The Compressive Multiplexer for Multi-Channel Compressive Sensing

J. P. Slavinsky,[r] Jason N. Laska,[r] Mark A. Davenport,[s] and Richard G. Baraniuk[r]

[r] Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005

[s] Department of Statistics, Stanford University, Stanford, CA 94305

ABSTRACT

The recently developed compressive sensing (CS) framework enables the design of sub-Nyquist analog-to-digital converters. Several architectures have been proposed for the acquisition of sparse signals in large swaths of bandwidth. In this paper we consider a more flexible *multi-channel* signal model consisting of several discontiguous channels where the occupancy of the combined bandwidth of the channels is sparse. We introduce a new compressive acquisition architecture, the compressive multiplexer (CMUX), to sample such signals. We demonstrate that our architecture is CS-feasible and suggest a simple implementation with numerous practical advantages.

Index Terms— compressive sensing, spectrum sensing, random demodulator, multiplexing, multichannel separation

1. INTRODUCTION

With modern signal processing firmly rooted in digital computation, efficient conversion from analog signals to digital representations is of fundamental importance. In recent years, the demand to acquire data at ever increasing bandwidths has imposed a burden on traditional analog-to-digital converters (ADCs) that rely on the Shannon-Nyquist sampling theorem.

The desire to circumvent the Shannon-Nyquist limitation has prompted a new signal acquisition framework, compressive sensing (CS), that demonstrates that signals can be acquired at lower sampling rates when they have have relatively few degrees of freedom. Furthermore, CS enables unique hardware architectures, since it espouses randomized linear sampling systems and non-linear, computational signal recovery.

This theory has spawned a number of new sub-Nyquist sampling architectures including the random demodulator (RD) [1], modulated wideband converter (MWC) [2], and random convolution [3], among others. The aim of these systems is to sample a wide swath of bandwidth at a rate significantly lower than twice its bandwidth, yet recover the salient signals within that bandwidth.

Our goal is to monitor multiple channels at a given time, for example, multiple spectral bands. The systems above can tackle this application by acquiring the entire bandwidth of interest. However, this approach is suboptimal in that these systems do not consider the case where channels of interest are discontiguous and possibly known *a priori*.

In this paper we propose a new architecture that we dub the *compressive multiplexer* (CMUX) for sub-Nyquist acquisition of multiple, discontiguous channels. Our design builds upon the conventional CS framework of Candes, Romberg, Tao, and Donoho [4–6] and provably satisfies the requirements set forth in that framework. Our architecture exhibits additional hallmarks: it uses only a single ADC; it can be assembled from easily available components; it requires simpler calibration than previous architectures; it reduces the rate of all components to below the Nyquist rate for the total acquired bandwidth; and it easily lends itself to specialized algorithms thanks to its structure. For the remainder of this paper, we will focus on the specific example of radio frequency (RF) channels for concreteness.

The CMUX effectively codes each channel with a near-orthogonal code and then combines the coded channels together in the following way. Each channel of interest is mixed down to baseband using a conventional RF analog tuner. We then randomly modulate the output of each channel by a low-rate, pseudo-random, "chipping" sequence of ± 1s. The chipping rate is greater than or equal to the Nyquist rate of any one channel. The randomly modulated channels are then summed together and sampled at the chipping rate.

To unmix the channels, we perform sparse recovery as developed in the CS framework. For the CMUX context, the CS recovery procedure can be thought of as *multi-channel separation* [7–9].

Unlike previous architectures, such as the RD and the MWC, our system does not require calibration of an analog low-pass filter or integrator. Rather, basic calibration can be simply achieved via the knowledge of a few resistor values. Also, unlike other parallel architectures, such as [2, 10], the CMUX only requires one ADC, rather than one per channel.

This paper is organized as follows. In Section 2 we review the CS framework and briefly discuss related CS architectures for sensing the RF spectrum. In Section 3 we introduce the CMUX architecture and propose an implementation that highlights the practical benefits of this new design. In Section 4 we discuss recovery algorithms that exploit the special structure of the CMUX. In Section 5 we demonstrate the simulated performance of the CMUX in both ideal and non-ideal scenarios. In Section 6 we conclude with a discussion about the connections to communication and adaptive sensing techniques.

