

A PROBABILISTIC ANALYSIS OF THE COMPRESSIVE MATCHED FILTER

Armin Eftekhari,^c Justin Romberg,^g and Michael B. Wakin^c

^cDivision of Engineering, Colorado School of Mines

^gSchool of Electrical and Computer Engineering, Georgia Tech

ABSTRACT

In this paper we study the “compressive matched filter,” a correlation-based technique for estimating the unknown delay and amplitude of a signal using only a small number of randomly chosen (and possibly noisy) frequency-domain samples of that signal. To study the performance of this estimator, we model its output as a random process and—borrowing from analytical techniques that have been used to derive state-of-the-art signal recovery bounds in the field of compressive sensing—we derive a lower bound on the number of samples needed to guarantee successful operation of the compressive matched filter. Our analysis allows the roles of time and frequency to be exchanged, and we study the particular problem of estimating the frequency of a pure sinusoidal tone from a small number of random samples in the time domain. Thus, for signals parameterized by an unknown translation in either the time or frequency domain, our theoretical bounds and experimental results confirm that random measurements provide an economical means for capturing and recovering such information.

Keywords— Compressive sensing, Matched filtering, Tone estimation, Random processes

1. INTRODUCTION

In the field of compressive sensing (CS) it is known that certain signals of high dimension but low complexity (namely, sparse signals) can be fully recovered from small numbers of random measurements. Several strong analytical bounds in CS have been derived using principles from the modern theory of empirical processes and the concentration of measure phenomenon. For example, the restricted isometry property, which is a sufficient condition for robust and stable signal recovery, can be recast as a supremum of a random process [10–14]. Such problems are the concern of a vast literature in statistics. One classic result, namely the Dudley inequality [4], provides a sharp upper bound for the expectation of the supremum of a random process. Such an expectation bound can then be extended to a probabilistic tail bound using arguments from the theory of concentration of measure [9].

The effectiveness of compressive measurements is not limited to problems involving the recovery of low-complexity sig-

nals. Compressive techniques can also be applied to variety of low-complexity inference problems (concerning possibly arbitrary high-dimensional signals) that do not require the full-scale signals to be reconstructed [1, 2, 7, 8]. In Section 2 of this paper, for example, we study the problem of estimating the carrier frequency of a pure sinusoidal tone from a small number of noisy, randomly chosen time-domain samples. We present strong bounds on the performance of a simple correlation-based technique for estimating the unknown frequency.

The roles of time and frequency can be easily exchanged in our analysis. This allows us to also consider the more general problem of matched filtering (i.e., estimating the unknown delay and amplitude of a known template signal) from a small number of noisy, randomly chosen frequency-domain samples. In Section 3, we provide the appropriate bounds for this problem. Our work complements past analysis of the so-called “smashed filter,” which viewed the estimation problem as a search along a parametric manifold embedded in the low-dimensional measurement space [1, 2]. In contrast, our methods provide a more direct characterization of the correlation statistics and allow the consideration of continuous-time signals.

Due to space limitations, we are unable to include proofs of our theorems in this paper. However, we note that our bounds have been derived by adopting an analytical framework very similar to those that have proved useful for studying CS signal recovery. The interested reader is referred to [5] for detailed proofs of our results based on the Dudley inequality and supporting concentration of measure arguments.

2. TONE ESTIMATION

2.1. Problem Setup and Noise-free Analysis

To set the stage for a general characterization of the compressive matched filter, we first consider the slightly narrower problem of tone estimation from a small number of randomly chosen samples in time. More precisely, suppose we observe samples of a pure complex exponential $A \cdot e^{i\omega_0 t}$ with fixed but unknown frequency $\omega_0 \in \Omega$ and amplitude $A \in \mathbb{C}$. Let $t_1, t_2, \dots, t_m \in T$ denote sampling times drawn randomly from a uniform distribution on an interval $T = [-t_{\max}, t_{\max}]$, and form the vector of observations $y \in \mathbb{C}^m$ as

$$y[k] = A \cdot e^{i\omega_0 t_k}, \quad k = 1, 2, \dots, m.$$

Given y , we are interested in estimating $\omega_0 \in \Omega$.

