CHOCS: A Framework for Estimating Compressive Higher Order Cyclostationary Statistics

Chia Wei Lim and Michael B. Wakin

Department of Electrical Engineering and Computer Science, Colorado School of Mines
1500 Illinois Street, Golden, CO 80401, USA

ABSTRACT

The framework of computing Higher Order Cyclostationary Statistics (HOCS) from an incoming signal has proven useful in a variety of applications over the past half century, from Automatic Modulation Recognition (AMR) to Time Difference of Arrival (TDOA) estimation. Much more recently, a theory known as Compressive Sensing (CS) has emerged that enables 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 HOCS features used for AMR are compressible in the Fourier domain, and hence, that AMR of various linearly modulated signals is possible by estimating the same HOCS features from nonuniform compressive samples. We provide analytical support for the accurate approximation of HOCS features from nonuniform samples and derive practical rules for classification of modulation type using these samples based on simulated data.

Keywords: Higher order cyclostationary statistics, Compressive sensing, Compressive signal processing, Automatic modulation recognition, cyclic moments, cyclic cumulants

1. INTRODUCTION

By applying quadratic transformations to cyclostationary signals, one can compute what are known as second order cyclostationary statistics (SOCS). Typically, one uses SOCS to generate informative periodicities in the signal. By raising a cyclostationary 2Phase-Shift Keying (2PSK) signal to the power of two, for example, one can generate periodicities in the signal at twice its carrier frequency.[1] The pioneering work of Gardner et al.,[2,3] in the formulation of the theory of *higher order cyclostationary statistics* (HOCS) extends the theory of SOCS to higher orders (greater than 2). This has led to a wide variety of applications such as Automatic Modulation Recognition (AMR) and Time Difference of Arrival (TDOA) estimation of communication signals using HOCS. Previously,[3,4] HOCS were shown to have excellent signal selectivity in the sense of their ability to exhibit discriminating features (magnitudes and phases of the generated sine waves) when estimated from a linearly modulated cyclostationary signal.

The theory of Compressive Sensing (CS) or Compressive Sampling has in recent years shown great potential for enabling sub-Nyquist rate sampling of certain high-bandwidth signals. CS exploits the fact that many high-dimensional signals can be efficiently captured in a small number of parameters. More specifically, many signals are *sparse* or *compressible* when transformed into an appropriate domain. For example, a signal whose spectrum is not dense may be compressible when transformed into the frequency domain. CS allows such signals to be captured using a small number of non-adaptive (usually random) measurements, and the requisite number of measurements scales with the complexity of the signal (e.g., the total amount of occupied spectrum) rather than its gross bandwidth.

Early works in the theory of CS by Candès et al.[5] established the feasibility of the use of a random nonuniform sampler to acquire a signal that is compressible in the frequency domain. Subsequent reconstruction of the signal (from its compressed form) is possible by solving a convex optimization problem.

Further author information: (Send correspondence to C.W.L.)
C.W.L. e-mail: clim@mines.edu, M.B.W. e-mail: mwakin@mines.edu
This work was partially supported by DSO National Laboratories Singapore.

1.1 Motivation

While HOCS has proven to be useful, certain conditions must be met for successful HOCS estimation. In general, the higher the order of the HOCS features, the higher the oversampling of the signal is required (this is necessary to prevent aliasing). The estimation accuracy of HOCS features also improves with the duration of the signal since the HOCS theory relies on the usual statistical expectation operator, $\mathbb{E}[\cdot]$.

In short, these two restrictions translate to huge sampling burdens and data storage requirements on the front-end signal acquisition hardware if one wishes to subsequently estimate HOCS features from the acquired signal. Since HOCS features are compressible in the frequency domain, one could envision using CS to reduce the sampling rate. However, if we were to use a traditional CS formulation, reconstruction would still be required, and this would incur additional processing latency. In contrast, we propose to utilize techniques from a sub-field of CS known as *compressive signal processing* (CSP),[6] which will allow us to directly estimate the HOCS features without the need for full-scale signal reconstruction and potentially to do so with even fewer measurements.

Hence in this paper, extending our work in Reference 7, we consider the task of estimating HOCS features directly from a small set of nonuniform compressive samples (NCS). As a result, we propose a framework for estimating *compressive higher order cyclostationary statistics* (CHOCS).

1.2 Paper Organization

In the next section, we first introduce a model for the incoming signal, which is that of a linearly modulated baseband communication signal. Next we provide literature reviews on the theory of HOCS and on the NCS sampling protocol. Following that, we formally state the problem we intend to solve and define the environments when the problem arises.

In Section 3, we argue the feasibility of applying CS theory to estimate HOCS with a representative experiment and adapt conventional HOCS using NCS, culminating in CHOCS. In Section 4 we conclude with a discussion of a range of simulation results when using CHOCS to solve the AMR problem.

2. PROBLEM FORMULATION

2.1 Signal Model

Throughout this paper, we assume the incoming signal is a linearly modulated baseband communication signal of the form

$$s(t) = \sum_{k=-\infty}^{\infty} a s_k p(t - kT - t_0) e^{j(2\pi \Delta f_c t + \theta_c)}, \tag{1}$$

where a is the signal amplitude, s_k is the kth transmitted symbol, $p(t)$ is the signaling pulse shape, T is the symbol period, t_0 is the symbol timing offset, Δf_c is the residual carrier frequency offset, and θ_c is the carrier phase. We make the common assumption that $\Delta f_c \ll \frac{1}{T}$, and we also assume that the signal symbols, s_k, are independent and identically distributed (IID), uniformly chosen from some finite dictionary depending on the modulation type of the signal.

2.2 HOCS Preliminaries

2.2.1 Fraction-of-Time Probability (FOT)

In the theory of HOCS, the mathematical framework defined for analysis is known as the fraction-of-time (FOT) probability framework. In the FOT probability framework, statistical parameters are no longer defined in terms of the ensemble average of multiple realizations of the stochastic signal (as they would be under the stochastic processes framework), but rather in terms of infinite time averages, based on a single realization of the signal. The infinite time averaging operation defines probabilistic functions such as moments and cumulants of a random variable on the infinite duration signal.

