb{R}^M$ with $M \ll N$. The CS theory states that there exists an overmeasuring factor $c > 1$ such that only $M := cK$ incoherent measurements y are required to reconstruct x with high probability [1, 2]. That is, just cK incoherent measurements encode all of the salient information in the K-sparse signal x.

2.2. Reconstruction from incoherent projections

The amount of overmeasuring required depends on the (nonlinear) reconstruction algorithm. Most of the existing literature on CS [1, 2, 4, 6, 8] has concentrated on optimization-based methods for signal
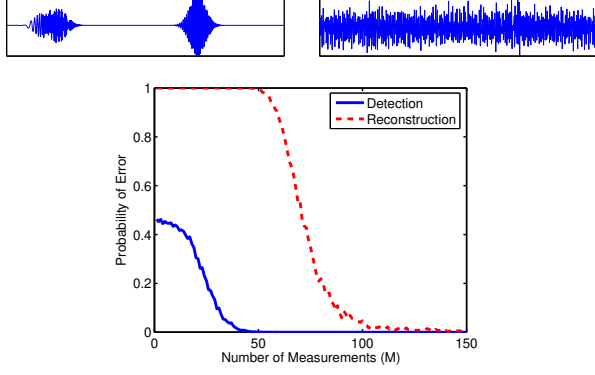
Fig. 1. *(Top) Sample wideband* chirp *signal and same chirp embedded in strong narrowband interference. (Bottom) Probability of error to reconstruct and detect chirp signals embedded in strong sinusoidal interference* (SIR $= -6$ dB) *using greedy algorithms. In this case, detection requires* $3\times$ *fewer measurements and* $4\times$ *fewer computations than reconstruction for an equivalent probability of success.*

recovery, in particular ℓ_1 minimization. The ℓ_1 approach seeks a set of sparse coefficients $\widehat{\theta}$ by solving the linear program [9]

$$\widehat{\theta} = \arg \min_{\theta} \|\theta\|_1 \quad \text{subject to} \quad \mathbf{V}\theta = y,$$

where $\mathbf{V} = \mathbf{\Phi}\mathbf{\Psi}$ is the *holographic* basis. Greedy reconstruction algorithms build up a signal approximation iteratively by making locally optimal decisions [3,7,10]. In this paper, we focus on Matching Pursuit (MP) [11], an efficient greedy algorithm that selects basis vectors one-by-one from a dictionary to optimize the signal approximation at each step. In its application to CS, MP seeks a sparse representation of the measurement vector y in the dictionary $\{v_i\}$ consisting of column vectors from the holographic basis $\mathbf{V}$.

MP ALGORITHM FOR CS RECONSTRUCTION
1. Initialize the residual $r_0 = y$ and the approximation $\widehat{\theta} = 0$, $\widehat{\theta} \in \mathbb{R}^Z$. Initialize the iteration counter $t = 1$.
2. Select the dictionary vector that maximizes the value of the projection of the residual onto $\mathbf{V}$

$$n_t = \arg \max_{i=1,\ldots,Z} \frac{\langle r_{t-1}, v_i\rangle}{\|v_i\|}. \tag{1}$$

3. Update the residual and the estimate of the coefficient for the selected vector

$$r_t = r_{t-1} - \frac{\langle r_{t-1}, v_{n_t}\rangle}{\|v_{n_t}\|^2}v_{n_t}, \tag{2}$$

$$\widehat{\theta}_{n_t} = \widehat{\theta}_{n_t} + \frac{\langle r_{t-1}, v_{n_t}\rangle}{\|v_{n_t}\|^2}. \tag{3}$$

4. Increment t. If $t < T$ and $\|r_t\|_2 > \epsilon\|y\|_2$, then go to Step 2; otherwise, go to Step 5.
5. Obtain the signal estimate $\widehat{x} = \mathbf{\Psi}\widehat{\theta}$.

The parameter ϵ sets the target error level for convergence, and T sets the maximum number of algorithm steps.

