Iranian Journal of Science & Technology, Transaction B: Engineering, Vol. 34, No. B3, pp 257-274
Printed in The Islamic Republic of Iran, 2010

MODULATION CLASSIFICATION FOR BURST-MODE QAM SIGNALS IN MULTIPATH FADING CHANNELS*

H. R. NIKOOFAR AND A. R. SHARAFAT**

Dept. of Electrical and Computer Engineering, Tarbiat Modares University,
P.O. Box 14155-4838, Tehran, Iran
Email: sharafat@modares.ac.ir

Abstract– We present a new and efficient method for identifying the modulation type of a bursty QAM signal in the presence of additive white Gaussian noise (AWGN) in unknown fading channels and with unknown carrier phase. Our approach is based on an iterative combination of *blind equalization* and *soft-clustering* techniques, utilizes the constant modulus algorithm (CMA) for an initial reconstruction of the received signal constellation, and employs a nonparametric soft clustering method to identify a small set of possible modulations based on the partition coefficient (PC) criterion. We then use an iterative approach in utilizing the decision adjusted modulus algorithm (DAMA) to refine our choices until we reach two hypotheses. Besides, we propose a parallel implementation of the proposed method, and provide an asymptotic analysis as well as simulation results to demonstrate the validity and usefulness of the proposed method. Simulation results indicate that our proposed method achieves a significant gain as compared to the cumulant-based approach for burst mode transmissions in fading channels.

Keywords– Blind equalization, modulation classification, multipath fading channel, soft clustering

1. INTRODUCTION

Recent advances in wireless communications include concepts such as software defined radio (SDR) systems and cognitive radio networks. In both systems, a smart receiver that can adaptively tune into various unknown modulation formats and time-varying environments is required. In conventional receivers, supplementary information such as training and/or signalling data are periodically sent by the transmitter, which allows the receiver to recognize the transmitted signals with acceptable quality. However, transmission of such redundant and supplementary information may either not be possible, or would reduce the efficiency of transmission; therefore, in some systems, a blind adaptive algorithm for identifying the modulation type of the received signal is desirable. Furthermore, modulation classification is an active research area in the design of receivers for non-cooperative communication systems.

As defined in [1], *a modulation classifier (MC) is a system that, given a measurement* $r(t)$ $0 \le t \le T_o$, *recognizes the modulation type of* $r(t)$ *from N possible modulations* $\mathcal{A} = \{C_1, C_2, \ldots, C_N\}$. The received signal $r(t)$ is typically considered as a modulated signal received through, and corrupted by the communication channel and additive noise

$$r(t) = f_{ch}(s(t), h_c) + z(t) \tag{1}$$

where $s(t)$ is the modulated signal, h_c is the channel impulse response, and $z(t)$ is the additive white Gaussian noise. As quadrature amplitude modulation (QAM) is widely used in wireless communication

*Received by the editors October 20, 2009; Accepted June 7, 2010.
**Corresponding author

systems [2-3], we concentrate on the treatment of unknown linearly modulated QAM signals denoted by $s(t)$ in (1).

Modulation classification of QAM signals within a set of candidate constellations $\mathcal{A}$ has been studied in the past decade (see, e.g., [1, 3-7] and references therein). In doing so, two main approaches have been broadly followed, namely decision theoretic algorithms and pattern recognition schemes [4]. In the decision theoretic approach, the problem is formulated by a multiple hypothesis testing approach, and the decision is made based on the likelihood function for the received signal under each hypothesis. In [8], a quasi log-likelihood ratio classifier (qLLRC) is developed to distinguish BPSK from QPSK modulations in additive white Gaussian noise (AWGN) environments assuming a small value for the signal-to-noise ratio (SNR). In [9] and [10], this approach is extended to MPSK and QAM signals. In [11], the sequential probability ratio test (SPRT) is proposed for hypothesis testing to classify QAM signals. In [1] and [12], another likelihood-based approach is formulated that can be applied to any digital amplitude-phase modulation.

In maximum-likelihood (ML) classifiers, noise is assumed to be AWGN. The drawbacks of ML classifiers include very high computational complexity, and lack of robustness to model mismatches such as phase and frequency offsets, residual channel effects, timing errors, non-Gaussian noise distributions, and symbol sets with unequal sizes [13]. Such effects, which emanate from blind acquisition, synchronization, and equalization, are not easily modelled in the ML framework.

