BEATING NYQUIST THROUGH CORRELATIONS: A CONSTRAINED RANDOM DEMODULATOR FOR SAMPLING OF SPARSE BANDLIMITED SIGNALS

Andrew Harms[†]*, Waheed U. Bajwa*[‡]*, and Robert Calderbank*[†‡]

[†]Princeton University, Princeton, NJ 08544
[‡]Duke University, Durham, NC 27708

ABSTRACT

Technological constraints severely limit the rate at which analog-to-digital converters can reliably sample signals. Recently, Tropp et al. proposed an architecture, termed the *random demodulator* (RD), that attempts to overcome this obstacle for sparse bandlimited signals. One integral component of the RD architecture is a white noise-like, bipolar modulating waveform that changes polarity at a rate equal to the signal bandwidth. Since there is a hardware limitation to how fast analog waveforms can change polarity without undergoing shape distortion, this leads to the RD also having a constraint on the maximum allowable bandwidth. In this paper, an extension of the RD, termed the *constrained* random demodulator (CRD), is proposed that bypasses this bottleneck by replacing the original modulating waveform with a *run-length limited* (RLL) modulating waveform that changes polarity at a slower rate than the signal bandwidth. One of the main contributions of the paper is establishing that the CRD, despite employing a modulating waveform with correlations, enjoys some theoretical guarantees for certain RLL waveforms. In addition, for a given sampling rate and rate of change in the modulating waveform polarity, numerical simulations confirm that the CRD, using an appropriate RLL waveform, can sample a signal with an even wider bandwidth without a significant loss in performance.

1. INTRODUCTION

One of the defining characteristics of analog-to-digital converters (ADCs) is the tradeoff between sampling rate and resolution. This tradeoff exists, in part, because the capacitors used to build ADC circuits take time to switch between charged and uncharged states, forcing designers to limit either the sampling rate or the resolution of an ADC [1, 2]. A rule of thumb for this rate–resolution tradeoff is that a doubling of the sampling rate causes a 1 bit reduction in the ADC resolution; in other words, $2^B \cdot f_s = P$, where B denotes the effective number of bits (ENOB)—a measure of ADC resolution, f_s denotes the sampling rate, and the constant P is determined by the state-of-the-art in ADC technology. Unfortunately, the constant P in ADC technology increases at a much slower pace than that followed by Moore's law for microprocessors [1, 2]. This forces many applications to push the current ADC technology to the limit. For example, software-defined radios require sampling rate on the order of 1 GHz and therefore can only manage resolution of 10 ENOB using today's ADC technology [1].

Fortunately, the rate–resolution tradeoff of the ADC technology can be circumvented by exploiting prior knowledge of additional structure in signals. One such additional structure is *signal sparsity*; it has been known for some time that bandlimited signals that are sparse in the frequency domain can be sampled at a rate that is much smaller than the Nyquist rate [3]. This old idea has been revisited in the past few years given the recent theoretical triumphs of compressed sensing [4]. In particular, while several techniques have been put forward for sampling sparse bandlimited signals at sub-Nyquist rates, three candidate architectures that rely primarily on recent developments in compressed sensing are *chirp sampling* [5], *Xampling* [6], and *Random Demodulator* [7]. Our focus in this paper is on the sampling of bandlimited signals that can be well-approximated through a small number of frequency tones and the Random Demodulator (RD) architecture seems particularly well-suited for this specific problem, including near-optimal guarantees for robustness against noise. Therefore we concentrate on the RD architecture in this exposition, although some of the ideas presented also appear to be of relevance to the Xampling architecture.

