

Automatic Modulation Recognition for Spectrum Sensing using Nonuniform Compressive Samples

Chia Wei Lim and Michael B. Wakin

Department of Electrical Engineering and Computer Science
Colorado School of Mines

Abstract—The theory of Compressive Sensing (CS) has enabled the efficient acquisition of high-bandwidth (but sparse) signals via nonuniform low-rate sampling protocols. While most work in CS has focused on reconstructing the high-bandwidth signals from nonuniform low-rate samples, in this work, we consider the task of inferring the modulation of a communications signal directly in the compressed domain, without requiring signal reconstruction. We show that the N^{th} power nonlinear features used for Automatic Modulation Recognition (AMR) are compressible in the Fourier domain, and hence, that AMR of M -ary Phase-Shift-Keying (MPSK) modulated signals is possible by applying the same nonlinear transformation on nonuniform compressive samples. We provide analytical support for the accurate approximation of AMR features from nonuniform samples, present practical rules for classification of modulation type using these samples, and validate our proposed rules on simulated data.

I. INTRODUCTION

A. Overview

Automatic modulation recognition (AMR)—the process of inferring a received signal’s modulation type by exploiting certain signal features—has varied applications in spectral monitoring, surveillance, and spectrum sensing in cognitive radios. With a wide variety of approaches available for a wide variety of modulation types, AMR is a well-established research arena [1], [3]. To apply most of the classical AMR techniques, however, Nyquist-rate (or faster) samples of the received analog signal must be available. When the signal has high bandwidth, this can place a significant burden on the front-end sampler, even though in AMR one wishes only to extract a very simple piece of information from the data.

Meanwhile, in a separate research field known as Compressive Sensing (CS) [5], [6], it has been shown that certain high-bandwidth signals can actually be recovered from small numbers of random, nonuniform samples collected at a rate far below Nyquist. In particular, such reconstruction is possible for signals that—within their high bandwidth—have very few large Fourier coefficients. (Such signals are said to be *sparse* in the frequency domain.) Reconstruction in CS from the low-rate samples is typically achieved by solving a regularized inverse problem with a sparsity-promoting signal penalty; the number of measurements required for reconstruction scales with the sparsity level of the signal.

In order to significantly reduce the measurement burden compared to classical Nyquist-rate AMR strategies, in this

paper we consider the task of solving the AMR problem from a small set of nonuniform compressive samples (NCS). While one natural approach to solving this problem would be to fully reconstruct a Nyquist-rate sample vector from the NCS and then apply a classical AMR algorithm to this vector, we ask the question of whether such full-scale reconstruction is truly necessary. Because we only wish to extract a very simple piece of information from the data, we propose to identify this information in a more direct fashion that requires only a crude estimate of the reconstructed signal. Our proposed technique is computationally very simple and requires far fewer NCS than would be needed to perfectly reconstruct the Nyquist-rate sample vector. This work could potentially enable AMR for signals with much higher bandwidth than is currently possible or lead to new radio system designs with smaller footprints and lower power requirements.

B. AMR for Spectrum Sensing

Cognitive radios change transmission or reception parameters to communicate efficiently and avoid interference with non-cooperative users present in the network. This capability is achieved via active monitoring of several factors, such as the radio frequency spectrum, user behavior, and the network state [2]. Hence, cognitive radios must be able to characterize the RF spectrum accurately despite having little or no a priori knowledge of the received signal characteristics. Among the important signal parameters that one might want to identify are symbol rate, carrier frequency, and modulation type. In this paper, we focus our discussion on recognizing the modulation type, although our proposed technique can also be used to estimate the carrier frequency.

For our NCS-based modulation classifier, we propose to use a decision theoretic technique relying on features obtained from the N^{th} power nonlinear transformation (NPT) [4]. In more classical settings, the NPT has been used for AMR on signals with Amplitude Shift Keying (ASK), Phase Shift Keying (PSK), and Minimum Shift Keying (MSK) modulation types. In this paper we focus on signals with M -ary Phase-Shift-Keying (MPSK) modulation. We detail our signal model in Section II and explain the NPT more fully in Section III.

C. Representative Experiment

The use of the NPT for AMR requires positive detection of features (peaks) in the spectrum of a signal which has undergone the NPT. Depending on the location of the peaks,

Email: clim@mines.edu, mwakin@mines.edu. This work was partially supported by DSO National Laboratories Singapore.