Compared to decision theoretic approaches, pattern recognition approaches take another viewpoint and attempt to extract some discriminating features from a received signal to identify the modulation type by using statistical classifiers (e.g., the K-nearest-neighbour classifier) [4]. In order to classify QAM signals in AWGN environments, some features such as constellation magnitude [14], higher-order cumulants [3], wavelets [5], cyclostationary properties [15], phase histogram [16], regular distance between constellation points [17-18], first zero-crossing location of the characteristic function [19], and constellation shape [20-22] are utilized. In [17], the regular distance between constellation points and the associated constellation size is used together with the characteristic function (2-D Fourier transform of the observed symbols' histogram) or 1-D Radon transform to estimate the constellation size. This feature can be successfully extracted for square and cross-shaped constellations when SNR is high and the residual ISI is negligible. In [19], the sensitivity of this feature to SNR variations is studied, and the first zero crossing (FZC) of the characteristic function is used as a more reliable method for estimating the constellation size. However, both methods are inappropriate for diamond-shaped constellations (e.g., V.29) with frequency offsets, and imperfect equalization conditions.

Approaches that exploit the constellation magnitude are insensitive to frequency offset and phase variations. However, their performances degrade with increase of noise level and constellation size. Among the aforementioned techniques, the constellation classification method proposed in [3], which uses the estimated fourth order cumulant of AWGN channel output, is promising due to its invariance to constellation rotation, its low computational complexity, and its robust performance, even at low SNR values. This method, however, requires an accurate estimate of noise variance in the channel and needs excessive data for correct classification.

Clustering algorithms in pattern recognition techniques [22] have also been applied for constellation identification. These include the agglomerative [23], the divisive [24], the hard C-means [12], and the fuzzy C-means [20] cluster analysis. All of these techniques are sensitive to SNR as well as to parameter values. Moreover, they allow cluster centres to scatter in all directions during the clustering process, which implies that the limited available information on candidate constellations is ignored [25].

All of the aforementioned methods for classification of QAM signals assume AWGN channels, and their performances may drastically deteriorate in multipath fading environments. In [17], a cascade

architecture that consists of a blind equalizer and a modulation classifier is employed for mitigating the impact of inter-symbol interference (ISI) that emanates from channel effects, as shown in Fig. 1. This concept is also used in [10] and [21]. We employ the same structure in this paper.

The main contribution of this paper is the development of a novel modulation classification technique for identifying the modulation type and the size of burst mode QAM signals in the presence of AWGN, multipath fading channel, and unknown phase offsets. Our proposed scheme is iterative and is based on a combination of blind equalization and soft-clustering classification techniques.

We introduce an iterative approach in the conventional blind equalizer-classifier cascade structure (BECC) by using the classifier's soft-output to iteratively adjust the equalizer's settings. By soft-output we mean that the classifier does not decide on one class as its final result (i.e., hard-output), but associates a degree of resemblance with each class. We also exploit the constellation information in our soft-clustering algorithm.

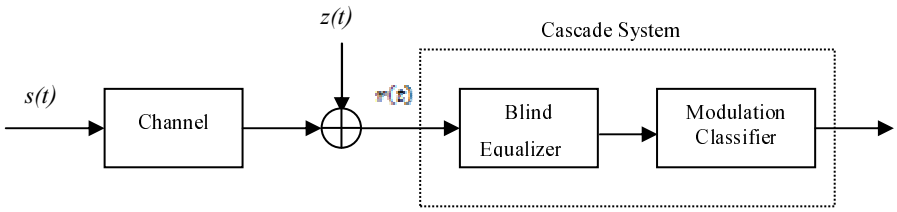


Fig.1. Conventional blind equalizer-classifier cascade structure

The rest of this paper is organized as follows. In Section II, we present the problem statement, derive the discrete-time models for a multipath dispersive channel and received signal, and introduce our assumptions. In Section III, we explain the new iterative BECC structure along with the corresponding blind equalization algorithms. In Section IV, the proposed scheme for modulation classification is presented. Our soft-clustering approach for constellation identification is in Section V. In Section VI, we present asymptotic performance analysis of our classifier for bursty signals at high SNR values. Simulation results are in Section VII, and conclusions are in Section VIII.