1.1. Our Contributions

One integral component of the RD architecture is a white noise-like, bipolar modulating waveform that changes polarity at a rate equal to the signal bandwidth. Since there is a hardware limitation to how fast analog waveforms can change polarity without undergoing shape distortion, the RD also has a constraint on the maximum allowable signal bandwidth. This bottleneck is reminiscent of the challenges faced by researchers in the early days of magnetic recording systems. In magnetic disks, 0's and 1's are stored by changing the polarization of the recording media and the read head reports transitions in recorded data as alternating positive and negative peaks in read-back voltage. Increasing the recording density on a magnetic disk by packing more bits in a region causes the read-back voltage to rapidly change polarity, leading to significant distortions in the peak amplitudes, among other things, and causes a large number of read errors.

In order to overcome this challenge in magnetic recording systems, Tang and Bahl [8] introduced the idea of *Run-Length Limited* (RLL) sequences in which run-length constraints describe the minimum separation, d, and maximum separation, k, between transitions from one state to another. The idea in the case of magnetic recording being that one can use (d, k) RLL binary sequences to increase the number of bits written on the disk by a factor of $(d + 1)$ without affecting the read-back fidelity. Note that there is a rate loss associated with converting arbitrary binary sequences to (d, k) RLL binary sequences and the major breakthrough in magnetic recording was that the rate loss associated with certain (d, k) sequences is smaller than $d + 1$, leading to a net increase in recording density; we refer the reader to [9] for further details on this topic.

In this paper, we make use of the lessons learned from the re-

[0]This work was supported by the Office of Naval Research under Grant N00014-08-1-1110, by the Air Force Office of Scientific Research under Grants FA9550-09-1-0551 and FA 9550-09-1-0643 and by NSF under Grant DMS-0914892.

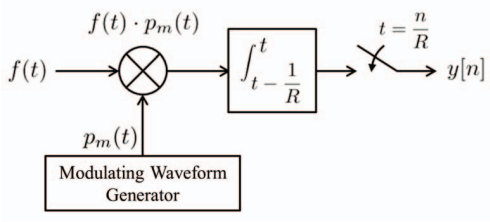


Fig. 1. Block diagram of the (constrained) random demodulator [7].

search on magnetic recording systems and propose an extension of the RD architecture, termed the *Constrained* Random Demodulator (CRD), that replaces the original RD modulating waveform with a (d, k) RLL modulating waveform. This is quite similar in spirit to the use of (d, k) sequences in magnetic recording systems and clearly leads to an increase in the operating bandwidth of the RD by a factor of $(d+1)$ without any hardware modifications. This increase in the operating bandwidth however comes at the cost of introducing statistical dependence across the modulating waveform. One of our main contributions is establishing that the CRD, despite employing a modulating waveform with correlations, enjoys some theoretical guarantees for certain RLL waveforms that are quite similar to the original RD architecture. In addition, one would expect an increase in the operating bandwidth to lead to an overall reduction in the allowable sparsity of the bandlimited signal. However, for a given sampling rate and rate of change in the modulating waveform polarity, we find through numerical simulations that the CRD using an appropriate RLL waveform can sample a signal with approximately $25\% \sim 35\%$ more bandwidth without a significant reduction in the signal sparsity.

2. BACKGROUND: THE RANDOM DEMODULATOR

In this section, we briefly review some of the key characteristics of the RD architecture as they pertain to the sampling of sparse bandlimited signals. We refer the reader to [7] for a comprehensive overview of this architecture.

The basic purpose of the RD is to take samples at a sub-Nyquist rate and still be able to reconstruct signals that are periodic, limited in bandwidth to W Hz, and are completely described by a total of $S \ll W$ tones. In other words, a signal $f(t)$ being fed as an input to the RD takes the parametric form

$$f(t) = \sum_{\omega \in \Omega} a_\omega e^{-2\pi i \omega t}, \; t \in [0, 1) \tag{1}$$

where $\Omega \subset \{0, \pm 1, ..., \pm(W/2 - 1), W/2\}$ is a set of S integer-valued frequencies and $\{a_\omega : \omega \in \Omega\}$ is a set of complex-valued amplitudes. In order to acquire this sparse bandlimited signal $f(t)$, the RD performs three basic actions as illustrated in Fig. 1. First, it multiplies $f(t)$ with a modulating waveform $p_m(t)$ that is given by