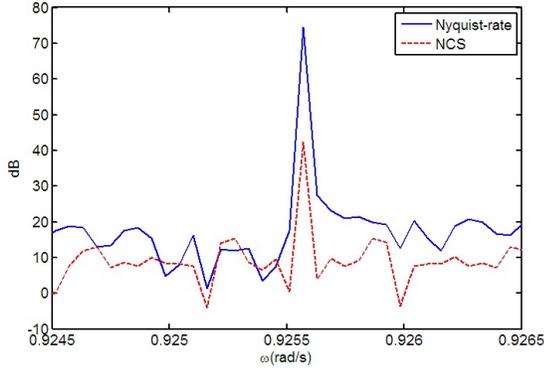


Fig. 1: Comparison of classification features generated from Nyquist-rate samples and from NCS. The presence of a peak in the dashed curve at the proper frequency confirms that AMR may indeed be possible from NCS without requiring full-scale, exact signal recovery.

different modulation types can be inferred. Here we show a simple experiment to demonstrate that such peaks can indeed be generated and detected from small numbers of NCS; the complete details of our technique are presented in Sections IV through VII, and we present several additional experiments in Section VIII.

Figure 1 shows two spectral estimates constructed for a 5000-symbol portion of a Binary Phase Shift Keying (2PSK) signal having carrier frequency $1.574 \times$ the symbol rate. The Nyquist-rate curve (top) is the result of applying the NPT on a collection of 100000 uniform time samples (this sampling rate is $\approx 9.64 \times$ the Nyquist rate). A discriminating feature—the prominent peak—is clearly visible in this spectrum. The NCS curve (bottom, dashed), on the other hand, was generated by applying our modified NPT technique to a collection of only 2700 NCS (this sampling rate is $\approx \frac{1}{2} \times$ the symbol rate and $\approx \frac{1}{4} \times$ the Nyquist rate). As we explain, the presence of a peak in the NCS curve at the proper frequency confirms that AMR may indeed be possible from NCS without requiring full-scale, exact signal recovery, although because the spectral estimate from NCS is imperfect, the strength of the peak is significantly lower than the peak in the Nyquist-rate curve.

In the typical CS setting, one would acquire NCS and subsequently reconstruct the underlying signal from its NCS. However, additional algorithmic complexity would be incurred due to the reconstruction process. Hence, in this work, we propose a low-complexity classifier that involves estimating the spectral peaks directly from the NCS. In the sections that follow, we provide analytical justification for the success of our technique using CS principles, and we empirically quantify the carrier-to-noise ratio margin loss of our estimator compared to one that uses full Nyquist-rate samples.

II. SIGNAL MODEL AND PROBLEM STATEMENT

We assume that a received signal has the general MPSK analytic form

$$r(t) = \sum_{n=-\infty}^{\infty} A g(t - nT - t_0) e^{j2\pi \frac{(m_n - 1)}{M}} e^{j(2\pi f_c t + \theta_c)} + v(t), \quad (1)$$

where A is the signal amplitude, t_0 is the timing offset, T is the symbol period, $M \in \{2, 4, 8\}$ is the number of unique

phases used, $m_n \in \{1, 2, \dots, M\}$ is the n^{th} transmitted symbol, f_c is the carrier frequency, θ_c is the carrier phase, $g(t) = \frac{\sin(\pi t/T) \cos(\pi \beta t/T)}{\pi t/T \sqrt{1 - 4\beta^2 t^2/T^2}}$ is the Root Raised Cosine Pulse Shape (RRC), β is the roll-off factor, and $v(t)$ is complex noise.

In the AMR problem, we are interested in determining M , the number of unique phases used in PSK modulation, from samples of $r(t)$. In Section III we discuss the NPT technique for AMR proposed in [4] that uses Nyquist-rate samples. In Sections IV through VII, we adapt this technique to operate using only NCS.

III. GENERATION OF FEATURES FROM NYQUIST SAMPLES

In [4], the use of the NPT was covered rigorously and here we focus on its application to MPSK signals. The idea behind the NPT is that by raising an MPSK signal to an appropriate power, one can convert the finite number of symbol phases to a constant phase and thus completely remove the effects of phase modulation due to data. Subsequent analysis in the Fourier domain gives a dominant spike at product of the carrier frequency and the raised power, and this spike acts as a useful feature for subsequent classification.