2. PROBLEM STATEMENT AND MODELLING

Given a M_i-QAM received signal $r(t)$ in the presence of additive white Gaussian noise, multipath fading, and absence of special side-knowledge (as explained later in this section), we wish to find the constellation size (M_i) and constellation shape. Specifically, we are interested in bursty signals, in which the burst length, the symbol rate, and the constellation size are varied in time.

In each burst, the transmitted signal is

$$s(t) = \mathrm{Re}\left\{ A\sum_{k=0}^{K-1} a_k^{(i)} g(t-kT) e^{j2\pi f_c t} \right\} \tag{2}$$

where f_c is the carrier frequency, T is the symbol duration, $g(t)$ represents the pulse shaping filter, A is the signal amplitude, and $\{a_k^{(i)}\}_{k=0}^{K-1}$ is a symbol sequence of length K drawn from one of the N constellations $\mathbb{A}=\{C_1,C_2,,...,C_N\}$. Each constellation C_i is characterized by M_i-array alphabet $\{S_1^i,S_2^i,\cdots,S_{M_i}^i\}$, i.e., $C_i=\{S_1^i,S_2^i,\cdots,S_{M_i}^i\}$. Our focus is limited to 8-PSK, 8QAM-v29, 16-QAM, 16-QAM-v29, 32-QAM, 64-QAM, 128-QAM, and 256-QAM modulations.

Figure 2 shows the block diagram of the communication system. The continuous time signal $r(t)$ observed at the receiver input is modelled as a distorted version of $s(t)$, degraded by the band-limited

multipath fading channel effects, and additive white Gaussian noise. The complex envelope of the received signal, after down-converting by a local carrier, is

$$r(t) = v(t) + z(t) = e^{j(2\pi\Delta f_0 t + \theta_c)} \int_{-\infty}^{+\infty} h(\tau, t)s(t - t_d - \tau)d\tau + z(t) \tag{3}$$

where $h(\tau,t)$ is the baseband equivalent of the channel impulse response (that may be time-varying), t_d is the timing epoch, Δf_0 is the residual carrier frequency (also known as frequency offset), θ_c is the carrier phase, and $z(t)$ is the base-band equivalent of the receiving noise.

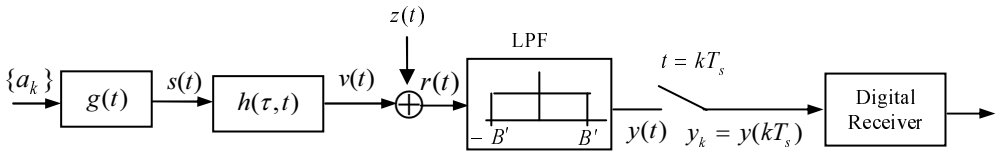


Fig. 2. Communication scheme in a multipath fading channel [26]

A multipath fading channel can be modelled as a time varying system [26]. The equivalent base band impulse response $h(\tau,t)$ represents the output of such a channel at time *t* if an input impulse is applied at time $t-\tau$. A tapped-delay-line (TDL) is often used to model a time varying multipath fading channel [26, 27], i.e.,

$$h(t,\tau) = \sum_{m=0}^{L_c} h_m(t)\delta(\tau - mT_s) \tag{4}$$

where T_s is the sampling period that satisfies $T_s \le 1/B$ (in which *B* is the signal bandwidth) and is chosen such that $T=N_sT_s$, *T* is the symbol period, N_s is the number of samples per symbol, (L_c+1) is the length of the channel in terms of T_s, and $h_m(t) = T_s h(t, mT_s)$ is the channel tap coefficient. The discrete-time model of the received signal, after ideal low-pass filtering with bandwidth $B' = B + f_D$ (where f_D is the maximum Doppler frequency of the multipath fading channel), and fractionally-spaced sampling at $t=kT_s$ is

$$y_k = y(kT_s) = e^{j(2\pi\Delta f_o T_s k + \theta_c)} \sum_{m=0}^{L_c} h_m(kT_s)s(kT_s - mT_s - t_d) + n(kT_s) \tag{5}$$

where *n(t)* is the noise at the output of the low-pass filter.