$$p_m(t) = \sum_{n=0}^{W-1} \varepsilon_n 1_{\left[\frac{n}{W}, \frac{n+1}{W}\right)}(t) \tag{2}$$

where the discrete-time *modulating sequence* (MS) $\{\varepsilon_n\}$ independently takes values $+1$ or -1 with probability $1/2$ each. Next, it low-pass filters the continuous-time product $f(t) \cdot p_m(t)$. Finally, it takes samples at the output of the low-pass filter at a rate of $R \ll W$.

One of the major contributions of [7] is that it expresses the actions of the RD on a continuous-time, sparse bandlimited signal $f(t)$ in terms of the actions of an $R \times W$ matrix Φ_{RD} on a vector $\alpha \in \mathbb{C}^W$ that has only S nonzero entries. Specifically, let $\mathrm{x} \in \mathbb{C}^W$ denote a Nyquist-sampled version of the continuous-time input signal $f(t)$. Then it is easy to conclude from (1) that x can be written as $\mathrm{x} = F\alpha$, where the matrix $F = \frac{1}{\sqrt{W}} \left[e^{-2\pi i n \omega / W}\right]_{(n,\omega)}$ denotes a (normalized) discrete Fourier transform matrix and $\alpha \in \mathbb{C}^W$ has only S nonzero entries corresponding to the amplitudes of the nonzero frequencies in $f(t)$. Now note that the effect of the modulating waveform on $f(t)$ in discrete-time is equivalent to multiplying a $W \times W$ diagonal matrix $D = \mathrm{diag}(\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_{W-1})$ with $\mathrm{x} = F\alpha$. Further, the effect of the low-pass filter on $f(t) \cdot p_m(t)$ in discrete-time is equivalent to multiplying an $R \times W$ matrix H, which has W/R consecutive ones starting at position $rW/R + 1$ in the r^{th} row of H, with $DF\alpha$.[1] Therefore, if one collects R samples at the output of the RD into a vector $\mathrm{y} \in \mathbb{C}^R$, then it follows from the preceding discussion that $\mathrm{y} = HDF\alpha = \Phi_{RD} \cdot \alpha$, where we have that the random demodulator matrix $\Phi_{RD} = HDF$.

Given the discrete-time representation $\mathrm{y} = \Phi_{RD} \cdot \alpha$, recovering the continuous-time signal $f(t)$ described in (1) is equivalent to recovering the S-sparse vector α from y. In this regard, the primary objective of the RD is to guarantee that α can be recovered from y even when the sampling rate R is far below the Nyquist rate W. Fortunately, recent theoretical developments in the area of compressed sensing have provided us with numerous greedy as well as convex optimization based methods that are guaranteed to recover α (or a good approximation of α) from y as long as the *sensing matrix* Φ_{RD} can be shown to satisfy certain geometrical properties [4]. The highlight of [7] in this regard is that the RD matrix is explicitly shown to satisfy the requisite geometrical properties as long as the sampling rate R scales linearly with the number of frequency tones S in the signal and (poly)logarithmically with the signal bandwidth W.

3. THE CONSTRAINED RANDOM DEMODULATOR

With the RD, it is possible to sample a sparse bandlimited signal at a significantly lower rate than the Nyquist rate. Still at issue though is the fact that the RD requires creation of a modulating waveform that changes polarity at the Nyquist rate. Given the nature of analog electronics, there is a hard bandwidth limit beyond which such waveforms cannot be generated without shape distortion. Stated differently, the RD makes use of a modulating waveform with an unconstrained $(d, k) = (0, \infty)$ RLL MS that in turn determines the maximum operating bandwidth of the RD architecture. The basic idea behind the CRD proposed here is to replace the unconstrained MS of the RD with an RLL MS with $d > 0$, which increases the operating bandwidth of the architecture by a factor of $(d + 1)$ without any changes to the hardware technology. In terms of the system equation, the CRD has the diagonal of the matrix D comprised of the RLL MS.