Email: aeftekha@mines.edu, jrom@ece.gatech.edu, mwakin@mines.edu. This work was partially supported by NSF Grants CCF-0830320 and CNS-0910592 and by DARPA grant FA8650-08-C-7853.

One natural strategy is to use the least-squares estimate for ω_0 , which is given by

$$\hat{\omega}_0 = \arg \max_{\omega \in \Omega} |\langle y, \psi_\omega \rangle|, \quad (1)$$

where $\psi_\omega \in \mathbb{C}^m$ is the test vector given by

$$\psi_\omega[k] = e^{i\omega t_k}, \quad k = 1, 2, \dots, m.$$

Having computed $\hat{\omega}_0$, it is also straightforward to compute a least-squares estimate for A .

Due to the randomness of the sample times, the output of the correlation-based estimator in (1) is a random variable. In order to study the performance of this estimator, we define the random process $X(\omega) := \langle y, \psi_\omega \rangle$ on Ω . To ensure the accuracy and robustness of our estimator, we would like guarantee that the empirical supremum of $|X(\omega)|$ happens close to ω_0 and that there is an ample margin between the peak of this process and the values of $|X(\omega)|$ away from ω_0 . Later in this section, we extend our analysis to account for the measurement noise.

To get a rough idea of how $X(\omega)$ behaves, we first compute the mean of this process:

$$\mathbb{E}X(\omega) = A \cdot m \cdot \text{sinc}(0.5|T|(\omega_0 - \omega)), \quad (2)$$

where $\text{sinc}(\alpha) := \sin(\alpha)/\alpha$. Informally, then, one expects that on average $|X(\omega)|$ will have a clear peak near ω_0 (where the sinc function has its main lobe) and decay away from this point (following the sidelobes of the sinc). This is indeed likely to be the case even for a small number of samples, as evidenced by the following theorem.

Theorem 1 [5, Theorem 1] *Suppose that $|\Omega||T| \geq 3$. Then*

$$\mathbb{E} \sup_{\omega \in \Omega} |X(\omega) - \mathbb{E}X(\omega)| \leq C_1 \cdot |A| \cdot \sqrt{m \log(|\Omega||T|)}, \quad (3)$$

where C_1 is a known universal constant.

Comparing (2) with (3), we see that while the peak of $|\mathbb{E}X(\omega)|$ scales with $|A|m$, the maximum expected deviation of $X(\omega)$ from its mean function scales only with $|A|\sqrt{m \log(|\Omega||T|)}$, and so even for small values of m , the peak of $|X(\omega)|$ can be expected to occur in the main lobe of the sinc function. Theorem 3 below translates this argument into a formal lower bound on the requisite number of measurements.

2.2. Robustness to Measurement Noise

Our analysis can be extended to account for additive observation noise. Let $n = [n_1, n_2, \dots, n_m]^T$ denote a vector of independent zero-mean complex-valued Gaussian random variables with variance σ_n^2 , and define the vector of noisy observations $y_n \in \mathbb{C}^m$ as

$$y_n[k] = A \cdot e^{i\omega_0 t_k} + n_k, \quad k = 1, 2, \dots, m.$$

By correlating y_n against the test vector ψ_ω for all $\omega \in \Omega$, we obtain the random process

$$X_n(\omega) := \langle y_n, \psi_\omega \rangle = X(\omega) + N(\omega),$$

where $X(\omega)$ is as defined above, and the zero-mean noise process is given by

$$N(\omega) := \langle n, \psi_\omega \rangle.$$

We can bound the peak of the noise process as follows.

Theorem 2 [5, Theorem 4] *Suppose that $|\Omega||T| \geq 3$. Then*

$$\mathbb{E} \sup_{\omega \in \Omega} |N(\omega)| \leq C_2 \cdot \sigma_n \cdot \sqrt{m \log(|\Omega||T|)},$$

where C_2 is a known universal constant.

Since the peak of $|\mathbb{E}X(\omega)|$ scales with $|A|m$, Theorem 2 suggests that the compressive estimator can withstand noise levels of $\sigma_n \sim |A|\sqrt{m \log^{-1}(|\Omega||T|)}$.