2.3. (Dis)advantages
The implications of CS are very promising. Instead of sampling a sparse signal N times and then compressing it, only $cK \ll N$ incoherent measurements suffice. CS with random measurements is advantageous for low-power and low-complexity sensors (such as in sensor networks) because it integrates sampling, compression and

encryption of many different kinds of signals [6, 7]. However, several significant challenges to CS-based signal reconstruction remain. In particular, (*i*) the overmeasuring factor c required for perfect reconstruction can be quite large, typically $c \approx \log_2(1 + N/K)$ for linear programming based reconstruction [7]; (*ii*) the computational complexity of a linear program or greedy algorithm for signal reconstruction is high, at least cubic in the signal length N for a linear program. Greedy algorithms use fewer computations, but require an even larger factor c.

3. GREEDY DETECTION FOR SPARSE SIGNALS

While the CS literature has focused almost exclusively on signal reconstruction, we now show that incoherent measurements can also be used to solve signal detection problems without ever reconstructing the signal. In the process, we will be able to save significantly on both the number of measurements and computational complexity (see Figure 1).

3.1. Detection problem setup
Suppose that the dictionary $\mathbf{\Psi}$ contains particular elements of interest that we wish to detect as components of x. For example, we might want to detect smooth signals, and $\mathbf{\Psi}$ might contain a basis of sinusoids or orthogonal polynomials. Let $\Omega \subset \{1, 2, \ldots, Z\}$ denote the set of *target indices* that represent these components of interest, and let $\mathbf{\Psi}_\Omega$ and θ_Ω denote the corresponding restrictions of the dictionary and coefficients, respectively. Given a set of incoherent measurements $y = \mathbf{\Phi}x$, we aim to determine whether or not x was generated using any of the target components in $\mathbf{\Psi}_\Omega$. That is, we must decide between two composite hypotheses:

$$\mathcal{H}_0 : \theta_\Omega = 0 \quad \text{vs.} \quad \mathcal{H}_1 : \theta_\Omega \neq 0. \tag{4}$$

3.2. Sparse signal detection
To begin, suppose that the signal x is provided directly. If $\mathbf{\Psi}$ is an orthonormal basis, then the solution to (4) can be obtained easily by *matched filtering*. That is, we can directly compute θ_Ω using inner products of x against the columns of $\mathbf{\Psi}_\Omega$ and then compare the components of θ_Ω to a threshold.

However, if $\mathbf{\Psi}$ is not orthogonal or if $\mathbf{\Psi}$ is a *redundant* dictionary, then inner products of x against the columns of $\mathbf{\Psi}_\Omega$ will not suffice for detection. Indeed, if $\mathbf{\Psi}$ is redundant then there likely will be infinitely many θ_Ω consistent with x. Difficulties also arise because the columns of $\mathbf{\Psi}$ are correlated with each other, so that the presence of one *interferes* with the detection of another.

This is analogous to *multiuser detection*, a classical problem in communications that is known to be NP-hard [12]. A practical iterative decoding algorithm, known as *successive cancelation* or *onion peeling*, is very similar in spirit to MP. These algorithms identify the strongest component of $\mathbf{\Psi}$ in x, remove it from x, and then proceed to find the next strongest component. Essentially, this invokes a model for x, namely that it has a sparse expansion in $\mathbf{\Psi}$. This suggests that for our detection problem we should employ a greedy algorithm such as MP from Section 2.2. We can then look for significant energy among the coefficients θ_Ω.

Thus, we now assume that instead of x we observe $y = \mathbf{\Phi}x$. In this case, y will have the same linear expansion among the columns of $\mathbf{V}$ that x has among the columns of $\mathbf{\Psi}$. This strongly motivates an MP approach to solving the sparse detection problem with incoherent measurements. In particular, we can seek to extract the *sufficient statistics* of interest from a small number of incoherent projections. As in CS reconstruction, random measurements provide in some sense a universal representation of the sufficient statistics for a wide range of signal classes.

It is important to note that, just as in any detection problem, it is not necessary to reconstruct precise *values* for the expansion coefficients θ_Ω. Rather, we generally only need to know whether there is a significant *contribution* from these elements. Moreover, there is *no* requirement to accurately reconstruct the coefficients θ_{Ω^c}. This allows us to reduce considerably the number of measurements and computations required when detecting compared to reconstructing.