There is of course a price to be paid for using RLL sequences to increase the addressable bandwidth. Specifically, recall that RLL sequences place constraints on separations between different states (transitions), which are characterized by the parameters d and k—the minimum and maximum separation, respectively. Therefore the price we pay is that the entries of an RLL MS are not statistically independent. However, the key insight here is that the dependence for certain RLL MS's is local and decays geometrically to zero as separation within the MS increases.

In the following, we will make use of this insight to characterize

[1] Throughout this paper, we assume WLOG that R divides W.

the geometry of the CRD matrix Φ_{CRD} in terms of the Restricted Isometry Property (RIP). The RIP of a matrix Φ is important for the recovery of signals using the techniques of compressive sensing. The RIP of order S with restricted isometry constant δ_S is satisfied for Φ if

$$\left| \frac{\|\Phi \mathrm{x}\|_2^2 - \|\mathrm{x}\|_2^2}{\|\mathrm{x}\|_2^2} \right| \leq \delta_S \tag{3}$$

with $\delta_S \in (0,1)$ and $\|\mathrm{x}\|_0 \leq S$. Stated alternatively, RIP requires that the singular values of every $W \times S$ submatrix of Φ satisfy $\sqrt{1-\delta_S} < \sigma_i < \sqrt{1+\delta_S}$ for $1 \leq i \leq S$. In the following, we use the "triple-bar" norm of [7] to describe the RIP condition. Given a matrix A, this norm captures the largest magnitude eigenvalue of any $S \times S$ principal submatrix of A: $|||A||| = \sup_{|\Omega| \leq S} \|A|_{\Omega \times \Omega}\|$. Therefore, (3) is satisfied if and only if $|||\Phi^* \Phi - I||| \leq \delta_S$.

In [7], independence of the MS and independence of the rows of Φ_{RD} are used to show that Φ_{RD} satisfies the RIP. For the CRD, however, we have statistical dependencies both within the MS and across the rows of Φ_{CRD} that must first be dealt with. To that end, we apply an argument similar to that used in [10] establishing the RIP of random Toeplitz matrices and show that Φ_{CRD} can satisfy the RIP if entries of the MS become independent when separated by a distance greater than $\ell < \infty$. In the following section, we argue the validity of this assumption by discussing RLL MS's that (approximately) exhibit this behavior.

The forthcoming analysis relies on the matrix $\Delta = \{\Delta_{\alpha\omega}\}_{\alpha,\omega=0,\ldots,W-1}$ determined by the correlation characteristics of the MS. It has entries $\Delta_{\alpha\omega} = \sum_{j \neq k} \mathbb{E}[\varepsilon_j \varepsilon_k] \eta_{jk} f_{j\alpha}^* f_{k\omega}$ where $f_{j\alpha}$ is the (j,α)th entry of F, $\eta_{jk} = \langle h_j, h_k \rangle$, and h_j is the jth column of H. For the RD, $\Delta = 0$, whereas we require that $|||\Delta||| < 1$ for the CRD to satisfy RIP. To this end, we give the following theorem for which a proof will be provided in a future journal paper.[2]

Theorem 1 (RIP for the CRD)**.** *Let Φ_{CRD} be a $R \times W$ CRD matrix using a (d,k) RLL MS with maximum dependence distance ℓ and $|||\Delta||| < \delta$ for a fixed $\delta \in (0,1)$. Next, suppose that R satisfies*

$$R \geq \ell^3 (\delta - |||\Delta|||)^{-2} \cdot C \cdot S \log^6(W)$$

where C is a fixed constant. Then with probability $O(1 - W^{-1})$ the CRD matrix Φ_{CRD} satisfies the RIP of order S with constant $\delta_S \leq \delta$.