It can be shown that the arguments in both Theorems 1 and 2 hold not only on average, but also with a small failure probability. In other words, except with a small probability, both $|X(\omega) - \mathbb{E}X(\omega)|$ and $|N(\omega)|$ are guaranteed to be uniformly small over Ω . This enables us to quantify the number of samples needed to ensure, with high probability, a clear margin between the peak of $|X_n(\omega)|$ and the function away from the peak. From this, we can quantify the performance of the least-squares estimator as follows.

Theorem 3 [5, Corollary 8] *Let $\delta > 0$. Suppose that $|\Omega||T| \geq 3$ and that*

$$m \geq C_3 \cdot \log(|\Omega||T|/\delta) \cdot \max\left(1, \frac{\sigma_n^2}{|A|^2}\right), \quad (4)$$

where C_3 is a known universal constant. Then with probability at least $1 - \delta$, the maximum value of $|X_n(\omega)|$ must be attained for some $\hat{\omega}_0$ within the interval $[\omega_0 - 2\pi|T|^{-1}, \omega_0 + 2\pi|T|^{-1}]$.

2.3. Simulations

We illustrate the concentration behavior of the compressive least-squares estimator with a short collection of simulations. To begin, we set $\omega_0 = 100$ and consider the pure complex exponential signal $Ae^{i\omega_0 t}$ with $A = 1$. We collect m random samples of this signal over the interval $T = [-1/2, 1/2]$. For various values of m , Figure 1 plots the resulting correlation statistics $|X(\omega)|$ (solid blue line) as well as $|\mathbb{E}X(\omega)|$ (dashed red line—a shifted, scaled sinc function) over the domain $\Omega = [-300, 300]$. We observe the expected concentration of $|X(\omega)|$ around $|\mathbb{E}X(\omega)|$, with the relative maximum deviation decreasing as a function of m .

Next, we fix $m = 50$ and add to the measurements complex-valued Gaussian noise with $\sigma_n = c|A|\sqrt{m \log^{-1}(|\Omega||T|)}$. For various values of c , Figure 2 plots the resulting correlation statistics $|X_n(\omega)|$ (solid blue line) as well as $|\mathbb{E}X_n(\omega)|$ (dashed red line—a shifted, scaled sinc function exactly equal to $|\mathbb{E}X(\omega)|$) over the domain $\Omega = [-300, 300]$. We again observe the expected concentration of $|X_n(\omega)|$ around $|\mathbb{E}X_n(\omega)|$, with the relative deviation increasing as a function of c .

As discussed in the preceding sections, the correlation statistics $|X_n(\omega)|$ will be useful for estimating ω_0 when the peak of

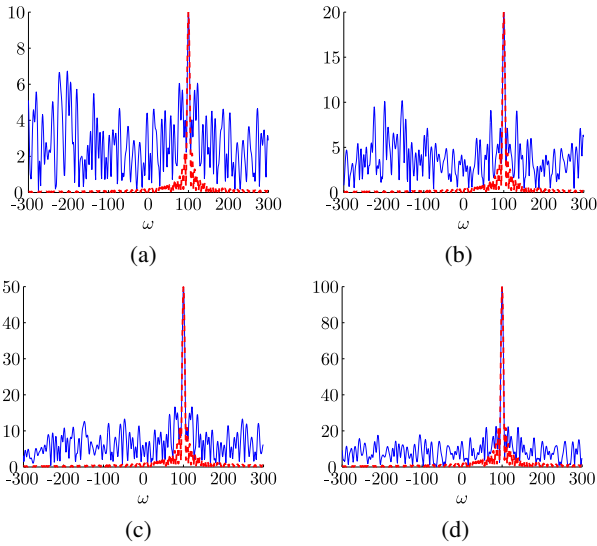


Fig. 1. The correlation statistics $|X(\omega)|$ (solid blue line) as well as $|\mathbb{E}X(\omega)|$ (dashed red line) for (a) $m = 10$, (b) $m = 20$, (c) $m = 50$, and (d) $m = 100$ noiseless random samples. The empirical correlation statistics become more tightly clustered around the mean function with increasing m . The coincidence of the peaks of the red and blue curves means that ω_0 can be accurately determined from the random samples.