The following are assumed in this paper:

AS1. Constellation size is unknown, however it is in a given set, e.g., $\mathcal{A}1$= {8PSK, 8QAM}, $\mathcal{A}2$={V.29c, 16QAM}, $\mathcal{A}3$={8QAM, 8PSK, 16QAM, V.29c}, $\mathcal{A}4$= {128QAM, 256QAM}, or $\mathcal{A}5$={16QAM, V.29c, 32QAM}. The proposed scheme in this paper is also valid for any $\mathcal{A}_i \cup \mathcal{A}_j$.

AS2. Lack of special side-knowledge means that parameter values $A, \Delta f_0, \theta_c, t_d$, AWGN noise power, and channel impulse response $h(\tau,t)$ in (2)-(5) are generally unknown. However, we estimate the symbol rate $R = 1/T$ and the carrier frequency offset Δf_0 by using the cyclostationary characteristics of the received signal [28-30]. Symbol timing is recovered by standard fractional sampling as in (5).

AS3. There is no direct path between the transmitter and the receiver. Based on the central limit theorem, $h(\tau,t)$ can be modelled as a complex-valued Gaussian random process [31], whose

mean value is zero, $|h(\tau,t)|$ at any instant t is Rayleigh distributed [26], and the corresponding wireless channel is a Rayleigh multipath fading channel.

AS4. Using (5), the channel coefficients at time $t=kT_s$ are

$$\mathbf{h}_k = \left[h_{0,k}, h_{1,k}, \cdots, h_{L_c,k}\right]^T \tag{6}$$

where $h_{m,k} = h_m(kT_s)$, and $[.]^T$ denotes transposition. Assuming uncorrelated scattering [31], channel tap coefficients are treated as a zero-mean complex Gaussian random process with the following covariance matrix [26]

$$\mathbf{R}_j = \tfrac{1}{2} E\left[\mathbf{h}_{k+j}\mathbf{h}_k^{\ H}\right] = J_0(2\pi f_D j T_s)\boldsymbol{\Sigma} \tag{7}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, f_D is the maximum Doppler frequency, and $(.)^H$ denotes the Hermitian of a matrix. The diagonal matrix $\sum$ is

$$\boldsymbol{\Sigma} = diag\left[\sigma_0^2, \sigma_1^2, \cdots, \sigma_{L_c}^2\right] \tag{8}$$

where $\sigma_m^2 = \tfrac{1}{2} E\left[\left|h_{m,k}\right|^2\right]$ is obtained from the channel multipath intensity profile (MIP), also known as the multipath delay-power profile [26-27].

AS5. Data is transmitted burst-by-burst, and the channel is slow fading (i.e., $f_d T < 0.01$), such that the channel tap coefficients do not change during each burst (i.e., $h_{m,k} = h_m$ and $\mathbf{R}_j = \boldsymbol{\Sigma}$). However, the coefficients' values may change from one burst to another.

AS6. Channel noise $z(t)$ is additive, and is modelled as a circularly symmetric zero-mean white complex Gaussian random process having an autocorrelation function $R_z(\tau) = N_0\delta(t)$, where N_0 is the noise power spectral density. Thus, $n_k = n(kT_s)$ in (5) is a circularly symmetric discrete complex Gaussian random process with variance $\sigma_g^2 = E[|n_k|^2] = 2B'N_0$. The average SNR is [27]

$$SNR_{\ AVERAGE} = \frac{\frac{1}{2}A^2T\sum_{m=0}^{L_c} E\left[\left|h_{m,k}\right|^2\right]}{\sigma_g^2} = \frac{A^2T\sum_{m=0}^{L_c}\sigma_m^2}{\sigma_g^2}. \tag{9}$$

In order to increase the utilization of the channel's capacity, the transmitter is assumed to adapt its symbol rate and modulation type to channel conditions [2, 32]. That is, when the transmitter finds a channel in which higher transmission rates are permissible, it switches to a higher order modulation. This means that under such circumstances, channel conditions are more stable and the burst length is relatively long (there are more symbols available in one burst). In contrast, transmitting on lower order modulations means a bad fading channel and short burst length. Therefore, when the constellation size is large, bursts are relatively long and more symbols are transmitted in each burst.