4. "GOOD" RUN-LENGTH LIMITED SEQUENCES

The statement of Theorem 1 suggests the CRD requires an RLL MS for which a) $|||\Delta||| < 1$ and b) entries separated by more than ℓ are statistically independent. We examine these conditions for two classes of RLL MS's.

First, consider a rate-$\frac{1}{d+1}$ repetition-coded (RC) MS that is generated from a Rademacher MS by repeating each element d times. Such a MS satisfies the (d,∞) RLL constraints. The repeated entries of the sequence are perfectly correlated with each other while remaining independent of all other entries in the sequence, resulting in $\ell = d+1$, which is optimal. However, the Gram matrix $\Delta^* \Delta$ for a rate-$\frac{1}{2}$ RC MS is the identity matrix resulting in $|||\Delta||| = 1$, and $|||\Delta||| \geq 1$ for any lower rate RC MS. We conclude that the CRD matrix using a RC MS will not satisfy the RIP.

Second, consider a General RLL (G-RLL) MS generated from a Markov chain. The correlation within such a sequence decays geometrically, i.e. the autocorrelation function $R_\varepsilon(m) \sim \lambda^{|m|}$ with $\lambda < 1$ [9, 11]. The dependence in such a sequence is not entirely local, so we cannot find a finite separation that guarantees complete independence. However, the correlation decays geometrically to zero so we can find a separation that gives us near independence. We will show that these sequences perform well as MS's in the CRD and approximately satisfy Theorem 1.

[2]For $d = 0$, $\ell = 1$ and $\Delta = 0$ and the result of [7] is obtained.

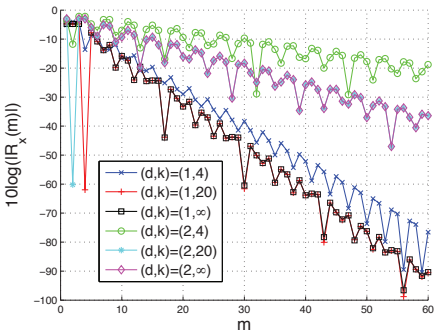


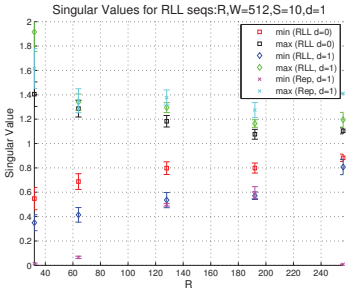
Fig. 2. Log-magnitude plot of the auto-correlation of an RLL code as a function of the time separation. Larger values of d and smaller values of k exhibit stronger correlation. The function for $k = 20$ is nearly identical to the function for $k = \infty$. For reference, an independent sequence is given by $d = 0$ and $k = \infty$.

To find the autocorrelation function for these G-RLL MS's, we need to understand the underlying Markov chain. The Markov chain is stationary and has $2k+2$ states described by a transition matrix $P = \{p_{ij}\}$ [9]. Calculation of the autocorrelation function $R_x(m)$ requires the vectors $\mathrm{a} : a_i = \sum_{u=1}^{2k+2} \pi_u p_{ui} y_{ui}$ and $\mathrm{b} : b_j = \sum_{v=1}^{2k+2} p_{jv} y_{jv}$ containing the weighted sum of symbols output on arriving at each state and on departing each state respectively. Here, π_u is the u^{th} entry of the stationary distribution for P, and y_{ij} is the symbol output on departing state i and arriving at state j. The autocorrelation is a function of only the separation between entries and is given by $R_x(m) = \mathrm{b}^T P^{(m-1)} \mathrm{a}$ [11].