2. BACKGROUND

2.1. Compressive sensing

CS is a new approach to signal sampling that aims to reproduce Shannon-Nyquist performance using a smaller number of samples. Over some fixed interval of time (or region of space), we obtain the CS measurements as

$$y = \overline{\Phi} x_a + e = \Phi x + e \tag{1}$$

where $y \in \mathbb{R}^M$, $e \in \mathbb{R}^M$ is a noise term, $\overline{\Phi}$ is the sampling operator, and Φ is the equivalent $M \times N$ matrix that operates on the vector x of N Nyquist-rate samples. We can often write $x = \Psi\alpha$ where α is K-sparse, i.e., it has K non-zero entries, and Ψ some transform matrix, in which case we consider the combined matrix $A = \Phi\Psi$.

To ensure stable recovery, we turn to the *restricted isometry property* (RIP), introduced by Candès and Tao [4]. We say that a matrix A satisfies the RIP of order K if there exists a constant, $\delta \in (0, 1)$, such that

$$(1 - \delta)\|\alpha\|_2^2 \leq \|A\alpha\|_2^2 \leq (1 + \delta)\|\alpha\|_2^2, \tag{2}$$

holds for all K-sparse α. In words, A acts as an approximate isometry on the set of vectors that are K-sparse. Provided that A satisfies the RIP of order $2K$ with $\delta \leq \sqrt{2} - 1$, *Basis Pursuit De-Noising* (BPDN),

$$\widehat{\alpha} = \arg\min_{\alpha} \ \|\alpha\|_1 \ \text{ s.t. } \|A\alpha - y\|_2 \leq \epsilon \tag{3}$$

The authors would like to thank Justin Romberg for his advice and insight on this problem. M.D. was supported by NSF DMS-1004718. J.S., J.L., and R.B. were supported by the grants NSF CCF-0431150, CCF-0728867, CCF-0926127, CNS-0435425, and CNS-0520280, DARPA/ONR N66001-08-1-2065, ONR N00014-07-1-0936, N00014-08-1-1067, N00014-08-1-1112, and N00014-08-1-1066, AFOSR FA9550-07-1-0301 and FA9550-09-1-0432, ARO MURI W911NF-07-1-0185 and W911NF-09-1-0383, and the Texas Instruments Leadership University Program.
email: {jps, laska, richb}@rice.edu, markad@stanford.edu — web: dsp.rice.edu

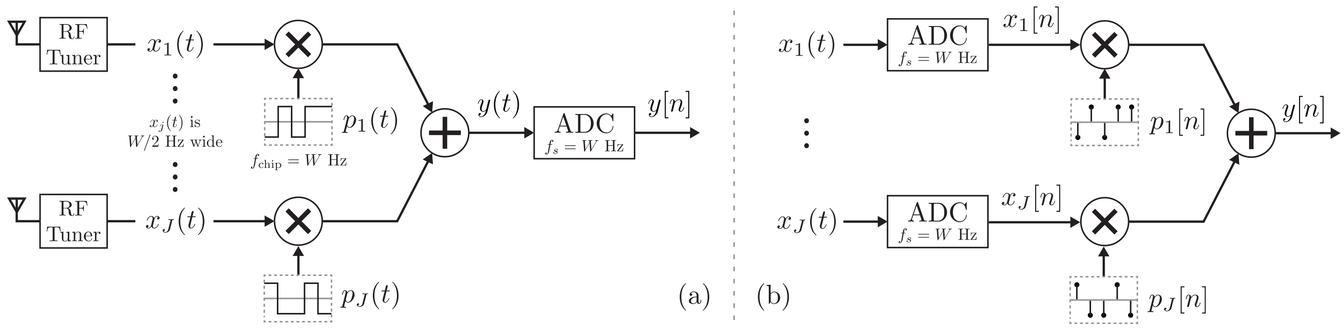

Fig. 1. (a) CMUX system diagram. Each of the J channels is spread by a different chipping sequence, and then summed and sampled. (b) CMUX equivalent system. The sampling operation is moved to the front of the system for the sake of analysis.

will yield a recovered signal $\widehat{\alpha}$ that satisfies $\|\widehat{\alpha} - \alpha\|_2 \leq C_0\epsilon$ provided that $\|e\|_2 \leq \epsilon$, where $C_0 > 1$ is a constant depending only on δ [5]. While convex optimization techniques like BPDN are powerful methods for CS signal recovery, there also exist a variety of alternative algorithms, such as greedy algorithms, that are used in practice and have comparable performance guarantees.