3.3. Incoherent detection and estimation algorithm (IDEA)

Based on the above motivation, we propose the Incoherent Detection and Estimation Algorithm (IDEA) for signals hidden in incoherent measurements. IDEA is based on the MP reconstruction algorithm from Section 2.2 with two important modifications. First, we set the number of iterations T to be much smaller than necessary for reconstruction. Second, we replace Step 5 with the following:

5. If $\|\widehat{\theta}_\Omega\|_\infty$ exceeds a threshold γ, detect $\mathcal{H}_1$; else choose $\mathcal{H}_0$.

Due to the smaller T, the vector $\widehat{\theta}$ might not accurately reconstruct the signal. However, it may still contain sufficient information for detection. Our detection decision is made simply by examining the components $\widehat{\theta}_\Omega$ and comparing the maximum coefficient to the threshold γ. We will see in the next section that, despite potential imprecision in $\widehat{\theta}$, the detection decision can be remarkably accurate. Indeed, the detection process can succeed even when M is far too small to recover x. Thus, the number of measurements can be scaled back significantly if detection, rather than reconstruction, is the ultimate goal.

4. CASE STUDY

4.1. Dictionary-based detection

IDEA is very well suited to detecting signals in the presence of interference and noise when the signals and interference can be sparsely represented in distinct, incoherent dictionaries. We formalize the problem as follows. We aim to distinguish between two hypotheses

$$\mathcal{H}_0: \quad x = n + \omega \quad \text{vs.} \quad \mathcal{H}_1: \quad x = s + n + \omega,$$

where s denotes the signal of interest (from some class of signals), n denotes the interference, and ω denotes additive white Gaussian noise with $\omega \sim \mathcal{N}(0, \sigma_\omega^2 \mathbf{I})$. Each component is sparse in some dictionary; that is, $s = \mathbf{\Psi}_s \theta_s$, $\|\theta_s\|_0 = K_s$,[2] and $n = \mathbf{\Psi}_n \theta_n$, $\|\theta_n\|_0 = K_n$; however, the noise is not sparse in either dictionary. We can restate the detection problem in terms of the *concatenated dictionaries* and coefficients, writing

$$x = [\mathbf{\Psi}_s \ \mathbf{\Psi}_n] \left[\begin{array}{c} \theta_s \\ \theta_n \end{array} \right] + \omega =: \mathbf{\Psi}\theta + \omega.$$

Now, from the measurements $y = \mathbf{\Phi}x$ (with $\mathbf{\Phi}$ incoherent with both Ψ_s and Ψ_n), we must distinguish between the two hypotheses

$$\mathcal{H}_0: \quad \theta_s = 0 \quad \text{vs.} \quad \mathcal{H}_1: \quad \theta_s \neq 0.$$

We set Ω in IDEA such that $\theta_\Omega = \theta_s$ to obtain detection decisions.

IDEA offers several advantages in this detection scenario. First, the sparsest approximation of y will tend to correctly separate the contributions from the signal and interference components thanks to the incoherency of the two dictionaries. Second, the additive noise is attenuated during sparse approximation since its energy is distributed over all of the expansion coefficients [11].

4.2. Wideband signals in strong narrowband interference

As a concrete example, we study the problem of detecting from random measurements the presence of weak wideband signals corrupted by strong interfering narrowband sources and additive noise.[3] This is a potentially difficult problem: The weakness of the wideband signal precludes an energy detection approach, and if the wideband and narrowband signals overlap in the frequency domain, then bandpass interference suppression may damage the signal beyond detectability. We seek to detect wideband signals that are frequency modulated *chirps*. Chirps are sparsely represented in a *chirplet* dictionary [13] that is incoherent with the Fourier basis that sparsifies narrowband signals. Hence, we can apply IDEA directly. We choose a chirplet dictionary for $\mathbf{\Psi}_s$ and the Fourier basis for $\mathbf{\Psi}_n$.

4.3. Simulations

We set the signal length to $N = 1024$ and construct a 432-element normalized chirplet dictionary consisting of 64-sample chirplets having 16 start times, 9 starting frequencies, and 3 chirp rates. When present, we generate the chirp signal $s = \mathbf{\Psi}_s \theta_s$ with $K_s = 5$, and we assign $\mathcal{N}(0, \sigma_s^2)$ coefficients to the nonzero elements of θ_s. For the interference we set $K_n = 6$ and assign $\mathcal{N}(0, \sigma_n^2)$ coefficients to its nonzero elements. The $M \times N$ measurement matrix $\mathbf{\Phi}$ contains i.i.d. $\mathcal{N}(0, 1)$ entries. Since the number of measurements required for signal reconstruction is proportional to $K_s + K_n$, the detection results will extend directly to other sparsity levels when the number of measurements increases appropriately.