To be more precise, when a *noiseless* MPSK signal is raised to some integer power N , where $N = kM$ for some integer k , the resulting signal can be shown to have the following form:

$$(r(t))^N = \sum_{n=-\infty}^{\infty} A^N g^N(t - nT - t_0) e^{j(2\pi k M f_c t + k \theta_c)} + \sum_{n=-\infty}^{\infty} \text{Cross Terms}. \quad (2)$$

What is remarkable about (2) is that in this expression, the effect of carrier phase modulation has been removed. The result is a strong, constant-phase sinusoid at frequency $N f_c$.

Figures 2(a) and (b) show the spectrum of a 2PSK signal raised to the powers 2 and 4, respectively; a dominant peak is apparent in both cases because all carrier phase modulations have been converted to a constant phase. Figures 2(c) and (d) show the spectrum of a 4PSK signal raised to the powers 2 and 4, respectively; a dominant peak is visible only in the second plot, and it occurs at $4f_c$ where $f_c = 0.4628$ rad/s. Figures 2(e) and (f) show the spectrum of an 8PSK signal raised to the powers 2 and 4, respectively; no peaks are visible because the carrier phase modulation has decreased only from 8 unique phases to 4 and 2 phases, respectively (rather than 1). Thus, for an 8PSK signal, it is necessary to raise the signal to a power ≥ 8 to generate peaks in its spectrum.

The secondary peaks visible in the spectrum of $(r(t))^N$ such as those occurring to the left and right of the dominant (strongest) peak in Fig. 2(b) are due to symbol rate related periodicity after the NPT is applied on the received signal.

IV. GENERATION OF FEATURES FROM NCS

A. CS, NCS, and Reconstruction

CS has in recent years shown promise as a method for efficient signal acquisition. The idea behind CS is that certain