A key theoretical CS result is that by drawing only $M = O(K\log(N/K))$ *random* rows, we obtain a Φ that satisfies the RIP of order $2K$ with high probability (and thus A has RIP if Ψ is an orthonormal basis). As we will see, similar guarantees are possible for highly structured measurement systems as well.

2.2. Compressive sampling architectures

Several hardware architectures have been proposed and implemented to perform CS in practical settings with analog signals. Selected examples include the random demodulator (RD), random filtering, random convolution, and the modulated wideband converter [1–3]. These systems aim to capture a large portion of bandwidth with fewer samples than Shannon would prescribe.

We briefly describe the RD as an example of such a system [1]. The input analog signal is modulated by a ± 1 "chipping sequence" operating at or above the Nyquist rate and integrated. The output of the integrator is sampled, and the integrator is reset after each sample. The ideal integrator with reset can be replaced by a low pass filter.

The components of the RD are typical for many of the aforementioned architectures. We also note that other parallelized architectures to date require the use of multiple ADCs [2, 10].

3. CMUX: A NEW MULTI-CHANNEL ARCHITECTURE

3.1. System description

The CMUX acquires J independent signal channels, each of bandwidth $W/2$ Hz, into a single stream of samples running at the Nyquist rate (W Hz) of any one channel. As shown in Figure 1(a), each channel is first mixed down to baseband to obtain $x_j(t)$ and then modulated by a pseudo-random ± 1 chipping sequence $p_j(t)$ with chipping frequency W Hz. The spread channels are then summed and sampled once per chip by a single ADC. It is important to note that the summation occurs across the channels and not over time (in contrast to previous systems [1–3]).

Without loss of generality, the CMUX can be written as a $W \times JW$ matrix Φ, formed by concatenating diagonal $W \times W$ submatrices Φ_j, $j = 1, \cdots, J$. For the sake of analysis, we will consider the elements along the diagonals to be ± 1 Rademacher variables. As an example, let $J = 3$ and $W = 3$. Then the Φ matrix might look like

$$\Phi = \Big[\underbrace{\begin{matrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{matrix}}_{\Phi_1} \;\; \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{matrix}}_{\Phi_2} \;\; \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{matrix}}_{\Phi_3} \Big] \quad (4)$$

We consider signals that are *jointly sparse* over the combined bandwidth of the spectrum channels. The sparsity basis Ψ for this model is a $JW \times JW$ block diagonal matrix with $W \times W$ DFT bases along the diagonal. Thus, we aim to recover a K-sparse vector $\alpha \in \mathbb{R}^{JW}$ such that $y = A\alpha$, where A is the union of orthonormal bases

$$A = [\Phi_1\mathcal{F}, \Phi_2\mathcal{F}, \cdots, \Phi_J\mathcal{F}], \quad (5)$$

and where $\mathcal{F}$ is the $W \times W$ unitary DFT matrix. For the remainder of this paper, the subscript j denotes the submatrix (or subvector) corresponding to channel j and the subscript $\backslash j$ denotes the submatrix (or subvector) corresponding to all channels except for j.

It has recently been demonstrated by Romberg that A of this form satisfy the RIP [7]. We present a modified version of the statement of the theorem (as suggested in [7]) for completeness:

Theorem 1 (Theorem 3.1 in [7])**.** *Let A be defined as in (5), and fix $\delta \in (0, 1)$. Then there exists C_0 such that when*

$$W \geq C_1 K \log^4(JW) \quad (6)$$

A satisfies the RIP of order K as in (2) with probability $1 - C_0^2/\delta C_1^2$, where C_0 is constant.

Note that the constant C_0 is the same as that in [7], and improved bounds on the probability may be obtained [8]. It is clear from this statement that for the total bandwidth $N = JW$, the number of possible channels can be upper bounded as $J \leq \frac{N}{K}\frac{1}{C_1 \log^4 N}$.

3.2. The CMUX and bandpass sampling

Real-world RF tuners often mix signals to an "intermediate frequency" (IF) instead of directly to baseband. For example, candidate CMUX tuners for RF applications use IFs between 22 MHz and 70 MHz. Sampling systems with IF signals typically complete the downconversion using a *bandpass sampling* technique that intentionally undersamples the IF signal so that its non-aliased image falls near 0 Hz.