2.2.2 Time-varying nth order Moments

The time-varying nth order moments of a signal $x(t)$ are defined[2,3] as

$$R_x(t,\boldsymbol{\tau})_{n,q} \triangleq E^{\{\alpha\}}\left[\prod_{i=1}^{n} x^{(*)_i}(t+\tau_i)\right], \tag{2}$$

where the ith factor $x(t+\tau_i)$ could be (optionally) conjugated and this conjugation is denoted by $(*)_i$. The variable q is the total number of conjugations and $\boldsymbol{\tau} \triangleq [\tau_1 \dots \tau_n]^T$ is a vector of lags. The sine wave extraction operator $E^{\{\alpha\}}[\cdot]$ is defined as

$$E^{\{\alpha\}}\left[g(t)\right] \triangleq \sum_{\forall \alpha : g_\alpha \neq 0} g_\alpha e^{j2\pi\alpha t}, \tag{3}$$

where

$$g_\alpha \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} g(t)e^{-j2\pi\alpha t}dt \equiv \langle g(t)e^{-j2\pi\alpha t}\rangle \tag{4}$$

extracts only the periodic component of the signal having frequency α. The argument to $E^{\{\alpha\}}[\cdot]$ in (2) is termed the *lag product* of the signal $x(t)$. In other words, given a signal $x(t)$, its time-varying nth order moments can be obtained by first analyzing its lag product for all periodic components and rebuilding its lag product only from these same periodic components.[‡]

The time-varying nth order moments of the signal $x(t)$ can be further expressed in terms of its lag-dependent Fourier coefficients (also known as its nth order cyclic moments) as

$$R_x(t,\boldsymbol{\tau})_{n,q} = \sum_{\alpha} R_x^{\alpha}(\boldsymbol{\tau})_{n,q} e^{j2\pi\alpha t}.$$

Therefore, the nth order cyclic moments satisfy

$$R_x^{\alpha}(\boldsymbol{\tau})_{n,q} = \langle R_x(t,\boldsymbol{\tau})_{n,q} e^{-j2\pi\alpha t}\rangle. \tag{5}$$

2.2.3 Time-varying nth order Cumulants

Before introducing the definition of time-varying nth order cumulants, for the sake of clarity, we state the definition of key parameters that the time-varying nth order cumulants are dependent on. Distinct partitions of the index set $\{1,2,\dots,n\}$ are referred to as D_n, d is the number of elements in a partition, and the set of indices belonging to a partition is denoted by $v_i,\ i \in \{1,\dots,d\}$. Table 1 shows the key parameters with their respective values when $n = 3$.

D_n	$\{1\}\{2\}\{3\}$	$\{1\ 2\}\{3\}$	$\{1\ 3\}\{2\}$	$\{2\ 3\}\{1\}$	$\{1\ 2\ 3\}$
d	3	2	2	2	1
v_i	$v_1 = \{1\}$, $v_2 = \{2\}$, $v_3 = \{3\}$	$v_1 = \{1,2\}$, $v_2 = \{3\}$	$v_1 = \{1,3\}$, $v_2 = \{3\}$	$v_1 = \{2,3\}$, $v_2 = \{1\}$	$v_1 = \{1,2,3\}$

Table 1: Key parameters and respective values for $n = 3$.

For a fixed value of n, the total number of distinct partitions is also known as the Bell number. All distinct partitions D_n can be generated using a modification of the recursive formula used to generate Bell numbers.[4]

Using the moment-to-cumulant conversion formula,[2] the time-varying nth order cumulants of a signal $x(t)$ are given by

$$C_x(t,\boldsymbol{\tau})_{n,q} = \sum_{D_n}\left[(-1)^{d-1}(d-1)!\prod_{i=1}^{d} R_x(t,\boldsymbol{\tau}_{v_i})_{n_i,q_i}\right], \tag{6}$$

[‡]The lag product may contain non periodic components as well.

where $n_i = |v_i|$ and

$$R_x(t, \boldsymbol{\tau}_{v_i})_{n_i,q_i} \triangleq E^{\{\alpha\}}\left[\prod_{k\in v_i} x^{(*)_k}(t+\tau_k)\right],$$

where $\boldsymbol{\tau}_{v_i}$ is the vector of lags in $\boldsymbol{\tau}$ with indices in v_i and $(*)_k$ again denotes optional conjugation of the kth factor $x(t+\tau_k)$. From (2), the time-varying nth order cumulants of the signal $x(t)$ can be shown[8] to have the following form:

$$C_x(t,\boldsymbol{\tau})_{n,q} = \left[a^n C_{s\,n,q} e^{j(2\pi\Delta f_c t+\theta_c)(n-2q)} e^{j2\pi\Delta f_c \sum_{m=1}^{n}(-)_m\tau_m} \prod_{\ell=1}^{n} p(t+\tau_\ell)\right] * \left[\sum_{k=-\infty}^{\infty}\delta(t-kT-t_0)\right], \quad (7)$$

where $C_{s\,n,q} = \sum_{D_n}(-1)^{d-1}(d-1)!\prod_{i=1}^{d} R_{s,v_i}$ is known as the symbol cumulants parameter, R_{s,v_i} are known as the symbol moments, $(-)_m$ denotes an optional minus sign depending on the optional conjugation and $*$ denotes the convolution operator. The symbol dependent parameters R_{s,v_i} can be computed from the symbols s_k in (1) using $\mathbb{E}\left[s_k^{|v_i|-q_i}(s_k^*)^{q_i}\right]$, where q_i is the number of indexes in the set v_i that are conjugated.

2.2.4 nth order Cyclic Cumulants

The nth order cyclic cumulants are the Fourier coefficients of the time-varying nth order cumulants,

$$C_x^{\beta}(\boldsymbol{\tau})_{n,q} \triangleq \langle C_x(t,\boldsymbol{\tau})_{n,q} e^{-j2\pi\beta t}\rangle.$$

It can be further shown[4] that

$$C_x^{\beta}(\boldsymbol{\tau})_{n,q} = \sum_{D_n}\left[(-1)^{d-1}(d-1)!\sum_{\boldsymbol{\alpha}^T\mathbf{1}=\beta}\prod_{i=1}^{d} R_x^{\alpha_i}(\boldsymbol{\tau}_{v_i})_{n_i,q_i}\right], \quad (8)$$

where $\boldsymbol{\alpha} = [\alpha_1 \ldots \alpha_C]^T$ with $\alpha_i : R_x^{\alpha_i}(\boldsymbol{\tau}_{v_i})_{n_i,q_i} \neq 0,\ \forall i \in \{1,\ldots,C\}$ referred to as the cycle frequencies and $\mathbf{1}$ is a column vector of length C having all entries equal to one. Thus, given the cyclic moments of a signal, one can use (8) to compute its corresponding cyclic cumulants. In general, C is not related to d and is dependent on the symbols s_k, whereas d is the number of factors formed from distinct partitions D_n.

From the definition of the time-varying nth order cyclic cumulants of $x(t)$, the nth order cyclic cumulants of the signal $x(t)$ can be shown[8] to have the following form:

$$C_x^{\beta}(\boldsymbol{\tau})_{n,q} = \frac{a^n}{T} C_{s\,n,q} e^{j\theta_c(n-2q)} e^{j2\pi\Delta f_c\sum_{m=1}^{n}(-)_m\tau_m} e^{-j2\pi\frac{kt_0}{T}}\left[\int_{-\infty}^{\infty}\left(\prod_{\ell=1}^{n} p(t+\tau_\ell)\right) e^{-j2\pi\frac{kt}{T}}dt\right], \quad (9)$$

where $\beta = \frac{k}{T} + (n-2q)\Delta f_c$, $k \in \mathbb{Z}$. Equation (9) gives an alternative form of the nth order cyclic cumulants as compared to (8) which relies on the moment-to-cumulant conversion formula. Hence, given the amplitude, symbol rate, symbol constellation of the modulation type, residual carrier frequency, timing offset and signaling pulse shape of a signal, one can also compute its cyclic cumulants using (9).