Detection vs. reconstruction: Given the measurements $y = \mathbf{\Phi}x$, we attempt to reconstruct x using MP; the probability of a reconstruction error as a function of the number of measurements M (averaged over 10,000 trials) is given in Figure 1. We define an error as failing to achieve a sufficiently small reconstruction error in the wideband signal s; hence $P_e = \Pr(\|s - \widehat{s}\|_2 > 10^{-3}\|s\|_2)$. Given the same measurements, we also attempt to detect the presence of a wideband component s; the probability of a detection error as a function of M (averaged over 10,000 trials) is also given in Figure 1. We use IDEA with $T = 15$ and $\epsilon = 0$ (we do not check for convergence) and set $\Pr(\mathcal{H}_0) = \Pr(\mathcal{H}_1) = 1/2$. We choose γ to minimize P_e based on Monte Carlo simulations. The chirps are embedded in strong interference; Figure 1 features Signal-to-Interference Ratio $\text{SIR} := 10\log_{10}(\sigma_s^2/\sigma_n^2) = -6$ dB and $\sigma_\omega = 0$. We see that low-P_e signal *detection* requires only about 50 measurements, while low-P_e *reconstruction* requires about 150 measurements. Moreover, each MP detection requires approximately $4\times$ fewer iterations than MP reconstruction. We note that a target P_e can be achieved with fewer iterations by obtaining more measurements, thus providing a valuable tradeoff.

Effect of interference: We now focus exclusively on detection performance. Figure 2(a) illustrates P_e as a function of M for several SIRs. For $M < 50$, detection performance degrades with the SIR. However, $M > 50$, detection performance remains largely unaffected. We believe that this is due to the general robustness of CS recovery – for $M > 50$ there seems to be enough information in the measurements to accurately estimate the interference components (those with the most energy). However, with few measurements, some of the interference energy is incorrectly assigned to the signal components, which corrupts performance.

Effect of noise: IDEA shares the same robustness to additive white Gaussian noise as CS reconstruction [1]. Figure 2(b) illustrates the detection performance in noise for different levels of the Signal-to-Noise Ratio (SNR) at the fixed $\text{SIR} = -6$ dB. We see a graceful performance degradation as the SNR decreases; indeed, when the power of the noise becomes comparable to that of the signal to be detected, most detection methods suffer.

---

[2]The ℓ_0 "norm" $\|\theta\|_0$ merely counts the number of nonzero components in the vector θ.

[3]Thanks to Dennis Healy for suggesting this case study and for many interesting discussions.
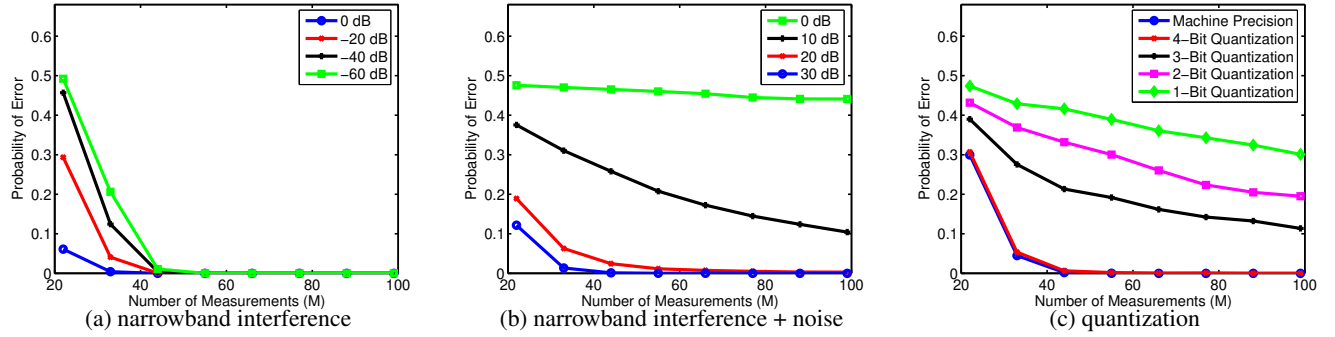
Fig. 2. *Performance of wideband chirp detection. (a) Strong narrowband interference; each curve is a different SIR. (b) Narrowband and white noise interference;* $\mathrm{SIR} = -6$ *dB, and each curve is a different SNR. (c) Effect of measurement quantization at* $\mathrm{SIR} = -20$ *dB; each curve is a different number of quantization levels. In all cases, detection performance degrades gracefully as the distortions increase.*