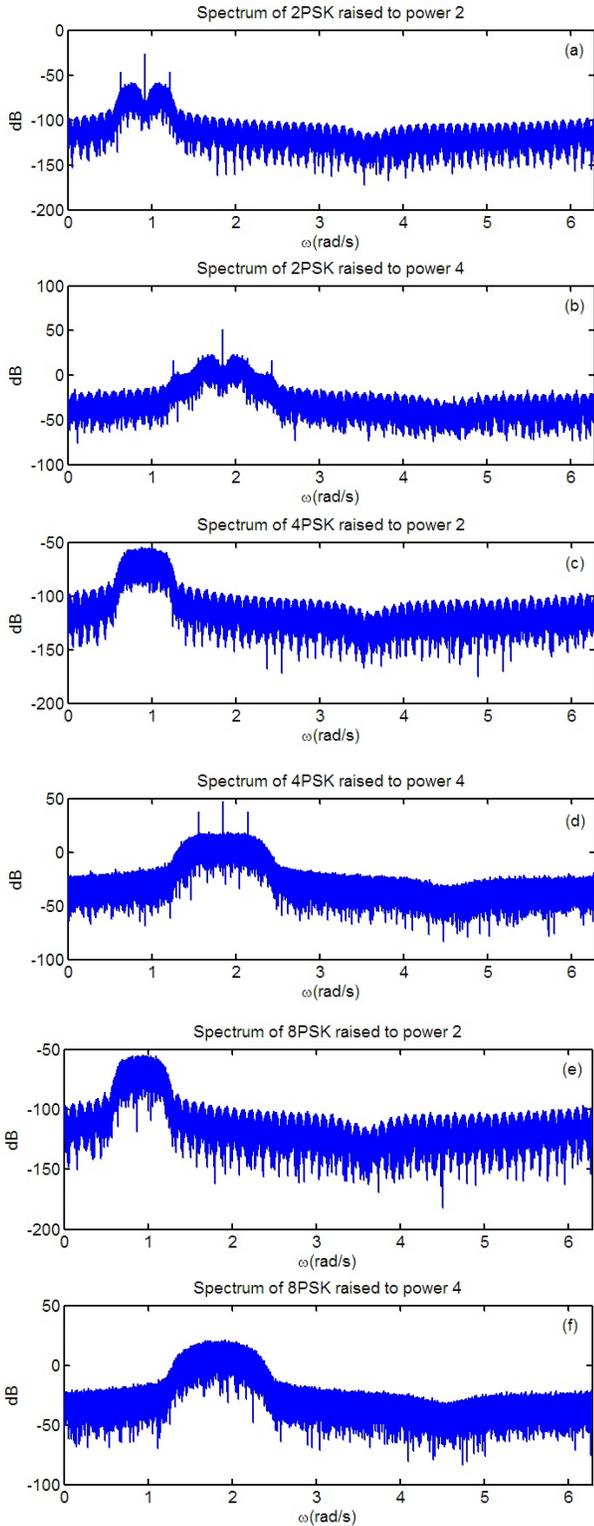


Fig. 2: (a) Spectrum of a 2PSK signal raised to the power 2; a dominant peak is apparent, suggesting that $N = 2$ is an integer multiple of the number M of phase modulation levels. (b) Spectrum of a 2PSK signal raised to the power 4; a dominant peak is apparent. (c) Spectrum of a 4PSK signal raised to the power 2; no peak is apparent, since $N = 2$ is not an integer multiple of $M = 4$. (d) Spectrum of a 4PSK signal raised to the power 4; a dominant peak is apparent. (e) Spectrum of an 8PSK signal raised to the power 2. (f) Spectrum of an 8PSK signal raised to the power 4; no peak is apparent for the final two plots since $N = 2$ and $N = 4$ are not integer multiples of $M = 8$.

high-bandwidth signals that obey *sparse* models can be recovered from a number of measurements that is far below what the Nyquist theorem would suggest. One way in which sparsity can be manifested is as follows: let z denote a hypothetical length- L vector of Nyquist-rate samples of some underlying analog signal. We say that z has an s -sparse representation in some $L \times L$ dictionary or basis Φ if we can write $z = \Phi\alpha$, where the length- L coefficient vector α has only $s \ll L$ nonzero entries.

The sparsity of α suggests that Nyquist-rate sampling of the analog signal may in fact be capturing redundant information. For example, it is possible to define a nonadaptive $P \times L$ compressive measurement matrix R with $P \ll L$ and actually recover the vector z from the length- P vector of measurements $y = Rz = R\Phi\alpha$. Defining the $P \times L$ matrix $A := R\Phi$, one can solve the undetermined system of equations $y = A\alpha$ if A satisfies a condition known as the *Restricted Isometry Property (RIP)* [5]. A matrix A is said to obey the RIP of order s with isometry constant $\delta_s \in (0, 1)$ if it satisfies

$$(1 - \delta_s)\|\alpha\|_2^2 \leq \|A\alpha\|_2^2 \leq (1 + \delta_s)\|\alpha\|_2^2 \quad (3)$$

for all s -sparse vectors α . Exact recovery of α from $y = A\alpha$ is possible via a sparsity-promoting ℓ_1 -minimization algorithm if A satisfies the RIP of order $2s$ with δ_{2s} sufficiently small. It has also been shown that the recovery is robust to noise, and approximate recovery is possible if the signal is nearly but not exactly sparse.

In the nonuniform compressive sampling scenario, an analog signal is sampled nonuniformly with random intervals between the sample times. Usually these random intervals are integer multiples of a fixed finer sampling interval (which is at least as small as the Nyquist-rate sampling interval). In our problem, let us define z to be a hypothetical length- L vector of uniform samples of the received signal $r(t)$:

$$z = [r(t_0) \ r(t_1) \ \cdots \ r(t_{L-1})]^T, \quad (4)$$

where $t_\ell = \ell T_s$ and T_s is a sampling interval that is equal to or smaller than the Nyquist-rate sampling interval. We assume that we do not measure z directly, but rather that we collect only a random subset of P samples from the L entries in z . We can represent the length- P vector of NCS as $y = Rz$, where R is a randomly generated binary selection matrix containing a single 1 on each row. For example, if $P = 3$, $L = 8$, and

$$R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

then we will collect the values $y[0] = z[1] = r(T_s)$, $y[1] = z[4] = r(4T_s)$, and $y[2] = z[6] = r(6T_s)$.

As a sparsifying basis, we choose Φ to be the $L \times L$ DFT synthesis matrix (i.e., one which implements the inverse DFT transform). Then we can write $z = \Phi\alpha$, where α is a length- L vector of DFT coefficients of z . Additionally, we can write

$$y = Rz = R\Phi\alpha.$$

It has previously been shown [10] that when Φ is an $L \times L$ DFT synthesis matrix and R is a $P \times L$ random selection

matrix, the composition $A = R\Phi$ of these two operators will satisfy the RIP of order $P/\log^4(L)$ with high probability. This means that from P random samples, one can recover signals that are nearly P -sparse in the DFT domain.