2.3 NCS Preliminaries

2.3.1 Sparsity

The key idea behind CS is the exploitation of the sparsity or compressibility of a signal. Specifically, the notion of sparsity can be explained as follows. Suppose we have a length L vector y consisting of entries which are samples of an analog signal acquired at its Nyquist rate. If we were to represent y in terms of an $L \times L$ dictionary (or basis) Φ, where $y = \Phi\eta$, y is said to have an s sparse representation in Φ if the length L coefficient vector η has $s \ll L$ nonzero entries. The vector y is said to be compressible in Φ if it can be well approximated by a sparse vector. Candès et al.[5] showed that if one were to make $P \ll L$ compressive measurements of sparse signal y resulting in a vector w via a $P \times L$ sensing matrix R such that $w = Ry = R\Phi\eta = A\eta$, exact recovery of

η (and therefore, y) from w may be possible. For example, it is possible to solve the under-determined system of equations (having P equations and L unknowns) if the $P \times L$ matrix satisfies the *Restricted Isometry Property* (*RIP*). A matrix $A = R\Phi$ satisfies the RIP of order s with isometry constant $\delta_s \in (0, 1)$ if the following holds for all s sparse η vectors

$$(1 - \delta_s)||\eta||_2 \leq ||A\eta||_2 \leq (1 + \delta_s)||\eta||_2.$$

Exact recovery of η from w can be achieved via an ℓ_1 minimization algorithm if A satisfies the RIP of order $2s$ and δ_{2s} is small. It was further shown that recovery is robust to noise and that recovery is approximate if η is not exactly sparse but compressible.

2.3.2 NCS

A commonly used[5] CS sampling protocol for frequency-sparse signals is NCS, and this is the sampling protocol that our CHOCS estimator will be built upon. In NCS, an analog signal is sampled with random time intervals between sampling instants. The random time intervals are usually integer multiples of a finer sampling interval (at least as fine as Nyquist rate of the signal). One can generate these random time intervals via a pseudo random number generator for simplicity. Based on our discussion in the previous section, we have the following relationship between our measurement (nonuniform compressive sampled) vector w and the matrix R

$$w = Ry,$$

where R is a binary $P \times L$ matrix containing 1's at the locations of the sample times. For the sake of clarity, if $P = 4$ and $L = 10$, one possible R matrix would be the following:

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

Hence in such an instance, we would have

$$w = \begin{bmatrix} w(1) & w(2) & w(3) & w(4) \end{bmatrix}^T = \begin{bmatrix} y(T_s) & y(4T_s) & y(5T_s) & y(10T_s) \end{bmatrix}^T,$$

where T_s is the finer sampling interval mentioned previously. One can easily see that entries $y(2T_s)$, $y(3T_s)$, $y(6T_s)$, $y(7T_s)$, $y(8T_s)$ and $y(9T_s)$ have been discarded by the NCS sampling protocol.

It was shown[9] that if Φ is the Discrete Fourier Transform (DFT) matrix and R is a random binary selection matrix as described above, then $A = R\Phi$ will satisfy the RIP of order $P/\log^4(L)$ with high probability. Hence, if η has a sparsity level up to approximately $P/\log^4(L)$, one can recover the vector η and consequently the vector y from only P random NCS measurements.

Having provided the necessary background on HOCS and NCS, we next formalize our problem statement. In this work, we aim to show that through the use of NCS (as a sampling protocol) and our proposed CHOCS estimator, it is possible to estimate HOCS of the incoming signal without the need for CS signal reconstruction.

3. CS AND HOCS

In order for CS to help in lifting the restrictions required for accurate estimation of HOCS features, these features must be concisely or mostly represented with a small number of non-zero coefficients belonging to some appropriate domain. Fig. 1 shows several magnitude plots of the Discrete Fourier Transform (DFT) of the lag product of a 2PSK signal with the corresponding values of n and q. The labeled peaks (also referred to as features) correspond to the various cyclic moments of the signal for each n and q. Evidently, the set of features (peaks) for each plot corresponds to just the 3 or 4 largest DFT coefficients of the signal's lag product and therefore, CS low rate sampling protocols (e.g., NCS) can be used to estimate HOCS in our scenario with the DFT basis as a sparsifying basis.

Fig. 2 shows (a zoomed portion of) the magnitude of the DFT of the lag product of a 2PSK signal with the top (solid blue) curve generated with signal samples acquired at the Nyquist rate. The bottom (red dashed) curve

gives the DFT of the lag product of the same 2PSK signal but generated from NCS, where only 20% of the total number of Nyquist rate samples are retained under NCS framework. A straightforward way of generating the bottom curve would be to use the same signal samples used to generate the top (solid blue) curve, set the sample values which would be discarded by the NCS sampling to zero, compute the lag product of this zero-padded signal, and then take the DFT of the resulting sequence (this is dealt with more rigorously in the next section). The distinct peak occurring in the same location as the peak for the Nyquist rate curve suggests the feasibility of using NCS to approximate the same peak, though the approximation is with another peak of significantly lower strength.

3.1 CHOCS

Returning to our previous discussion on NCS, suppose the vector y is the uniformly sampled Nyquist rate incoming signal of length L and we are computing its lag product with $\boldsymbol{\tau} = \mathbf{0}$ (the reason for this particular choice of $\boldsymbol{\tau}$ is further explained in the next section). We denote the lag product of the vector y by y_{lag},

$$y_{lag}[i] = (y[i])^{n-q}(y^*[i])^q, \tag{10}$$

for some n and q. Since we obtain the length P vector w when using NCS, we can define the following:

$$w_{lag}[i] = (w[i])^{n-q}(w^*[i])^q,$$

where w_{lag} denotes the lag product of the vector w. Due to the commutative relationship that the operator R (as previously defined) has with the lag product operation for $\boldsymbol{\tau} = \mathbf{0}$, we can also write the following

$$w_{lag} = Ry_{lag}. \tag{11}$$

Then, we can re-write (11) in the following form

$$w_{lag} = R\Phi\eta_{lag} = A\eta_{lag}, \tag{12}$$

where Φ is the $L \times L$ DFT matrix, the vector η_{lag} contains the DFT coefficients of the y_{lag} vector and $A = R\Phi$. Therefore, η_{lag} contains the cyclic moments of the incoming signal when referenced at the appropriate index. If, for a given sparsity level of η_{lag}, A satisfies the RIP, then exact reconstruction of y_{lag} would be possible given the observation vector w_{lag}. However, since we are considering the task of direct HOCS estimation without requiring prior signal reconstruction, we shall define the following estimate of η_{lag} from NCS as

$$\widehat{\eta}_{lag} = A^H w_{lag} = \Phi^H R^H w_{lag}, \tag{13}$$

where $(\cdot)^H$ denotes the conjugate transpose operator. The vector $\widehat{\eta}_{lag}$ may be computed by taking the DFT of the vector $R^H w_{lag}$. Previously,[10] it was shown that if the vector η_{lag} is s sparse and if A satisfies the RIP with order $s+1$, then the following guarantee holds for each $i \in \{1, 2, \ldots, L\}$:

$$|\widehat{\eta}_{lag}[i] - \eta_{lag}[i]| \leq \delta_{s+1}||\eta_{lag}||_2. \tag{14}$$