3. NEW BLIND EQUALIZATION-CLASSIFICATION CASCADE STRUCTURES

Modulation Classification (MC) in unknown dispersive channels in [13, 14, 17, 21] are based on a cascade equalizer-classifier set up as shown in Fig.1, and use of the constant modulus algorithm (CMA) [33] for blind equalization. In this, what follows is, we first review the strengths and weaknesses of CMA in the conventional blind equalizer-classifier cascade (BECC) structure, and state our motivation for its

improvement by using the decision adjusted modulus algorithm (DAMA) [34-35]. Subsequently, we introduce our novel iterative and parallel BECC structures.

The CMA equalizer has an adaptive linear transversal structure, as shown in Fig. 3. The equalizer tap coefficient vector $\mathbf{w}$ is chosen in such a way to minimize the CMA cost function, defined by [33]

$$\xi_{CMA} \overset{\Delta}{=} \frac{1}{4}\mathrm{E}[(|\hat{y}_k|^2 - R)^2], \tag{10}$$

where $\hat{y}_k$ is the output of the blind equalizer, and R is the dispersion constant [33]

$$R = \mathrm{E}[|a_k^{(i)}|^4] \Big/ \mathrm{E}[|a_k^{(i)}|^2]. \tag{11}$$

The expected values in (11) are taken over the constellation points $\{S_\ell^i\}_{\ell=1}^{M_i}$ assuming that all M_i points are equally probable. Using fractionally spaced CMA (FSE-CMA) [33] and LMS (least mean square) implementation of the algorithm, the equalizer tap coefficients are iteratively updated by

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mu \mathbf{x}_k^* \hat{y}_k (|\hat{y}_k|^2 - R)^2 \tag{12}$$

where $\mathbf{w}_k = [w_{0,k}, w_{1,k}, \cdots, w_{L_e,k}]^T$ is the equalizer tap coefficient vector, $\mathbf{x}_k = [x_{0,k}, x_{1,k}, \cdots, x_{L_e,k}]^T$ is the equalizer input vector at time k, μ is the step size, L_e is the length of FIR equalizer, and * denotes complex conjugate.

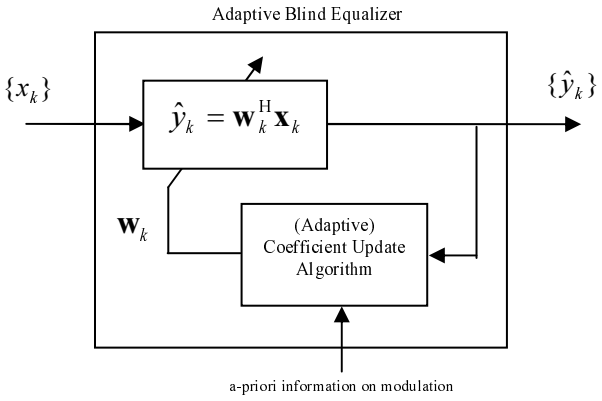


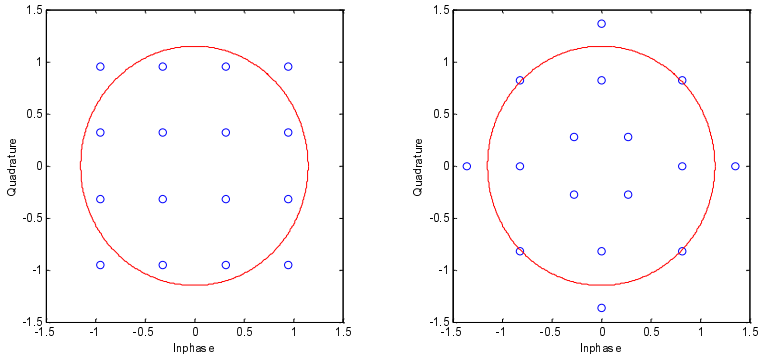
Fig. 3. Linear blind equalizer structure

The dispersion constant R in (11) depends on modulation type, and its values for different modulation types are shown in Table I. Figs. 4a and 4b show circles with radius $\sqrt{R}$ for 16QAM and 16QAM-V29c constellations, respectively.