In our problem of interest, where z represents uniform samples of an MPSK signal $r(t)$, the corresponding DFT coefficient vector α will not necessarily be sparse, due to the PSK modulation. As we discuss in the next subsection, however, the application of the NPT to the samples can lead us to a sparse coefficient vector to which we may apply CS inference techniques.

B. Sparsity and the NPT

Define z_N to be the length- L vector that would result if we had access to z and could apply the NPT to this vector:

$$z_N[\ell] := (z[\ell])^N, \quad \ell = 0, 1, \dots, L-1.$$

Letting α_N denote the DFT coefficients of z_N , we can write $z_N = \Phi\alpha_N$, and as discussed in Section III, if $N = kM$ for some integer k , we can expect α_N to be exactly (or more likely, approximately) sparse.

As an example, in Fig. 2(d), one can see that the spectrum of a 4PSK signal when raised to the power of 4 contains three prominent peaks. In fact, in this example, a full 33% of the total signal energy is captured in just the three largest DFT coefficients. While such spectra do not satisfy the technical definition of sparsity, they are approximately sparse in some sense, and so this suggests that CS-based recovery or inference techniques could possibly be used to recognize the presence of the prominent peaks. We discuss the application of such techniques in Section IV-C below.

Lastly, we remark that as with any spectral analysis, windowing of the samples is helpful for concentrating energy in the main lobe and attenuating energy in the sidelobes, especially when the dominant NPT peaks do not occur at integer multiples of the frequency resolution of the DFT.

C. The NPT with NCS

We have argued that z_N , the result of applying the NPT to z , is approximately sparse. Unfortunately, in our measurement scenario, we do not have direct access to z and cannot apply the NPT directly. Rather, we observe only $y = Rz$, where R is a random selection matrix as discussed in Section IV-A. What is remarkable about our setup, however, is that *the selection operator R commutes with the NPT*:¹ by applying the NPT to the compressive samples, we obtain the same measurements that would be obtained by taking compressive samples of the NPT version of z . More formally, define y_N to be the length- P vector that results by applying the NPT to y :

$$y_N[p] := (y[p])^N, \quad p = 0, 1, \dots, P-1.$$

We see immediately that we can write this vector y_N —which we can compute—as

$$y_N = Rz_N = R\Phi\alpha_N = A\alpha_N, \quad (5)$$

¹Such commutativity would not hold with many other compressive matrices used in CS, such as those fully populated with independent Gaussian entries.

despite the fact that we have no direct access to z_N itself.

D. Estimation of the NPT Spectrum

From (5), we see that it is possible to compute what are equivalent to compressive samples of the Nyquist-rate NPT vector, and the matrix A that relates this vector to its compressive samples is very likely to satisfy the RIP. Assuming a sufficient number of measurements P are collected (relative to the sparsity level of α_N), one could consider applying various CS recovery algorithms to exactly or approximately recover α_N from y_N , and then solve the AMR problem in a classical setting.

However, accurately recovering the vector α_N may not be necessary if we consider ideas from the sub-field of CS known as *compressive signal processing* (CSP) [7], where one attempts to solve various detection, classification, or estimation problems directly from compressive measurements without attempting to reconstruct the full-length signal. CSP techniques, when successful, can be much more computationally efficient than CS recovery, and they can succeed using far fewer measurements.

In the spirit of CSP, let us define the very simple estimate of α_N :

$$\hat{\alpha}_N = A^H y_N = \Phi^H R^H y_N,$$

where the superscript H denotes the conjugate transpose. The estimate $\hat{\alpha}_N$ can be computed simply by constructing a length- L vector containing the entries of y_N at the appropriate sample times (and zeros elsewhere), and taking the DFT of this vector.

CSP theory allows us to place certain guarantees on the accuracy of this estimate. In particular, if α_N is exactly s -sparse and A satisfies the RIP of order $s+1$, then it follows from results in [8] that

$$|\hat{\alpha}_N[\ell] - \alpha_N[\ell]| \leq \delta_{s+1} \|\alpha_N\|_2 \quad (6)$$

for all $\ell = 0, 1, \dots, L-1$. One implication of (6) is that, if α_N were strongly dominated by one coefficient (say, at position ℓ^*) such that $|\alpha_N[\ell^*]| > |\alpha_N[\ell]| + 2\delta_{s+1} \|\alpha_N\|_2$ for all $\ell \neq \ell^*$, it would follow that the peak of $|\hat{\alpha}_N[\ell]|$ must occur at position ℓ^* .