Therefore, it follows from (14) that the dominant coefficient of our estimate $\widehat{\eta}_{lag}$ (for example $\widehat{\eta}_{lag}[i^*]$) may also occur at the same location i^* in the vector η_{lag}. Though we do not expect η_{lag} to be only dominated by one coefficient as shown in Fig. 1, our simulation results show that $\widehat{\eta}_{lag}$ provides a low complexity estimate for η_{lag} since the NCS sampling protocol requires a lower sampling rate and fewer samples. In Fig. 2, the Nyquist-rate curve gives η_{lag} and the NCS curve gives $\widehat{\eta}_{lag}$ with the dominant peak ocurring at the same index location.

3.2 Features

We have argued the suitability of NCS as a sampling framework for generating HOCS features and now we discuss features suitable for solving the AMR problem. Based on (9), there are multiple dependencies relating the cyclic cumulants of a signal to its parameters. However, the only truly signal selective parameter is the symbol cumulants parameter $C_{s\,n,q}$. Hence the CHOCS features we shall use are:

$$F_x^0(\mathbf{0})_{n,\frac{n}{2}} := \frac{\left|C_x^0(\mathbf{0})_{n,\frac{n}{2}}\right|}{|C_x^0(\mathbf{0})_{2,1}|^{\frac{n}{2}}}, \quad \text{for } n \in \{4, 6, 8\}. \tag{15}$$

In practice, we denote an estimate of $F_x^0(\mathbf{0})_{n,\frac{n}{2}}$ from the incoming signal as $\widehat{F}_x^0(\mathbf{0})_{n,\frac{n}{2}}$ and compute $\widehat{F}_x^0(\mathbf{0})_{n,\frac{n}{2}}$ by first computing $\widehat{C}_x^0(\mathbf{0})_{n,\frac{n}{2}}$ (an estimate of $C_x^0(\mathbf{0})_{n,\frac{n}{2}}$) and $\widehat{C}_x^0(\mathbf{0})_{2,1}$ (an estimate of $C_x^0(\mathbf{0})_{2,1}$). More details about these computation are given in the next section. Therefore,

$$\widehat{F}_x^0(\mathbf{0})_{n,\frac{n}{2}} = \frac{\left|\widehat{C}_x^0(\mathbf{0})_{n,\frac{n}{2}}\right|}{\left|\widehat{C}_x^0(\mathbf{0})_{2,1}\right|^{\frac{n}{2}}}. \tag{16}$$

These set of normalized features (normalized with respect to $\left|C_x^0(\mathbf{0})_{2,1}\right|^{\frac{n}{2}}$) are chosen as they are independent of the underlying signal amplitude (a), residual carrier frequency offset (Δf_c), carrier phase offset (θ_c) and symbol timing offset (t_0). The lag vector ($\boldsymbol{\tau}$) was set to a vector of zeros ($\mathbf{0}$) since for the *Raised Cosine Pulse Shape* (RC), it has been mentioned[8] that the cyclic cumulant function achieves its maximum when $\boldsymbol{\tau} = \mathbf{0}$. We have also chosen cycle frequency $\beta = 0$ since this gives the largest value compared to other symbol rate harmonics. Using (9), it is straightforward to show that

$$F_x^0(\mathbf{0})_{n,\frac{n}{2}} = \frac{T^{\left(\frac{n}{2}-1\right)}\left|C_{s\,n,\frac{n}{2}}\right| \left|\int_{-\infty}^{\infty} p^n(t)dt\right|}{\left|C_{s\,2,1}\right|^{\frac{n}{2}} \left|\int_{-\infty}^{\infty} p^2(t)dt\right|^{\frac{n}{2}}}, \quad \text{for } n \in \{4,6,8\}, \tag{17}$$

which shows that the set of normalized CHOCS features selected are only dependent on the symbol period T, the symbol cumulants parameter $C_{s\,n,\frac{n}{2}}$, and the pulse shape $p(t)$. In general, for commonly used symbol dictionaries with unity variance, $C_{s\,2,1} = 1$.

4. RESULTS

Using our proposed CHOCS estimator, we performed Monte Carlo (MC) simulations for 2PSK, 4PSK, 8PSK and 16Quadrature-Amplitude-Modulation (16QAM) signal modulation types. The following signal parameters were kept constant for all simulations: signal amplitude ($a = 1$), residual carrier frequency offset ($\Delta f_c = 23.0625$Hz), symbol rate ($\frac{1}{T} = 12999.5625$ Hz), RC pulse shape with roll-off factor of 0.3, and signal duration of 1, and all of these parameters were assumed to be known. Since we intend to estimate CHOCS for up to order $n = 8$, the incoming signal needs to be at least 8×oversampled (i.e. the sampling frequency needs to be at least 8× its symbol rate for the signal model defined in Section 2.1) to prevent aliasing, hence a sampling frequency ($f_s = 131072$ Hz) was chosen for all simulations. For each signal type, 50 trials were performed.

The proposed CHOCS estimator was computed in the following manner.

1. Using the NCS vector w, compute its lag product using $w_{lag}[i] = (w[i])^{n-q}(w^*[i])^q, \forall i \in \{1,\ldots,P\}$ and for $n \in \{2,4,6,8\}$ and $q \in \{0,\ldots,\frac{n}{2}\}$.

2. For each n, q, compute the DFT ($\widehat{\eta}_{lag} = \Phi^H R^H w_{lag}$) of the lag product of the incoming signal.

3. Detect peaks (which are also estimates of the cyclic moments $R_x^\alpha(\mathbf{0})$) in $\widehat{\eta}_{lag}$. For each peak detected, there is a corresponding cycle frequency α. In our simulations, we have used the standard Cell Averaging Constant False Alarm Rate (CA-CFAR)[11] detector for detecting peaks with parameter $\epsilon = 6$dB and 40 noise cells.[†]

4. Compute estimates $\widehat{C}_x^0(\mathbf{0})_{n,\frac{n}{2}}$ and $\widehat{C}_x^0(\mathbf{0})_{2,1}$ using (8) based on all detected cyclic moments (the same peaks detected in the previous step), where for each peak detected, $\widehat{\eta}_{lag}[i_{detected}] = R_x^\alpha(\mathbf{0})_{n,q}$ and $i_{detected}$ corresponds to its index in $\widehat{\eta}_{lag}$.

5. Compute estimates of the desired CHOCS features $\widehat{F}_x^0(\mathbf{0})_{n,\frac{n}{2}}$ using (16).

[†]The interested reader is referred to Reference 11 for exact algorithm specifications.