While CMA can be successfully applied to QAM signals [34], its cost function (10) is better matched to PSK in the sense that $R^{(\mathrm{PSK})} = |a_\ell|^2$, and the cost function goes to zero at all signal points of a PSK constellation (when $R = R^{(\mathrm{PSK})}$). For a QAM constellation C_i, constellation points lie on the multiple radii $\{\lambda_d\}_{d=1}^{D_i}$, also known as multiple modulus, and typically $R \neq \lambda_d^2, \ \forall d, 1 \le d \le D_i$. Note that 16QAM-V29c, as shown in Fig. 4(b), is an exception. This means that (10) does not go to zero for any constellation point in the majority of standard QAM constellations. See, for example Fig. 5, where multiple modulus of 16QAM and 16QAM-V29c are shown. The mismatch between the constellation and the CMA's cost function causes excessive misadjustments for QAM signals [35].

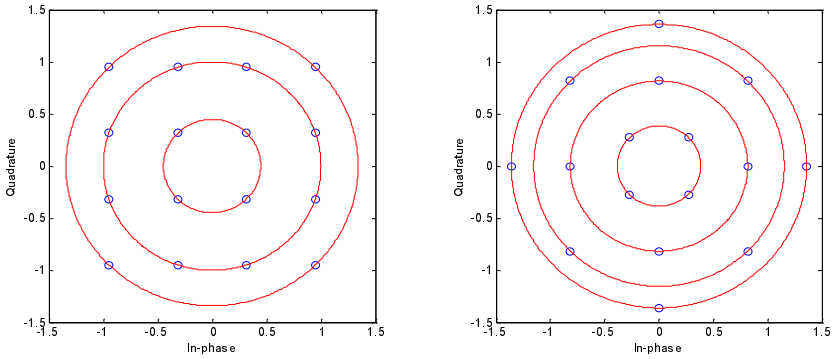
Table 1. Values of R for different QAM modulations

Modulation Type	R
8 PSK	1.0000
8 QAM-V29	1.405
16 QAM	1.320
16QAM-V.29c	1.4189
32 QAM	1.310
64 QAM	1.381
128 QAM	1.3427
256QAM	1.3953



(a) (b)

Fig. 4. Circles with radius equal to the square root of the dispersion constant for (a) Square 16QAM (b) Star shaped 16QAM-V.29c



(a) (b)

Fig. 5. Circles with radius equal to multi-modulus $\{\lambda_d\}_{d=1}^{D_i}$ of (a) 16QAM : $D=3, \lambda=\{0.4472,1,,1.3417\}$ (b) 16QAM-V.29c: $D=4, \lambda=\{0.3849,0.8165,1.1547,1.3607\}$

Since no prior knowledge on the underlying constellation is assumed, the CMA equalizer sets an initial arbitrary modulus R in cost function (10), regardless of the modulation type. As shown in [21], if the value of R is not far from its exact value $R^{(i)}$, it does not significantly impact misadjustment and speed of convergence of the CMA equalizer. For this reason, we use $R=1.3$ for the set of constellations in Table I.

By using the CMA equalizer in a conventional BECC structure, the residual channel effects are not sufficiently equalized for large constellation size QAM signals $(M_i>16)$ and the remaining ISI renders our estimate of the constellation shape inaccurate. Furthermore, due to slow convergence of CMA, its

performance is not satisfactory, especially for burst transmissions where the number of received symbols is limited.

To tackle this issue, we propose a novel structure that operates iteratively. Figure 6 shows our proposed structure, in which we perform joint equalization and classification repeatedly, and combine two different adaptive algorithms for blind equalization in sequence, namely the constant modulus algorithm (CMA) and the decision adjusted modulus algorithm (DAMA) in [35]. The main objective of employing DAMA instead of CMA in subsequent iterations is to reduce the misadjustment.