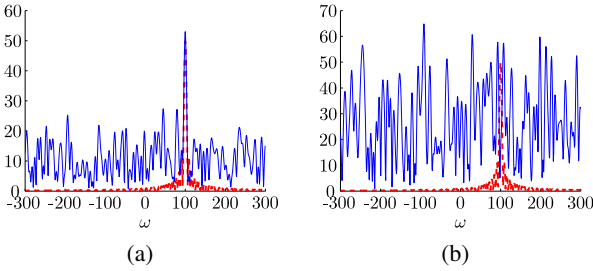


Fig. 2. The correlation statistics $|X_n(\omega)|$ (solid blue line) as well as $|\mathbb{E}X_n(\omega)|$ (dashed red line) for $m = 50$ samples corrupted by complex-valued Gaussian with $\sigma_n = c|A|\sqrt{m \log^{-1}(|\Omega||T|)}$, where (a) $c = 0.5$ and (b) $c = 1.5$. Increasing amounts of noise cause the amount of deviation between the two curves to increase; as a third point of comparison, Figure 1(c) corresponds to the noise level $\sigma_n = 0$.

this curve coincides with the main lobe of the mean function (a sinc centered at ω_0). However, as Figures 1 and 2 illustrate, the concentration of $X_n(\omega)$ about its mean depends both on the number of samples collected and on the level of noise. Thus, for various values of m and σ_n , we run 1000 trials described as follows. In each trial, we construct a random set of sample times and a random noise vector, compute the correlation statistics $|X_n(\omega)|$ over a grid of resolution $\frac{\pi}{4}$, and let $\hat{\omega}_0$ denote the frequency at which this quantity is maximized. Figure 3(a) plots, as a function of m and σ_n , the percentage of trials in which $|\hat{\omega}_0 - \omega_0| \leq 2\pi|T|^{-1}$, i.e., in which the empirical peak of $|X_n(\omega)|$ falls within the main lobe of the sinc function. In behavior that is consistent with (4), the compressive estimator withstands noise levels of $\sigma_n \approx 0.5|A|\sqrt{m \log^{-1}(|\Omega||T|)}$ across a range of m . (The vertical axis is logarithmic, so the

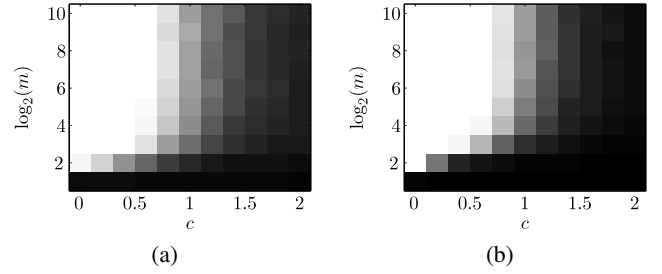


Fig. 3. The percentage of trials in which the frequency ω_0 is correctly estimated (to within $\pm 2\pi|T|^{-1}$) as a function of the number of samples m and the noise level $\sigma_n = c|A|\sqrt{m \log^{-1}(|\Omega||T|)}$. The experiments are run using the search bandwidth (a) $\Omega = [-300, 300]$ and (b) $\Omega = [-3000, 3000]$. White indicates 100% success, while black indicates 0% success.

tested number of samples m ranges from 2 to 1024.) Figure 3(b) reproduces the same experiment with a $10\times$ larger search bandwidth, i.e., with $\Omega = [-3000, 3000]$. Once again, the compressive estimator withstands noise levels of $\sigma_n \approx 0.5|A|\sqrt{m \log^{-1}(|\Omega||T|)}$ across a range of m .

3. TIME OF ARRIVAL ESTIMATION

3.1. Problem Setup and Noise-free Analysis

One can easily exchange the roles of time and frequency throughout our analysis. In doing so, the problem becomes that of estimating the unknown delay and amplitude of a known template signal using only a small number of randomly chosen (and possibly noisy) frequency-domain samples of that signal. For the special case of the tone estimation problem considered in Section 2, the so-called “template signal” corresponds to a sinc function. Our analysis, however, is not restricted to this choice of template function, and so in this section we study what we dub the *compressive matched filtering* problem at a more general level.