Fig. 3 shows the results of our MC simulations in the noiseless case. In the plots, the parameter $\gamma = \frac{P}{L}$ represents the fraction of the full Nyquist rate samples retained when using the NCS sampling protocol where P and L are as defined in Section 2.3.1. For the sake of clarity, when $\gamma = 0.1$, the sampling rate is $\frac{1}{10} \times f_s$ and $\approx 1\times$ the symbol rate. The plots show that the only feature that can be distinguished between the different signal types is $\widehat{F}_x^0(\mathbf{0})_{8,4}$ since it gives distinct values for different signal modulation types having the same parameters when $\gamma \geq 0.2$ (this corresponds to an average sampling rate of $\frac{1}{5} \times f_s$ and $\approx 2\times$ the symbol rate). Hence the proposed CHOCS estimator can be used for AMR of the four signal modulation types when $\gamma \geq 0.2$.

Next, additive white Gaussian noise (AWGN) was added to the signal types such that the resulting signal-to-noise ratio ($\frac{E_b}{N_0}$) is 5dB; all other signal parameters had the same values as before in the noiseless case. The resulting plots in Fig. 4 show that the only feature that is able to be distinguished between the different signal types is $\widehat{F}_x^0(\mathbf{0})_{6,3}$ since it gives distinct values for different signal modulation types having the same parameters when $\gamma \geq 0.2$ (this corresponds to a sampling rate of $\frac{1}{5} \times f_s$ and $\approx 2\times$ the symbol rate). Again, CHOCS is able to perform AMR of the four signal modulation types when $\gamma \geq 0.2$.

We note that the differing distinguishable features in the noiseless case and mediocre signal-to-noise ratio ($\frac{E_b}{N_0} = 5$dB) case suggest a possible sensitivity of the CHOCS features to noise. Therefore, in practice, estimation of the level of noise present in the incoming signal would be necessary in a pre-processing step.

Table 2 summarizes the range of values (in dB) of CHOCS features for the four signal modulation types considered for $\gamma = 0.2$. Classification rules for the four signal modulation types may be derived from these values.

Finally, in order to provide concrete evidence for CHOCS as a viable alternative estimator of HOCS, we attempt to estimate HOCS from Nyquist-rate samples of an incoming signal of duration of γ seconds. This is to provide a fair comparison of CHOCS and HOCS since CHOCS succeeds for an incoming signal of duration 1 second with $\gamma \geq 0.2$. Fig. 5 shows that accurate estimation of HOCS is not possible at all for the proposed HOCS features. Hence, with such a short signal duration, HOCS is not able to solve the AMR problem for the four signal modulation types.

5. CONCLUSION

Inspired by the field of CS, utilizing ideas from CSP, and motivated by the need to significantly reduce the analog front end acquisition burden, we have considered the task of solving the AMR problem between four signal modulation types using CHOCS. Specifically, through MC simulations, we have shown that NCS can be used to provide a significantly more efficient sampling protocol requiring only $\frac{1}{5}$ of the original sampling rate. We have also shown that even in the case of a mediocre signal-to-noise ratio situation ($\frac{E_b}{N_0} = 5$dB), CHOCS can still succeed. Our work could potentially be extended to enable AMR for much higher bandwidth signals than is currently possible due to sampling hardware limitations. In future work, we plan to increase the number of signal modulation types, incorporate other types of transmission channels besides AWGN and develop a theoretical framework for the proposed CHOCS estimator.

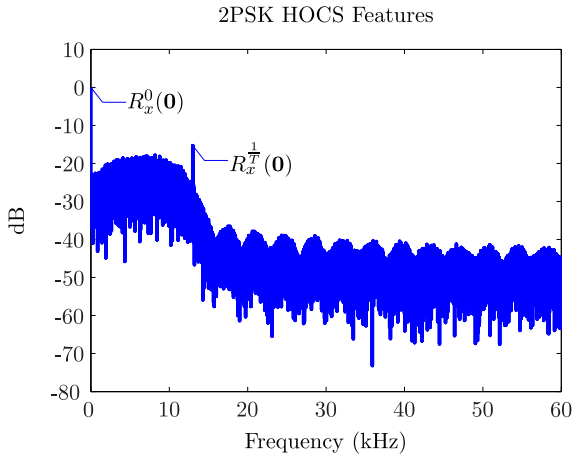
REFERENCES

1. Reichert, J., "Automatic classification of communication signals using higher order statistics," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 221–224, (1992).
2. Gardner, W. A. and Spooner, C. M., "The Cumulant Theory of Cyclostationary Time-Series, Part I: Foundation," *IEEE Transactions on Signal Processing* **42**, pp. 3387–3407, Dec. (1994).
3. Gardner, W. A. and Spooner, C. M., "The Cumulant Theory of Cyclostationary Time-Series, Part II: Development and Applications," *IEEE Transactions on Signal Processing* **42**, pp. 3409–3429, Dec. (1994).
4. Spooner, Chad M., *Theory and application of higher order cyclostationarity.* PhD thesis, University of California, (1992).
5. Candès, E. J., Romberg, J. and Tao, Terence, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Transactions on Information Theory* **52**, pp. 489–509, Feb. (2006).

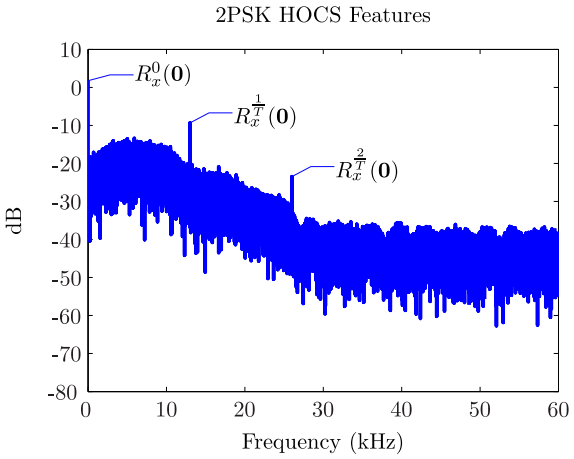
		Modulation Type			
		2PSK	4PSK	8PSK	16QAM
Noiseless	inf $\widehat{F}_x^0(\mathbf{0})_{8,4}$	27.86	24.90	24.71	26.44
	sup $\widehat{F}_x^0(\mathbf{0})_{8,4}$	28.18	25.03	24.84	26.63
$\frac{E_b}{N_0} = 5dB$	inf $\widehat{F}_x^0(\mathbf{0})_{6,3}$	17.7	15.91	14.83	15.46
	sup $\widehat{F}_x^0(\mathbf{0})_{6,3}$	18.26	16.39	15.38	15.83

Table 2: Range of values (in dB) of CHOCS features for $\gamma = 0.2$.