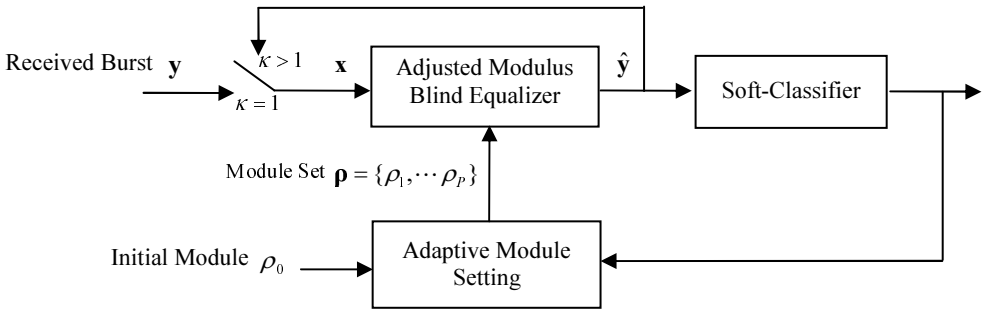


Fig. 6. Block diagram of the proposed iterative equalizer-classifier cascade

For unknown constellations, we use a modified DAMA as in [35]

$$\xi_{DAMA} = \frac{1}{4}\mathrm{E}\left[\min_{p}\left(\left|\left|\hat{y}_k\right|^2 - {\rho_p}^2\right|\right)^2\right],\ 1 \le p \le P, \tag{13}$$

where $\hat{y}_k$ is the output of the blind equalizer and $\{\rho_p\}_{p=1}^{P}$ are the arbitrary radii. The coefficient update equation for DAMA equalizer is

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mu \mathbf{x}_k^* \hat{y}_k (|\hat{y}_k|^2 - \rho_\nu^2), \tag{14}$$

where

$$\nu \overset{\Delta}{=} \arg\min_{p}\left(\left|\hat{y}_k\right|^2 - {\rho_p}^2\right). \tag{15}$$

4. PROPOSED SCHEME FOR MODULATION CLASSIFICATION

As shown in Fig. 6, we adjust the number of radii P and the values $\{\rho_i\}_{i=1}^{P}$ iteratively by using the soft classifier output. Box 1 describes the iterative procedure in this structure, and Figs. 7(a) and 7(b) show its corresponding block diagram.

In iteration 1, input burst goes through CMA equalizer (the E-Step in Box 1). In the next step, the soft classification algorithm is applied to the equalizer's output, and the degree of resemblance $\Re(C_i \mid \hat{\mathbf{y}})$ of the reconstructed constellation to every modulation type in the initial set of candidate modulations $\mathcal{A}^{(0)} = \mathcal{A}$ is calculated (the C-Step in Box 1). Now, we use a threshold to determine a smaller subsequent set of possible modulations $\mathcal{A}^{(1)}$ for the D-Step in Box 1. The combination of these steps (E, C, and D) is shown in Fig. 7a. We use the decision in the first iteration $\mathcal{A}^{(1)}$ in DAMA (the S-step in Box 1, and Fig 7b). In the second iteration and further, pre-equalized symbols are processed by the adjusted modulus DAMA equalizer followed by a soft-classifier and a threshold comparator as shown in Fig 7b. This procedure is terminated when we reach a two class problem, and the one that has a larger value of $\Re$ is selected.

Box 1. Iterative procedure for the BECC structure

I-Step: Initializing

Iteration index $\kappa = 0$,

Compute R for each modulation in the set $\mathcal{A}$:

$$R^{(i)} = E[|a_k^{(i)}|^4] \Big/ E[|a_k^{(i)}|^2], \quad a_k^{(i)} \in C_i, \ C_i \in \mathcal{A}$$

Set:

$$\rho_0 = \frac{1}{N}\sum_{i=1}^{N} R^{(i)} \qquad \boldsymbol{\rho}^{(0)} = \{\rho_0\}$$

$$\mathcal{A}^{(0)} = \mathcal{A}$$

Input Burst: $\mathbf{y} = [y_1, y_2, \cdots, y_K]^T$

$\mathbf{x} = \mathbf{y}$

E-Step: Blind Equalization

Iteration index $\kappa = \kappa + 1$

Input Burst: $\mathbf{x} = [x_1, x_2, \cdots, x_K]^T$

Reverse Burst : $\bar{\mathbf{x}} = [x_K, x_{K-1}, \cdots, x_1]^T$

Virtual Burst : $\hat{\mathbf{x}}_{\text{Virtual}} = [\mathbf{x}^T, \bar{\mathbf{x}}^T, \mathbf{x}^T, \bar{\mathbf{x}}^T, \mathbf{x}^T]^T$