In our settings we do not expect α_N to be exactly sparse or exactly dominated by just one coefficient. However, our experimental results confirm that $\hat{\alpha}_N$ provides a suitable, computationally efficient estimate for α_N even when the number of measurements P is very small. In Fig. 1, for example, the dashed red curve shows $|\hat{\alpha}_N|$ while the solid blue curve shows $|\alpha_N|$. We see that the estimate $|\hat{\alpha}_N|$ does indeed reveal a dominant peak and that the location of this peak coincides with the peak of $|\alpha_N|$.

V. DETECTION OF FEATURES

In order to achieve AMR, we must be able to detect the prominent features (peaks) previously identified in [4]. However, when computing $\hat{\alpha}_N$ from NCS, secondary peaks due to symbol rate periodicity may no longer be evident when $P \ll L$. Hence our focus is on implementing a robust detector with the goal of finding the dominant peak in $|\hat{\alpha}_N|$.

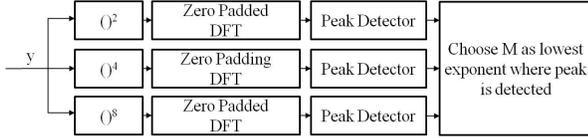


Fig. 3: Proposed classification scheme for determining M .

In particular, the Constant False Alarm Rate (CFAR) detector [9] is widely used for peak detection and here we propose to use the standard Cell Averaging CFAR (CA-CFAR) detector [9] for classification feature (peak) detection. For the sake of brevity, the interested reader is referred to [9] for exact algorithm specifications. In our simulations, we have used 1000 noise cells (NC) and 2 guard cells (GC).

VI. CLASSIFICATION RULE

Based on the discussions in Section III, when applying the NPT to an MPSK signal, a dominant spectral peak will be generated when the NPT power N is an integer multiple of M . In our simulations, we assume that we must choose between three possible modulation schemes: $M = 2$, $M = 4$, or $M = 8$. To solve this problem, we propose the classification procedure shown in Fig. 3.

VII. IMPACT OF NOISE

A. Generation of Features with Nyquist-Rate Samples

To illustrate the potential impact of measurement noise, observe that when a noisy MPSK signal of the form (1) is raised to the power $N = 2$, the resulting signal can be shown to have the following form:

$$(r(t))^2 = s_F(t) + s_{CT}(t) + 2s(t)v(t) + (v(t))^2 \quad (7)$$

where $v(t)$ is complex noise,

$$s(t) = \sum_{n=-\infty}^{\infty} Ag(t - nT - t_0)e^{j2\pi(m_n - 1)/M} e^{j(2\pi f_c t + \theta_c)},$$

$$s_F(t) = \sum_{n=-\infty}^{\infty} A^2 g^2(t - nT - t_0) e^{j[2\pi(2f_c t + 2\theta_c)]}, \text{ and}$$

$$s_{CT}(t) = \sum_{n=-\infty}^{\infty} \text{Cross Terms.}$$

When $M = 2$, $s_F(t)$ generates the required features (peaks) similar to the noiseless case described in (2). Compared to (2), however, (7) has 2 more terms: $2s(t)v(t)$ and $(v(t))^2$. These terms affect the features (peaks) by decreasing the overall power of the peaks relative to the noise floor. Similar conclusions hold for the cases $N = 4$ and $N = 8$.

B. Generation of Features with NCS

Applying the NPT to a noisy NCS sample vector y , we have

$$y_N = A(\alpha_N + \alpha_{N,\text{noise}}),$$

where A and α_N are as in (5), and $\alpha_{N,\text{noise}}$ represents the DFT coefficients of the 2 additional noise terms in (7).

Besides decreasing the overall power of the features (peaks) relative to the noise floor, $\alpha_{N,\text{noise}}$ diminishes the sparsity of

the desired spectrum. Because the number P of NCS required to accurately approximate the classification features depends on the sparsity level of $\alpha_N + \alpha_{N,\text{noise}}$ or the proximity of $\alpha_N + \alpha_{N,\text{noise}}$ to a sparse vector, it follows that we will require more measurements in order to ensure robust classification performance in the presence of noise. We quantify this tradeoff in our experiments in Section VIII, and we see that in a range of interesting problem scenarios the NCS-based classifier can succeed using sub-Nyquist sampling rates.