Most CS samplers ignore this issue, meaning that in practice they must consider a bandwidth that is $2f_{\text{IF}}$ Hz higher than necessary. The CMUX, however, can easily bandpass sample. Consider the alternative but equivalent CMUX system in Figure 1(b). This

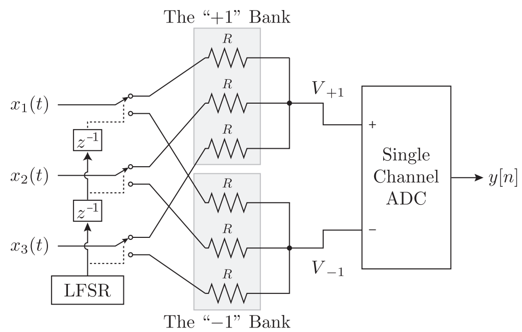

Fig. 2. Passive Averager CMUX (PA-CMUX) for $J = 3$.

figure shows that the sampling operation could arguably act directly on each channel input, which is exactly how bandpass sampling is performed.

3.3. The passive averager: A CMUX hardware concept

A major goal in implementing randomized CS hardware is to reduce the possible sources of hardware noise (i.e., achieve a simple design) so that it does not obscure the benefits achieved through sample rate reduction. To this end, we propose the Passive Averager CMUX (PA-CMUX).

As depicted in Figure 2, the PA-CMUX uses a single linear feedback shift register (LFSR) to generate the chipping sequence and J analog switches, two banks of resistors, and a single-channel ADC to perform the chipping sequence multiply and the instantaneous sum. The J uncorrelated chipping sequences are formed from delays of a single chipping sequence. Depending on the sign of the chipping sequence applied to it, each input signal $x_j(t)$ is routed by an analog switch to either a "$+1$" or "-1" bank of resistors. Each resistor bank consists of J resistors of nominally the same resistance; in practice, discrete resistors are unnecessary as each switch has a controlled output impedance.

Using Kirchoff's voltage and current laws, the voltage output from each bank, V_{+1}, V_{-1}, equals the average of the voltages that are fed into the bank, hence, inducing passive averaging. These voltages can be written as

$$V_{+1}[n] = \frac{\sum_{j \in P_{n,+1}} x_j[n]}{|P_{n,+1}|}, \quad V_{-1}[n] = \frac{\sum_{j \in P_{n,-1}} x_j[n]}{|P_{n,-1}|},$$

where $P_{n,+1}$ and $P_{n,-1}$ are the sets of channel indices being routed to the $+1$ bank or -1 bank at sample index n, respectively. Thus, we must rescale these voltages to obtain equivalent CMUX samples:

$$y[n] = |P_{n,+1}|\, V_{+1}[n] - |P_{n,-1}|\, V_{-1}[n].$$

While this could be achieved with two ADCs and gain components, we can also invert the averaging scale factors not during measurement, but during reconstruction. We simply compute the difference of the two averages, $y[n] = V_{+1}[n] - V_{-1}[n]$, and apply the averaging weights in Φ during reconstruction. The system designer can also calibrate the system by measuring the actual resistances along each signal pathway; these non-ideal values turn our averages into weighted averages.

The PA-CMUX relies on the fact that a single-channel ADC natively computes the difference between two voltages. In a typical setup, one of these voltages would be ground. However, in the PA-CMUX we can use it to compute the difference between the two averages.

3.4. Comparison with random demodulator

The CMUX compares favorably against the RD [1] on a number of fronts. CMUX computational models are more accurate and easier to calibrate due to absence of the analog filter. Additionally, since summation is performed over the channels and does not take place over time, the summation hardware is simpler. This sampling characteristic translates to a relaxed requirement on individual CMUX hardware components between samples (e.g. switching times); the RD's integrator enforces strict time-oriented performance requirements. Furthermore, the ability to bandpass sample significantly reduces the chipping and sampling frequencies for the CMUX as opposed to the RD. Even when ignoring the bandpass sampling issue, the CMUX's chipping sequences operate at a lower rate than the RD for the same total bandwidth. As power requirements for these components typically rise with the square of the frequency, a meaningful savings is achieved.

The multi-channel nature of the CMUX also brings benefits. The CMUX can grow its total bandwidth by adding channels without increasing the chipping and sampling rates. In RF scenarios, splitting the CMUX's target bandwidth across multiple RF tuners matches the fact that commercially-available tuners don't produce arbitrarily large bandwidths. And with access to multiple independent tuners, the CMUX can also allocate its bandwidth capacity where it is needed in the spectrum. The CMUX can also turn off unoccupied channels to improve performance; at an extreme, the CMUX reverts to a Nyquist sampler when all but one input channel is disabled.