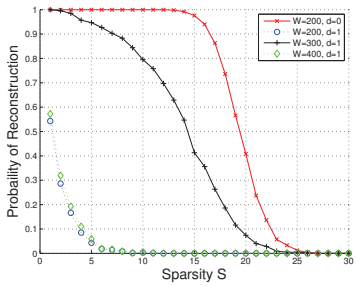
The autocorrelation is dependent on d, k, and the matrix P. Fig. 2 shows a plot of the auto-correlation function for several (d,k) sequences. We see the correlation indeed decays geometrically as m, the separation, increases. As an example, consider a G-RLL MS with $(d,k) = (1,20)$. The first few terms of the autocorrelation are given by $R_x(m) = [\frac{1}{3} \ \frac{-1}{3} \ \frac{-1}{3} \ 0 \ \frac{1}{6} \ \frac{1}{12} \ \ldots]$. For this autocorrelation function, we can carry out calculations to show that $|||\Delta||| < \sqrt{\frac{1}{3} + \frac{5}{6}\frac{R}{W}(S-1)}$. It is likely that optimizing over the choice of Markov chain used to generate the G-RLL MS will offer a smaller bound on $|||\Delta|||$, but that is beyond the scope of this paper.

5. NUMERICAL RESULTS AND DISCUSSION

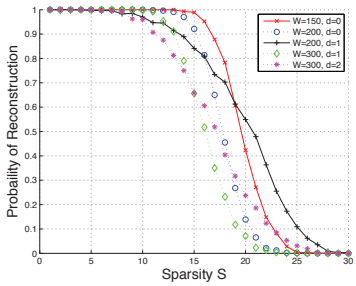
The RD is shown to satisfy RIP [7], and Theorem 1 tells us the CRD with certain RLL MS's also satisfies RIP. While these results do not hold for every RLL MS, numerical results suggest good performance if the correlation in the MS decays rapidly to zero and $|||\Delta||| < 1$. Fig. 3(a) shows the average minimum and maximum singular values, with error bars of two standard deviations, as a function of R obtained by evaluating sub-matrices of the system matrix using a G-RLL MS with $(d,k) = (0,\infty)$ and $(d,k) = (1,20)$ and a RC MS with $(d,k) = (1,\infty)$. The $d = 0$ points use an unconstrained MS and correspond to the RD. The $d = 1$ points are for the two CRD's. We can see a gap between the singular values for the CRD and the RD, and the CRD with the RC MS has some 0 singular



(a) Minimum and maximum singular values of size $W \times S$ sub-matrices of a CRD matrix as a function of R. The RLL d=0 points correspond to the RD of [7]. The RLL d=1 points correspond to a CRD with a $(1, 20)$ RLL MS and the Rep d=1 points correspond to a CRD with a $(1, \infty)$ RC MS. In each case, $W = 512$, $S = 10$, and 200 realization of the system are evaluated. We see a gap between the RD and CRD, but only the RC MS gives singular values of 0.



(b) RC MS: Probability of successful reconstruction vs. sparsity using the Basis Pursuit algorithm for 1000 instances of Φ for each set of parameters. The (x) marks are the RD with a 200Hz signal. The (o), (+), and ($\Diamond$) marks are the CRD with a rate-$\frac{1}{2}$ RC MS ((d, k)=$(1, \infty)$) and a 200Hz, 300Hz, and 400Hz signal. $R = 50$ and the transition width is fixed for each. The reconstruction performance is in all cases worse than with a G-RLL MS and is very poor for bandwidths of 200Hz and 400Hz.



(c) G-RLL MS: Probability of successful reconstruction vs. sparsity using the Basis Pursuit algorithm for 1000 instances of Φ for each set of parameters. The (x) and (o) marks are the RD with a 150Hz and 200Hz signal. The (+) and ($\Diamond$) marks are the CRD with (d, k)=$(1, 20)$ and a 200Hz and 300Hz signal. The (*) marks are the CRD with (d, k)=$(2, 20)$ and a 300Hz signal. $R = 50$ and the transition width is fixed for each. Reconstruction performance using the CRD is comparable to that of the RD, sacrificing only a small decrease in the allowed sparsity.