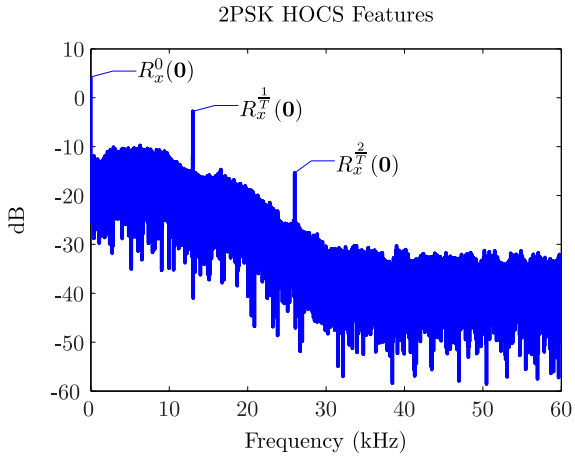
6. Davenport, M. A., Boufounos, P. T., Wakin, M. B. and Baranuik, R. G., "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing* **4**, pp. 445–460, Apr. (2010).
7. Lim, C. W. and Wakin, M. B., "Automatic Modulation Recognition for Spectrum Sensing using Nonuniform Compressive Samples," in *IEEE International Conference on Communications*, June (2012).
8. Dobre, O. A., Abdi, A., Bar-Ness, Y. and Su, W., "Higher-order cyclic cumulants for high order modulation classification," *MILCOM* **1**, pp. 112–117, Oct. (2003).
9. Rudelson, M. and Vershynin, R., "Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements," in *Information Sciences and Systems*, pp. 207–212, (2006).
10. Davenport, M. A. and Wakin, M. B., "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Transactions on Information Theory* **56**(9), pp. 4395–4401, (2010).
11. Gandhi, P. P. and Kassam, S. A., "Analysis of CFAR processors in homogeneous background," *IEEE Transactions on Aerospace and Electronics Systems* **24**(4), pp. 427–445, (1988).



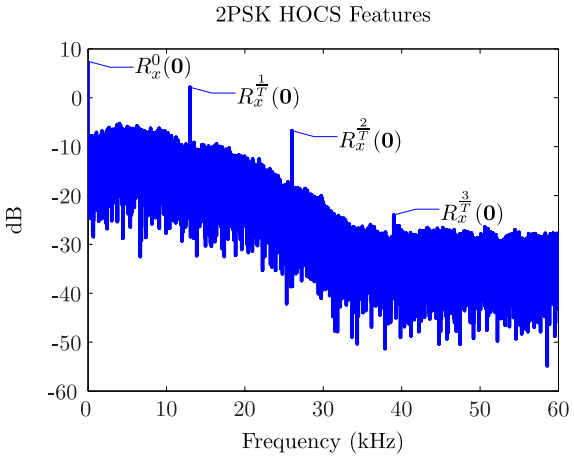
(a) $n = 2$, $q = 1$



(b) $n = 4$, $q = 2$



(c) $n = 6$, $q = 3$



(d) $n = 8$, $q = 4$

Figure 1: 2PSK HOCS Features.

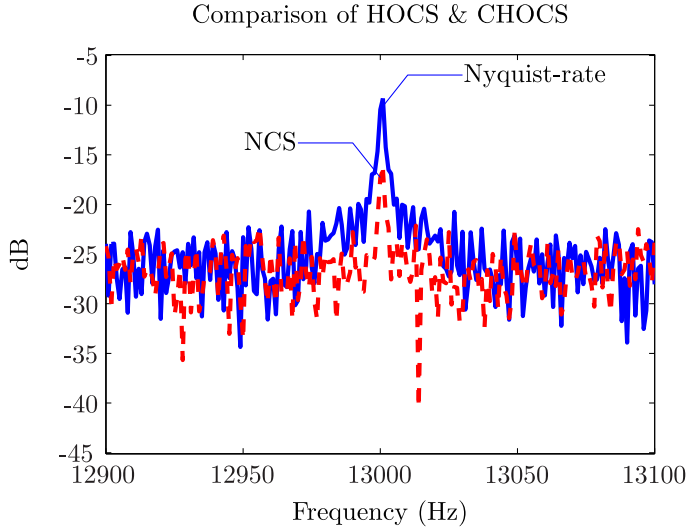
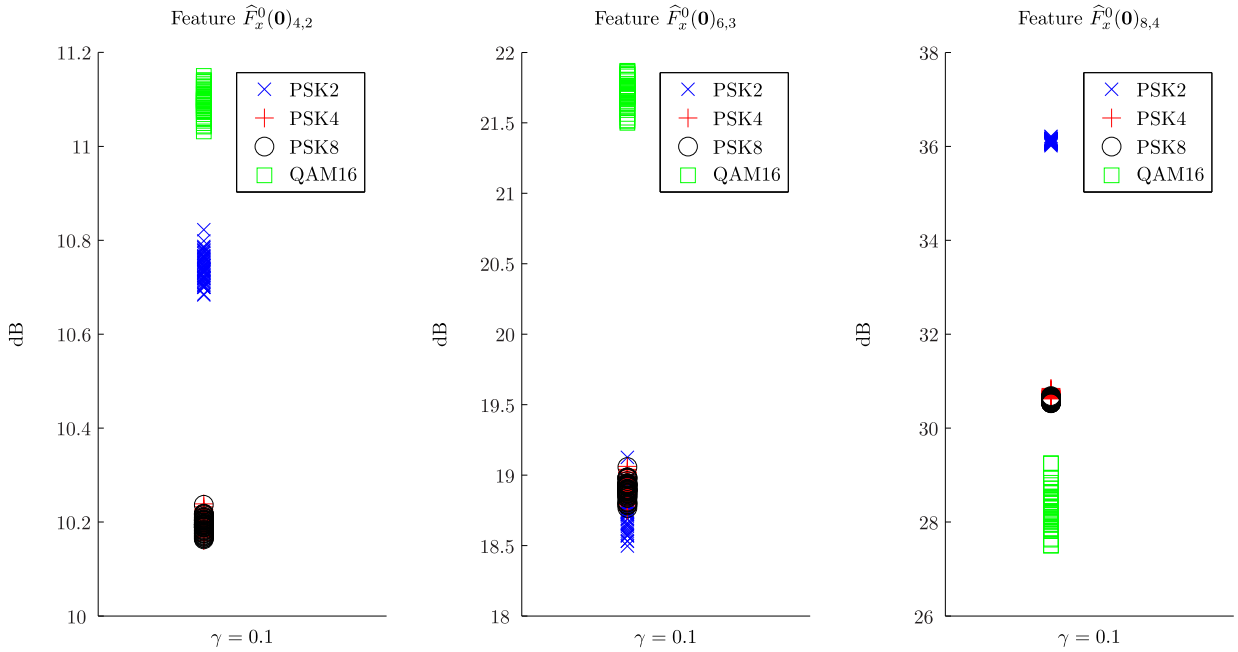


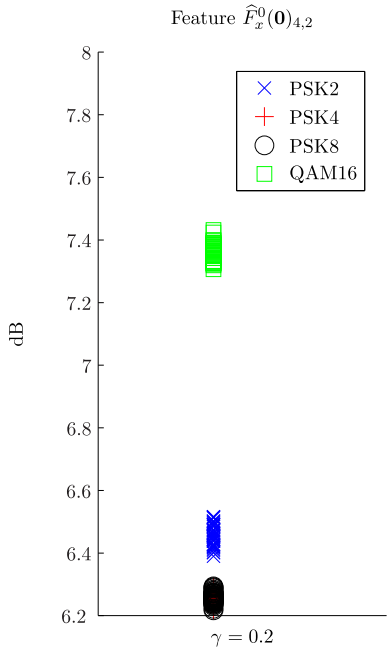
Figure 2: Comparison of HOCS and CHOCS.