C. Detection and Classification

When applying the CA-CFAR detector to features generated from noisy data, we only require a change in the threshold ϵ . Through an iterative approach, we have found that $\epsilon = 5$ dB is an appropriate margin to use in our range of noise levels. No modification to the proposed classifier in Fig. 3 is required.

VIII. RESULTS

Using our proposed detector and classifier, Fig. 4(a) shows a plot of the number of correct classifications for a 2PSK signal with varying carrier-to-noise ratios (CNR); a total of 1000 Monte Carlo simulations were performed for each CNR. The 2PSK signal contains 5000 symbols, uses an RRC pulse shape with $\beta = 0.3$, and has $f_c = 1.574 \times$ the symbol rate. The ‘‘Nyquist-rate’’ curve was generated by applying the NPT on a collection of $L = 21396$ uniform time samples (this sampling rate is $\approx 2 \times$ the Nyquist rate). The ‘‘NCS’’ curve in Fig. 4(a) was generated by applying our modified NPT technique to a collection of only $P = 1070$ NCS (this sampling rate is $\approx \frac{1}{5} \times$ the symbol rate and $\approx \frac{1}{10} \times$ the Nyquist rate). We see that AMR using sub-Nyquist NCS is indeed possible, but compared to its Nyquist-rate counterpart, requires approximately 5dB higher CNR for a given probability of successful classification. Similarly, the ‘‘NCS’’ curve in Fig. 4(b) was generated using just $P = 214$ NCS (this sampling rate is $\approx \frac{1}{25} \times$ the symbol rate and $\approx \frac{1}{50} \times$ the Nyquist rate). Here, compared to Nyquist-rate AMR, AMR using NCS requires approximately 8dB higher CNR for a given probability of successful classification.

Fig. 4(c) repeats this experiment using a 4PSK signal (all other simulation parameters such as β , L , P , and the number of symbols are unchanged). Here, we see that compared to its Nyquist-rate counterpart, AMR using NCS (with $P \approx \frac{L}{20}$) requires at least 6dB higher CNR for a given probability of successful classification. The margin loss is slightly greater (approximately 10dB) when $P \approx \frac{L}{100}$, as shown in Fig. 4(d). As compared to the 2PSK signal, there is a greater margin loss for the 4PSK signal due to the diminished compressibility of the 4PSK signal.

Fig. 4(e) again repeats the same experiment using an 8PSK signal having the same β , L , and number of symbols, but with $P \approx \frac{L}{5}$. Compared to its Nyquist-rate counterpart, we see that AMR using NCS (for an 8PSK signal) requires at least an average of 6dB higher CNR for a given probability of successful classification, and the margin can be as high as 10dB for some CNR. The average margin loss is slightly greater (approximately 12dB) when $P \approx \frac{L}{20}$, as shown in