There are of course some disadvantages. The CMUX undersampling factor is more restricted than in the RD. This factor is fixed at $J - J_{\text{off}}$, where J_{off} is the number of disabled input channels. Also, non-idealities inherent to the RF tuners (or equivalent) means that signals can fall out of coverage at channel edges.

4. ALGORITHMIC CHARCUTERIE

4.1. Trivial reconstruction

We can trivially produce an approximate recovery of any input channel i by multiplying that channel's chipping sequence against the output samples, i.e., $\widehat{x}_i = \Phi_i y$. It is clear from

$$\widehat{x}_i = \Phi_i y = \Phi_i \left(\sum_j \Phi_j x_j \right) = x_i + \sum_{i \neq j} \Phi_i \Phi_j x_j, \qquad (7)$$

that this approach yields the original channel x_j, plus a noise term that is the sum of the other channels spread by a new ± 1 sequence $\Phi_i \Phi_j$. Note that exact recovery is achieved when all non-zero coefficients are in a single channel. Furthermore, the trivial reconstruction can either be used by algorithms resilient to noise (correlation routines, PLLs, etc.).

4.2. Block coordinate relaxation (BCR)

The trivial reconstruction technique can be extended to perform joint reconstruction of all channels. One approach would be to approximate one channel as above, transform and threshold to keep the largest coefficients, subtract that channel's contribution from the measurements, and repeat this process with the other channels. Indeed, this is roughly the procedure of *block coordinate relaxation* (BCR) [11].

BCR provably solves the LASSO program

$$\alpha = \operatorname*{argmin}_{\alpha} \frac{1}{2} \|y - A\alpha\|_2^2 + \lambda \|\alpha\|_1 \qquad (8)$$

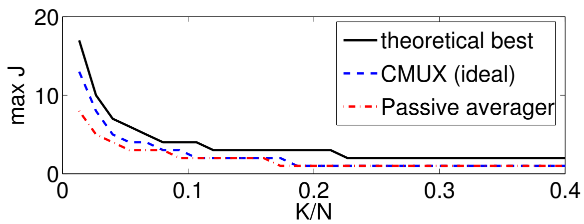

Fig. 3. Maximum number of channels J for fixed bandwidth $N = JW = 5000$.

when A is a union of orthonormal bases. For any ϵ in (3), there is an appropriate choice of λ such that the solutions to (3) and (8) are equivalent. We initialize by setting $r = y$ and the initial estimate $\alpha = 0$. The remaining steps are as follows: *i)* Choose a new block $j \in \{1, \cdots, J\}$. *ii)* Subtract the contribution of the current estimate (except from current block) from the measurements to update the residual, $r = y - A_{\setminus j}\alpha_{\setminus j}$. *iii)* Update the current block coefficients by soft-thresholding the DFT coefficients of the trivial reconstruction, $\alpha_j = \mathcal{S}(A_j^T r)$, where $\mathcal{S}(z) = z(|z| - \lambda)_+/|z|$, element-wise.

Note that for the CMUX, BCR uses exactly J FFTs and one soft-threshold operation of dimension W per iteration. Most other CS algorithms compute $A^T(y - Ax)$ in each iteration; thus these algorithms will require *at least* twice as many FFTs per iteration. Furthermore, the total number of iterations in BCR can be reduced by adaptively adjusting λ [9].

The soft-thresholding step of BCR projects the current channel estimate onto the ℓ_1-ball, thus "sparsely approximating" or "denoising" the channel estimate. We can extend this algorithm to recover non-integral frequencies, i.e., those oustide of the set of frequencies defined by the length-W DFT, by employing spectral CS [12] or BPDN-analysis [13] in place of soft-thresholding.

5. SIMULATIONS

5.1. Exactly sparse recovery

We wish to characterize the maximum number of channels required for exact recovery of sparse signals in simulation and compare this with the theoretical bound in Section 3.1. We fix $N = JW = 5000$ and vary K/N between 0 and 0.3. We perform 1000 reconstructions using SPGL1 [14] for each choice of J and record the maximum J such that 90% of the reconstructions yielded exact recovery.