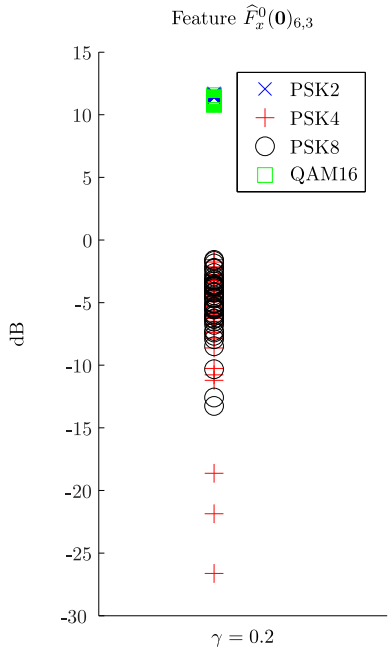
(a) $n = 4$, $q = 2$

(b) $n = 6$, $q = 3$

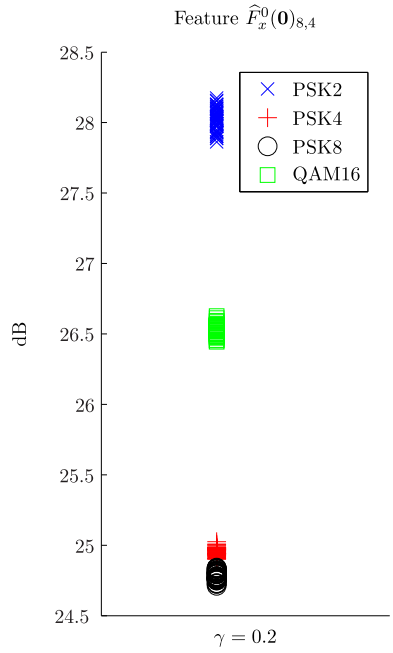
(c) $n = 8$, $q = 4$



(d) $n = 4$, $q = 2$

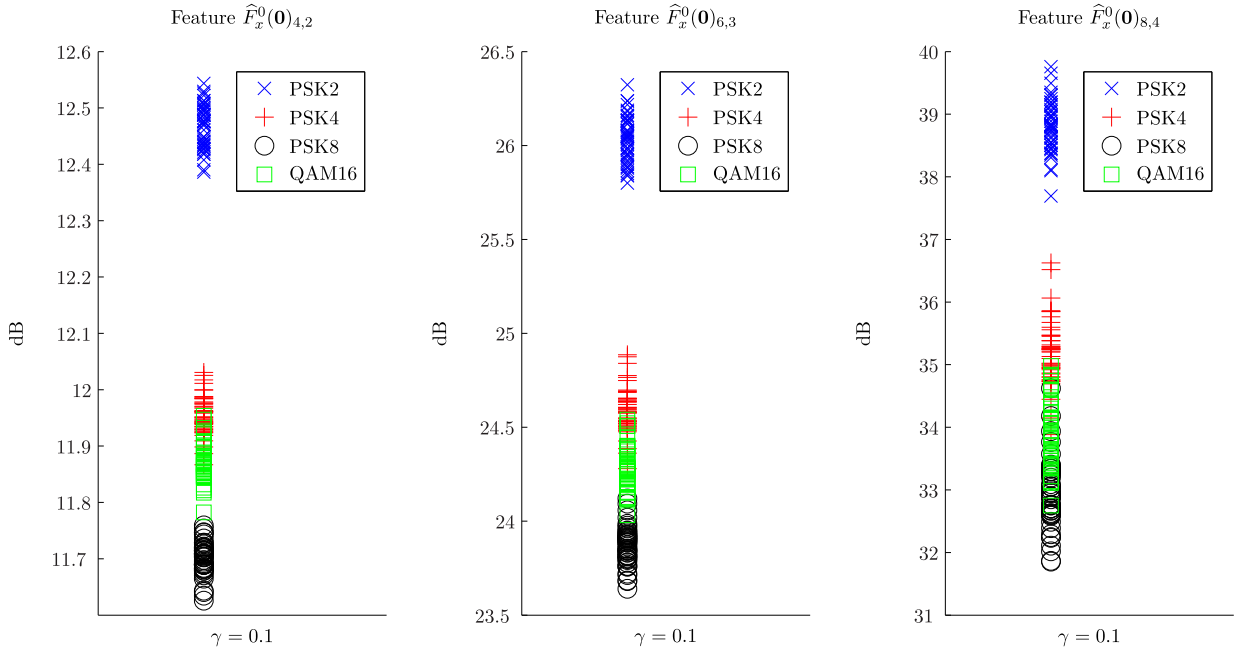


(e) $n = 6$, $q = 3$



(f) $n = 8$, $q = 4$

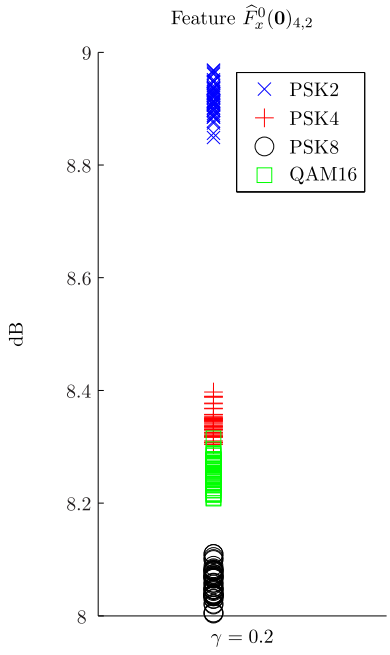
Figure 3: CHOCS Features for incoming signal with no noise.