Fig. 3. Numerical Results

values. The singular values for the CRD with the $(1, 20)$ G-RLL MS are, however, still bounded close to 1. This leads us to believe that the RIP is satisfied, and we will still see good reconstruction performance in practice using a G-RLL MS with $d > 0$.

Fig. 3(b) and Fig. 3(c) plot the probability of successful reconstruction as a function of the input signal sparsity, S. The curves were obtained with different RLL sequences and signal bandwidths W. Fig. 3(b) uses RC MS's which become independent if entries have separation more than $d+1$ but which have $|||\Delta||| \geq 1$. Fig. 3(c) uses G-RLL MS's which do not become completely independent but have geometrically decaying correlation and for which $|||\Delta||| < 1$. In both plots, the d=0 curve is the RD and provides the baseline for our comparison, and all curves were created with a random waveform that switches at the same rate. Fig. 3(b) tells us that using a RC MS does not result in good reconstruction performance, especially for certain bandwidths. Fig. 3(c) tells a different story about using a G-RLL MS. The CRD with d=1 offers comparable performance to the RD but with the benefit of acquiring an input signal with more bandwidth. Depending on the tolerance in reconstruction probability, it can provide up to a 33% increase in the acquirable bandwidth. Even at 50% greater bandwidth, the CRD only reduces the sparsity by $\sim$4 (25%). This shows numerically that observable bandwidth can be increased with a slight to no drop in the sparsity.

To conclude, the RD shows that a sparse bandlimited signal can be sampled not only based on the bandwidth of the signal, but also the sparsity of the signal. The underlying hardware also gives a minimum transition width of the random waveform, so when at this limit we are limited to viewing a bandwidth W with the RD. However, by using a CRD with an appropriate RLL MS, we can increase the bandwidth up to $W' \leq (d + 1)W$ if we are willing to incur a penalty in the sparsity of the signal. We give a theoretical justification based on the RIP, and numerical simulations suggest that only a small penalty is incurred for bandwidth increases of even 50%. Despite the sparsity penalty, we still gain a great advantage; the RD is limited by the hardware to viewing signals of a particular bandwidth while the CRD can look beyond this bandwidth.

6. REFERENCES

[1] R. H. Walden, "Analog-to-digital converters and associated IC technologies," in *Proc. IEEE CSICS*, Oct. 2008, pp. 1–2.

[2] B. Le, T. W. Rondeau, J. H. Reed, and C. W. Bostian, "Analog-to-digital converters: A review of the past, present, and future," *IEEE Signal Proc. Mag.*, pp. 69–77, Nov. 2005.

[3] D. Rife and R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Th.*, pp. 591–598, Sept. 1974.

[4] ," *IEEE Signal Processing Mag., Special Issue on Compressive Sampling*, vol. 25, no. 2, Mar. 2008.

[5] L. Applebaum, S. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Appl. Comp. Harmonic Anal.*, pp. 283–290, Sept. 2008.

[6] M. Mishali, Y. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," submitted to *IET J. Circuits, Devices and Systems*, Dec. 2009.

[7] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inf. Th.*, pp. 520–544, Jan. 2010.

[8] D. Tang and L. Bahl, "Block codes for a class of constrained noiseless channels," *Inform. Cont.*, pp. 436–461, Dec 1970.

[9] K. Immink, P. Siegel, and J. Wolf, "Codes for digital recorders," *IEEE Trans. Inf. Th.*, pp. 2260–2299, Oct 1998.

[10] W. U. Bajwa, *New Information Processing Theory and Methods for Exploiting Sparsity in Wireless Systems*, Ph.D. thesis, University of Wisconsin - Madison, 2009.

[11] G. Bilardi, R. Padovani, and G. Pierbon, "Spectral analysis of functions of Markov chains with applications," *IEEE Trans. Commun.*, pp. 853–861, July 1983.