Let $s(t)$ denote a known template signal with Fourier transform $\hat{s}(\omega)$ that is bandlimited to some range of frequencies $\Omega = [-\omega_{\max}, \omega_{\max}]$. Let $\omega_1, \omega_2, \dots, \omega_m$ denote sampling times drawn randomly from a uniform distribution on Ω , and suppose we observe m samples of the Fourier transform of the signal $A \cdot s(t - \tau_0)$ at the frequencies $\omega_1, \omega_2, \dots, \omega_m$. The delay τ_0 is unknown (but is assumed to belong to some known interval T), and the amplitude $A \in \mathbb{C}$ is unknown.

Form the vector of observations $y \in \mathbb{C}^m$ as $y[k] = A \cdot e^{-i\omega_k \tau_0} \hat{s}(\omega_k)$, $k = 1, 2, \dots, m$. Given y , the least-squares estimate for τ_0 is equal to $\hat{\tau}_0 = \arg \max_{\tau \in T} |\langle y, \psi_\tau \rangle|$, where $\psi_\tau \in \mathbb{C}^m$ is the test vector given by $\psi_\tau[k] = e^{-i\omega_k \tau} \hat{s}(\omega_k)$, $k = 1, 2, \dots, m$. Having computed $\hat{\tau}_0$, it is also straightforward to compute a least-squares estimate for A . The random process $X(\tau) = \langle y, \psi_\tau \rangle$ on T has mean function

$$\mathbb{E}X(\tau) = 2\pi A m |\Omega|^{-1} R_{ss}(\tau - \tau_0),$$

where $R_{ss}(\cdot) = (s(t) \star s^*(-t))(\cdot)$ denotes the autocorrelation

function of $s(t)$. (When s is a sinc function as arises in the tone estimation problem, R_{ss} will also be a sinc function.) This mean function is guaranteed to peak at $\tau = \tau_0$ with maximum amplitude $|\mathbb{E}X(\tau_0)| = |A|m|\Omega|^{-1}\|\hat{s}\|_2^2$; away from this value, the mean function may or may not have significant sidelobes, depending on the nature of the template signal.

Using similar arguments to those in Section 2.1, one can guarantee that, with very high probability, the maximum deviation of $X(\tau)$ from its mean function will be small when m is sufficiently large. Such results, however, will also now depend on the degree to which the spectrum $\hat{s}(\omega)$ of the template signal is spread over the sampling domain Ω . Not surprisingly, signals with a relatively flat spectrum over Ω require the smallest number of uniformly distributed random samples. On the other extreme, if the support of \hat{s} is restricted to only a small set within Ω , one must take many more uniform samples over Ω to gather the same amount of information. To quantify this effect, let us introduce

$$\mu_1 = \sqrt{|\Omega|} \cdot \|\hat{s}\|_4^2 \cdot \|\hat{s}\|_2^{-2} \quad \text{and} \quad \mu_2 = |\Omega| \cdot \|\hat{s}\|_\infty^2 \cdot \|\hat{s}\|_2^{-2}.$$

It follows from the Hölder inequality that $\mu_1, \mu_2 \geq 1$ with equality if $|\hat{s}(\omega)|$ is constant on Ω . In general, smaller values of μ_1, μ_2 correspond to more uniformly spread spectra. We can now state a counterpart of Theorem 1, generalized to account for arbitrary template signals.

Theorem 4 [5, Theorem 1] *Suppose that $|\Omega||T| \geq 3$. Then*

$$\begin{aligned} & \mathbb{E} \sup_{\tau \in T} |X(\tau) - 2\pi Am|\Omega|^{-1}R_{ss}(\tau - \tau_0)| \\ & \leq C_1 \cdot \mu_1 |A||\Omega|^{-1} \|\hat{s}\|_2^2 \cdot \sqrt{m \log(|\Omega||T|)}. \end{aligned} \quad (5)$$

3.2. Robustness to Measurement Noise

Once again, our analysis can be extended to account for additive measurement noise through the appropriate definition of the noisy random process $X_n(\tau)$. It turns out that the compressive matched filter can withstand noise levels of $\sigma_n \sim |A|\|\hat{s}\|_2 \sqrt{m|\Omega|^{-1} \log^{-1}(|\Omega||T|)}$. Furthermore, expectation bounds such as (5) can be replaced with strong probabilistic statements. These arguments allow us to quantify the number of samples needed to control the maximum deviation of $X_n(\tau)$ from its mean function, which equals the scaled and shifted autocorrelation function of the template signal. This result is particularly interesting when the underlying autocorrelation function $R_{ss}(\tau)$ has one main peak (a “main lobe”) centered at $\tau = 0$ and is relatively small away from the origin (small “side lobes”). When this condition is met, it is possible to guarantee a clear separation between the peak of $|X_n(\tau)|$ near τ_0 and the values of $|X_n(\tau)|$ away from this peak. This notion is formalized in the next result.