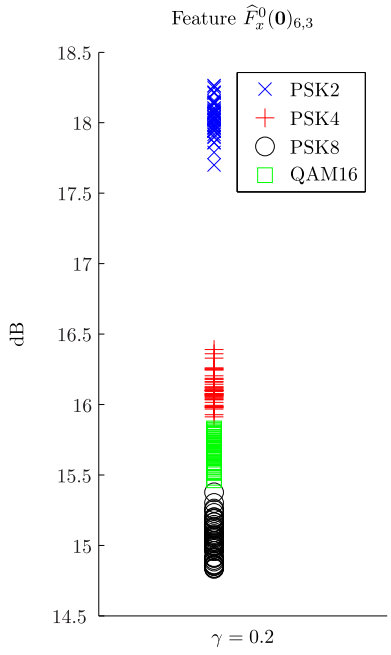
(a) $n = 4$, $q = 2$

(b) $n = 6$, $q = 3$

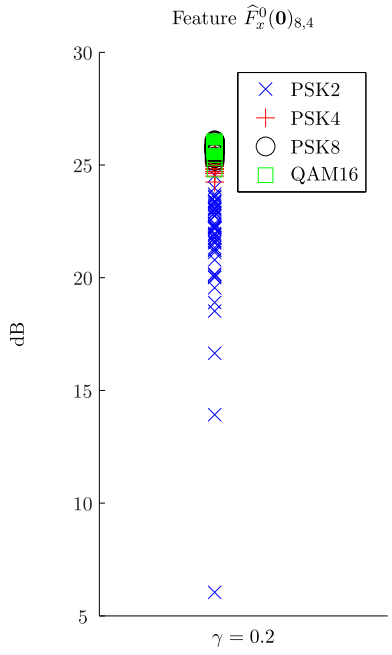
(c) $n = 8$, $q = 4$



(d) $n = 4$, $q = 2$

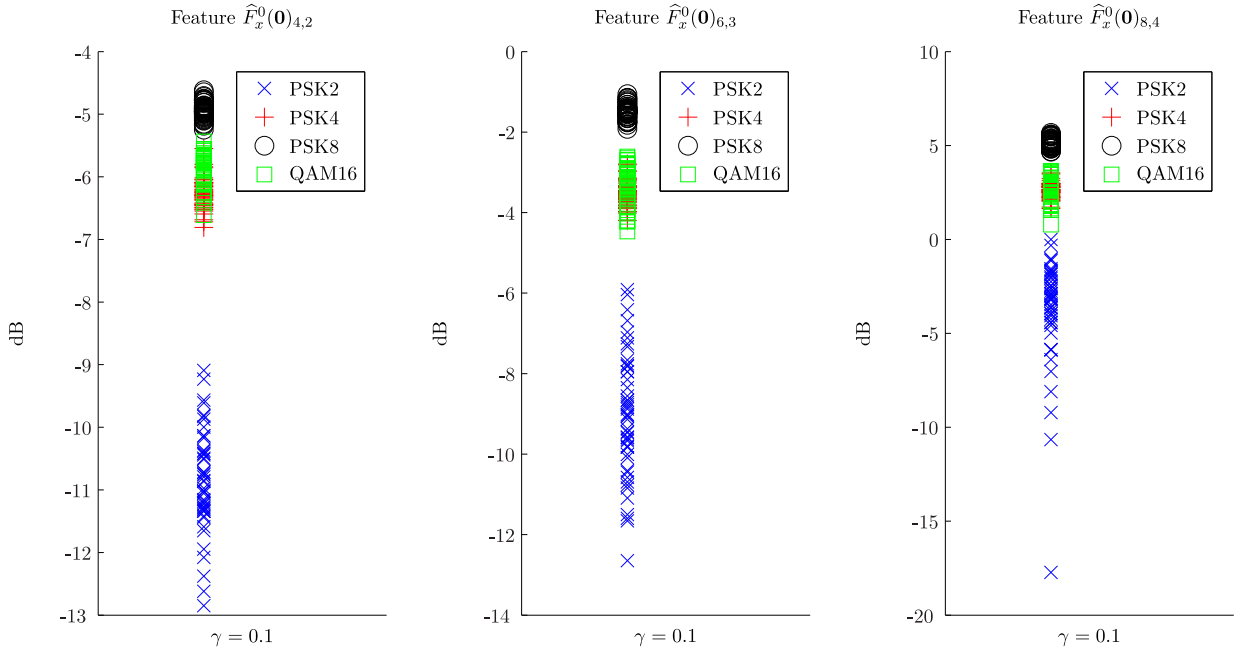


(e) $n = 6$, $q = 3$

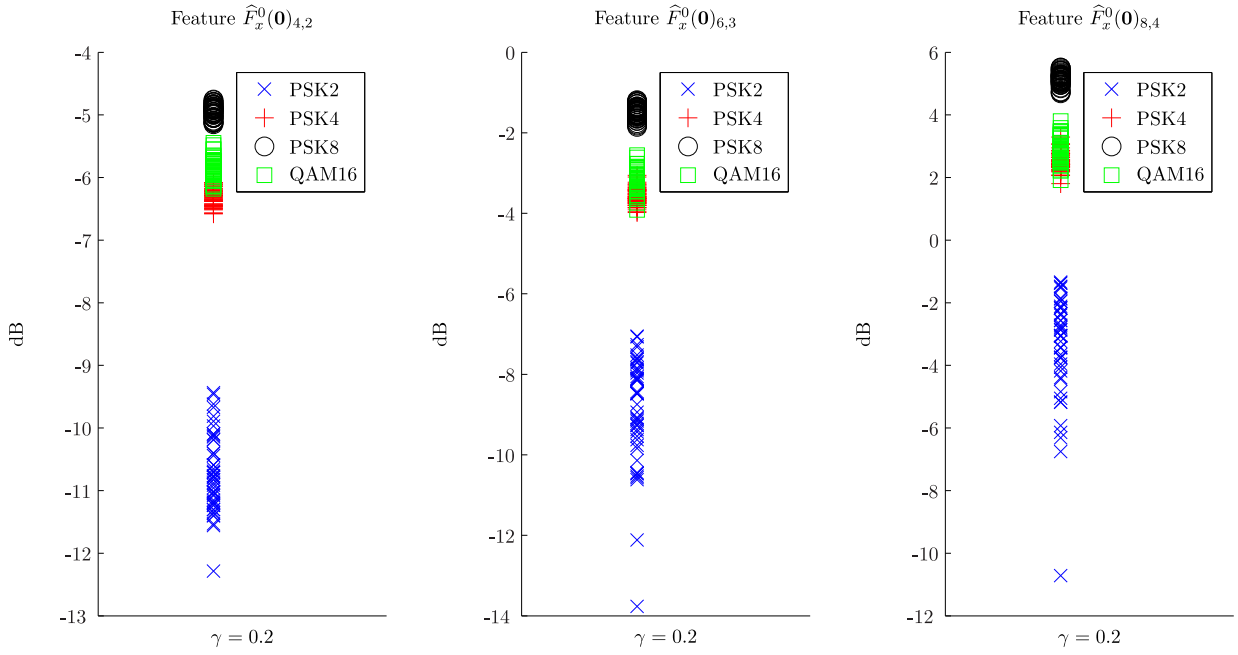


(f) $n = 8$, $q = 4$

Figure 4: CHOCS Features for incoming signal with $\frac{E_b}{N_0} = 5$dB.



(a) $n = 4$, $q = 2$ (b) $n = 6$, $q = 3$ (c) $n = 8$, $q = 4$



(d) $n = 4$, $q = 2$ (e) $n = 6$, $q = 3$ (f) $n = 8$, $q = 4$

Figure 5: HOCS Features for incoming signal with $\frac{E_b}{N_0} = 5$dB and duration γ seconds.