Effect of quantization: Figure 2(c) illustrates the detection performance for different levels of quantization of the measurements, with a fixed $\mathrm{SIR} = -20$ dB and no noise. Note in particular that the detection performance is remarkably robust with 4-bit (16 level) quantization; we expect the acceptable level of quantization to be dependent on the SIR and SNR of the received signal.

5. EXTENSIONS TO CLASSIFICATION

The properties of incoherent measurements allow us to formulate a simple algorithm for *sparse signal classification*. Consider a signal of interest x of length N and sparsity $\leq K$ in one of C bases (or dictionaries). Each basis represents a signal class. Assume that the different bases are incoherent with each other. Our goal in this classification problem is to determine which class best fits the signal. If the signal were available, then we could perform sparse approximation using each basis and then choose the class giving the sparsest representation. This would require all N signal samples to make the decision.

However, thanks to their universality, one set of cK random measurements suffices to find the sparsest representation of x from the C classes. This problem can be solved using a greedy algorithm. Orthogonal Matching Pursuit (OMP) [3] tracks the signal sparsity as it proceeds; the number of OMP iterations equals the sparsity of x in the corresponding basis. Therefore, by simultaneously running OMPs with each of the C bases, we can assign the signal to the class whose OMP iteration terminates first. The incoherence between the bases guarantees that only one class will have a sparse representation for the signal. Another option would be to run either OMP or MP for a (small) fixed number of iterations with each basis and then assign the signal to the class resulting in the smallest residual. As in IDEA, we expect that we can even reduce the number of measurements below the cK required for high-quality sparse approximation.

6. CONCLUSIONS

IDEA provides reliable detection performance from just a few incoherent measurements when the signals of interest are sparse or compressible in some basis or dictionary. In addition to its efficiency gains over CS reconstruction in terms of the number of measurements and computations required, IDEA shares the many known benefits of CS reconstruction, including the *universality* of random measurements, progressivity, and resilience to noise and quantization [1–7].

There are many opportunities for future research. The extension of IDEA to other signal + interference scenarios is straightfor-

ward. When the sparse signal decomposition can be parameterized, i.e., when each signal dictionary vector $\psi_i = f(\beta_i)$ with β a parameter vector, the CS methodology enables new algorithms for parameter estimation and other statistical signal processing tasks. An especially promising application is in CS acquisition of streaming signals; detection experiments with the random filtering approach of [14] found little to no performance degradation for streaming signals. Future work includes theoretical analysis on the required number of measurements for reliable detection analogous to similar results for reconstruction (see [7] and the references therein).

7. REFERENCES

[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," 2004, Preprint.

[2] D. Donoho, "Compressed sensing," 2004, Preprint.

[3] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," 2005, Preprint.

[4] E. Candès and T. Tao, "Near optimal signal recovery from random projections and universal encoding strategies," 2004, Preprint.

[5] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," 2005, Preprint.

[6] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," 2005, Preprint.

[7] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," 2005, Preprint.

[8] D. Donoho and Y. Tsaig, "Extensions of compressed sensing," 2004, Preprint.

[9] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comp.*, vol. 20, no. 1, pp. 33–61, 1998.

[10] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Fast reconstruction of piecewise smooth signals from random projections," in *Online Proc. of Workshop on Signal Proc. with Adaptive Sparse Structured Representations (SPARS)*, 2005.

[11] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, 1993.

[12] S. Verdú, *Multiuser Detection*, Cambridge Univ. Press, 1998.

[13] R. G. Baraniuk and D. L. Jones, "Shear madness: New orthogonal bases and frames using chirp functions," *IEEE Trans. Signal Proc.*, vol. 41, no. 12, pp. 3543–3549, 1993.

[14] J. A. Tropp, M. B. Wakin, M. F. Duarte, D. Baron, and R. G. Baraniuk, "Random filters for compressive sampling and reconstruction," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, 2006.