Theorem 5 [5, Corollary 6] *Suppose there exist constants $\alpha_1 \in [0, 1)$ and $\alpha_2 > 0$ such that $|R_{ss}(\tau)| \leq \alpha_1 R_{ss}(0)$ for all $|\tau| > \alpha_2$. Suppose also that $|\Omega||T| \geq 3$ and that*

$$m \geq C_4 \cdot \frac{\log(|\Omega||T|/\delta)}{(1 - \alpha_1)^2} \cdot \max \left(\mu_1^2, \mu_2, \frac{\sigma_n^2 |\Omega|}{|A|^2 \|\hat{s}\|_2^2} \right),$$

where C_4 is a known universal constant. Then with probability at least $1 - \delta$, the maximum value of $|X_n(\tau)|$ must be attained for some $\hat{\tau}_0$ within the interval $[\tau_0 - \alpha_2, \tau_0 + \alpha_2]$.

4. CONCLUSIONS

In line with other results from CS (such as [3, 6, 14], which consider the recovery of frequency-sparse and other continuous-time signals), our work helps confirm that random measurements can provide an efficient, robust, and universal means for capturing signal information. In particular, for signals parameterized by a translation in either the time or frequency domain, our theoretical bounds and experimental results confirm that the unknown parameters can be recovered via a simple least-squares estimator from a number of random measurements that is only logarithmic in the Nyquist bandwidth. Because we do not require the signal to be recovered, our measurement bounds show no dependence on the sparsity level of the signal. All of our results in this paper have been proved using a probabilistic analytical framework very similar to those that have proven successful for studying the signal recovery problem in CS [10–14]. By extending this framework, one could consider incorporating additional measurement nonidealities such as clock jitter.

5. REFERENCES

- [1] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk. Signal Processing with Compressive Measurements. *IEEE J. Sel. Top. Signal Proc.*, 4:445–460, 2010.
- [2] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk. The smashed filter for compressive classification and target recognition. *Proc. Computational Imaging V at SPIE Electronic Imaging*, 2007.
- [3] M. F. Duarte and R. G. Baraniuk. Spectral compressive sensing. 2010. Preprint.
- [4] R. M. Dudley. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Functional Analysis*, 1(3):290–330, 1967.
- [5] A. Eftekhari, J. Romberg, and M.B. Wakin. Matched filtering from limited frequency samples. *Arxiv preprint arXiv:1101.2713*, 2011.
- [6] Y. C. Eldar. Compressed sensing of analog signals in shift-invariant spaces. *IEEE Trans. Signal Process.*, 57(8):2986–2997, 2009.
- [7] J. Haupt, R. Castro, R. Nowak, G. Fudge, and A. Yeh. Compressive sampling for signal classification. In *Proc. Asilomar Conf. on Signals, Systems, Comp.*, pages 1430–1434, 2007.
- [8] J. Haupt and R. Nowak. Compressive sampling for signal detection. In *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Proc. (ICASSP)*, Apr. 2007.
- [9] M. Ledoux. *The Concentration of Measure Phenomenon*. American Mathematical Society, 2005.
- [10] H. Rauhut. Circulant and Toeplitz matrices in compressed sensing. In *Proc. SPARS’09*, Saint-Malo, France, 2009.
- [11] H. Rauhut, J. Romberg, and J. Tropp. Restricted isometries for partial random circulant matrices. *Submitted to Appl. and Comp. Harm. Analysis*, October 2010.
- [12] J. Romberg. A uniform uncertainty principle for Gaussian circulant matrices. In *Proc. Int. Conf. Digital Signal Process.*, 2009.
- [13] J. Romberg and R. Neelamani. Sparse channel separation using random probes. 2010. Preprint.
- [14] 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(1):520–544, 2010.