Figure 3 demonstrates the results of this experiment. The dashed line depicts the experimental performance for an ideal CMUX given by (4); the dash-dotted line depicts the performance of the PA-CMUX given in Section 3.3 with resistors deviating randomly up to 20% from their intended values. The solid depicts the curve $J = N/(K \log(N/K))$ exactly, the best possible performance for a CS system, (note that this is better than the bound given in Section 3.1). This simulation demonstrates that in both cases, typical CMUX behavior is close to an ideal CS system thus appears to outperform the theoretical guarantees.

5.2. Practical RF example

In this example we simulate two FM modulated voice signals, each approximately 12 kHz wide. The voice signals live in two different 400 kHz wide channels. There are 5 total input channels, making the total observed bandwidth 2 MHz. All channels have noise such that the voice signals have an SNR of 30 dB.

PSDs of the signal in one channel are shown in Figure 4. The original signal is depicted by solid lines, the trivial recovery is depicted by a dash-dotted line, and the CS recovery is performed with

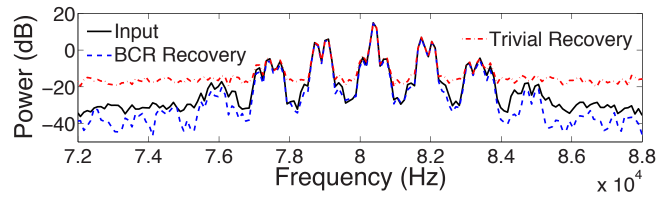

Fig. 4. Power spectral density of one 12 kHz-wide signal.

BCR as described earlier. The plot demonstrates that the largest energy portion of the spectrum can be recovered, even by the trivial method. However, the BCR recovery is significantly more accurate and has a lower noise floor than even the original signal (due to its more direct exploitation of sparsity).

6. DISCUSSION

In this paper we have introduced a compressive analog-to-digital converter that caters specifically to multi-channel RF signals. The key to our design is to code each channel and then combine them. We apply multi-channel separation for sparse signals for recovery. Our architecture permits simple calibration, requires only one ADC, and supports practical techniques such as bandpass sampling.

The CMUX is analogous to coded digital communications schemes such as code division multiple access (CDMA). Rather than coding the signal with orthogonal codes and transmitting into the same channel, we code the sensed channels and record them onto the same samples. Of course one difference is that the data we sense is analog, while in CDMA, the data is digital.

The CMUX also shares an interesting connection with compressive distilled sensing (CDS) [15]. Suppose that we detect that only a subset of channels have activity. We can then choose to disable some of the channels signals, thereby increasing the SNR of the channels of interest, a similar tradeoff to that found in CDS.

References

[1] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse, bandlimited signals," *IEEE Trans. Inform. Theory*, 2009.

[2] M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. of Selected Topics on Signal Proc.*, vol. 4, no. 2, pp. 375–391, 2010.

[3] J. Romberg, "Compressive sensing by random convolution," *SIAM J. on IMaging Science*, vol. 2, no. 4, pp. 1098–1128, Nov. 2009.

[4] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[5] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[6] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 6, no. 4, pp. 1289–1306, 2006.

[7] J. Romberg, "Multiple channel estimation using spectrally random probes," in *SPIE Wavelets XIII*, Aug. 2009.

[8] J. Romberg and R. Neelamani, "Sparse channel separation using random probes," *submitted to Inverse Problems*, Feb. 2010.

[9] J. Bobin, Y. Moudden, and J. Starck, "Morphological diversity and sparsity for multichannel data restoration," *J. Math Imaging Vis*, 2008.

[10] Z. Yu, S. Hoyos, and B. Sadler, "Mixed-signal parallel compressed sensing and reception for cognitive raio," in *International Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, 2008.

[11] S. Sardy, A. Bruce, and P. Tseng, "Block coordinate relaxation methods for nonparametric wavelet denoising," *J. of Comp. and Graph. Stat.*, vol. 9, no. 2, pp. 361–379, Jun. 2000.

[12] M. Duarte and R. Baraniuk, "Spectral compressive sensing," *submitted*, 2009.

[13] E. Candès, Y. Eldar, and D. Needell, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, 2010.

[14] E. van den Berg and M. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM J. on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.

[15] J. Haupt, R. Baraniuk, R. Castro, and R. Nowak, "Compressive distilled sensing: Sparse recovery using adaptivity in compressive measurements," in *Proc. 43 Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2009.