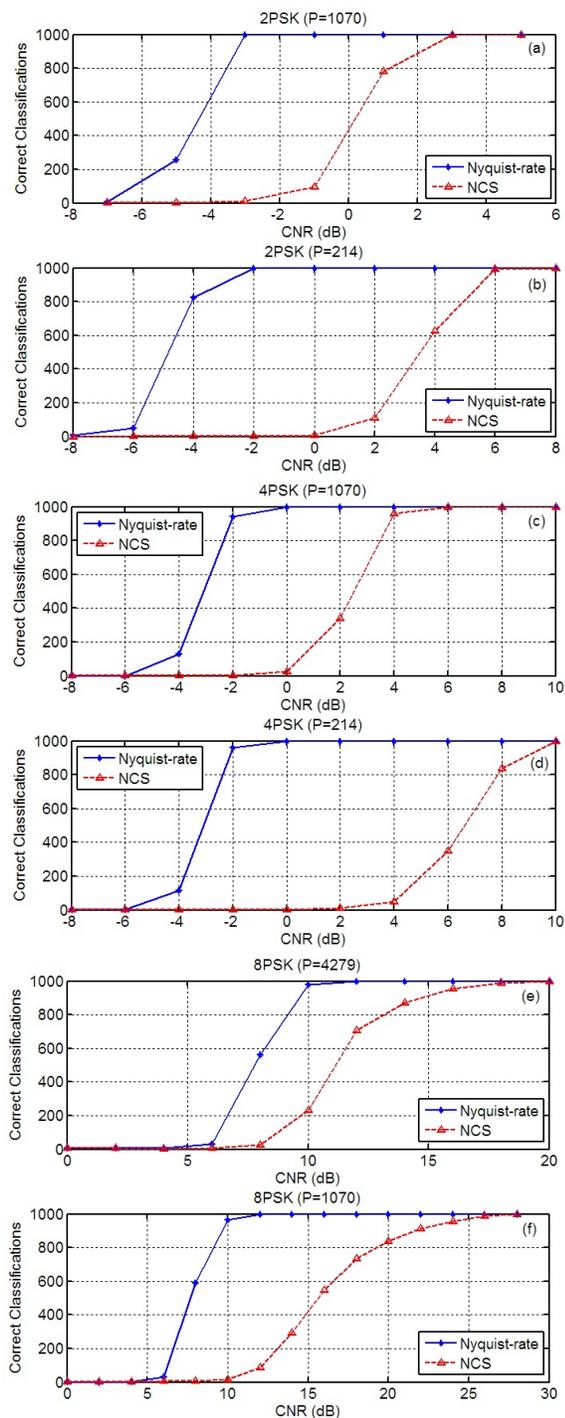


Fig. 4: Number of correct classifications (out of 1000 trials) versus carrier-to-noise ratio for (a) 2PSK signal with $P = 1070$ NCS, (b) 2PSK signal with $P = 214$ NCS, (c) 4PSK signal with $P = 1070$ NCS, (d) 4PSK signal with $P = 214$ NCS, (e) 8PSK signal with $P = 4279$ NCS, (f) 8PSK signal with $P = 1070$ NCS.

Fig. 4(d). As compared to the 2PSK and 4PSK signals, there is a greater margin loss for the 8PSK signal as it has the lowest compressibility when the NPT is applied. In general, we expect the CNR margins to increase with increasing values of M and do not expect the proposed classifier to be viable in practice for MPSK signals where $M \geq 16$. This does not

mean, however, that a more sophisticated (and algorithmically complex) classifier could not succeed for larger M using NCS.

IX. CONCLUSION

In order to significantly reduce the measurement burden compared to classical Nyquist-rate AMR strategies, and inspired by the CS research field, we have considered the task of solving the AMR problem from a small set of NCS. Specifically, we have proposed a simple estimation scheme for generating NPT features from NCS, a simple detection scheme for identifying peaks in the estimated NPT spectrum, and a simple classification rule for identifying the modulation parameter M . Our experimental results confirm that AMR using sub-Nyquist (even far sub-Nyquist) NCS is indeed possible and viable, but compared to its Nyquist-rate counterpart, AMR using NCS requires somewhat higher CNR for a given probability of successful classification. As a general rule, reduction in the number of NCS used for AMR will result in a greater margin loss; alternatively, increasing the number of NCS used for AMR improves the robustness to noise.

Our proposed technique is computationally very simple and requires far fewer NCS than would be needed to perfectly reconstruct the Nyquist-rate sample vector. This work could potentially enable AMR for signals with much higher bandwidth than is currently possible or lead to new radio system designs with smaller footprints and lower power requirements. In future work, we plan to extend our techniques to incorporate other types of classification features and to exploit cyclostationary properties of the signal.

REFERENCES

- [1] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Communications*, vol. 1, pp. 137–156, Apr. 2007.
- [2] S. Bilén, A. Price, O. Azarmanesh, and J. Urbina, "Modulation classification for radio interoperability via SDR," *SDR Forum Technical Conference*, Apr. 2007.
- [3] M. Wikström, "A survey of modulation classification methods for QAM signals," *Methodology Report*, Swedish Defence Research Agency, Mar. 2005.
- [4] J. Reichert, "Automatic classification of communication signals using higher order statistics," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp.221–224, Mar. 1992.
- [5] E. J. Candés, J. K. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete Fourier information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [6] E. J. Candés and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [7] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp.445–460, Apr. 2010.
- [8] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp.4395–4401, Sept. 2010.
- [9] P. P. Gandhi and S. A. Kassam, "Analysis of CFAR processors in homogeneous background," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 4, pp.427–445, July 1988.
- [10] M. Rudelson and R. Vershynin, "Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements," *IEEE Conference on Information Sciences and Systems*, 2006.