Run {Adaptive DAMA Equalization with $\boldsymbol{\rho} = \boldsymbol{\rho}^{(\kappa-1)}$

Input : $\hat{\mathbf{x}}_{\text{Virtual}}$

Output: $\hat{\mathbf{y}}_{\text{Virtual}} = [\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_{5K}]^T$ }

$\hat{\mathbf{y}} = [\hat{y}_{4K+1}, \hat{y}_{4K+2}, \cdots, \hat{y}_{5K}]^T$

$\mathbf{x} = \hat{\mathbf{y}}$

C-Step: Soft-Classification

Run {Soft-Classification Algorithm

Input: $\hat{\mathbf{y}}$ (recovered constellation)

Output: $\Re(C_i \mid \hat{\mathbf{y}})$ (degree of resemblance (likelihood) of recovered constellation with respect to constellation C_i in the set $\mathcal{A}^{(\kappa-1)}$) }

D-Step: Narrowing Down Modulations Set

$\mathcal{A}^{(\kappa)} = \{C_i \mid \Re(C_i \mid \hat{\mathbf{y}}) \geq \gamma^{(\kappa)}, C_i \in \mathcal{A}^{(\kappa-1)}\}$

if there are two remaining modulations

$$i^* = \arg\max_i \{\Re(C_i \mid \hat{\mathbf{y}}) \mid C_i \in \mathcal{A}^{(\kappa)}\}$$

Break (%END OF CLASSIFICATION}

End

S-Step : Adaptive Modulus Setting

Select adjusted-modulus $\boldsymbol{\rho}^{(\kappa)} = \{\rho_0, \rho_1, \cdots, \rho_P\}$ that better matches the remaining modulations in the set $\mathcal{A}^{(\kappa)}$

Go to **E-Step** (next iteration)

In the first iteration (Fig 7a), CMA is used as a pre-equalizer, and despite its slow convergence and relatively high misadjustment (especially for large QAM constellations), serves two purposes. First, it reduces channel effects on the received signal, thus enabling the classifier to reduce the number of possible modulations for subsequent iterations. Second, it sufficiently equalizes amplitude distortions caused by fading, so that DAMA can converge [35]. In fact, DAMA requires that amplitude distortions be sufficiently equalized.

For a fast procedure, we use parallel DAMA structures as in Fig. 8, in which the adjusted-modulus $\{\rho_p\}_{p=1}^{P}$ of DAMA equalizers are matched to the exact multi-modulus $\{\lambda_d\}_{d=1}^{D}$ of the corresponding modulation C_i with a view to improving the performance of the equalizer in removing the channel effects [35]. Of course, this requires additional computations in the receiver.

Since the burst length is usually shorter than what is required for convergence of the blind equalizer, we reuse the captured burst in a forward/reverse scenario to generate a virtual burst that is long enough for the blind equalizer to converge (see E-step in Box 1). In [36], it is shown that if the number of symbols in a burst exceeds the number of possible channel outputs, convergence of the blind equalizer can be achieved similar to the case of infinite data length.

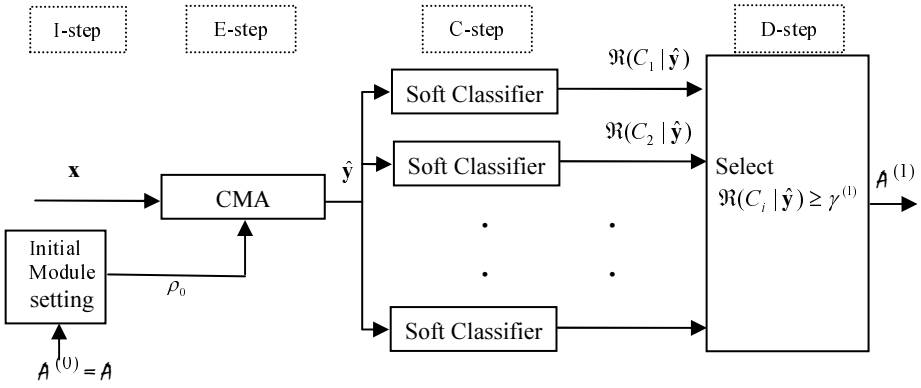


Fig. 7a. Block diagram of *I-D Steps* in the first iteration of iterative BECC structure

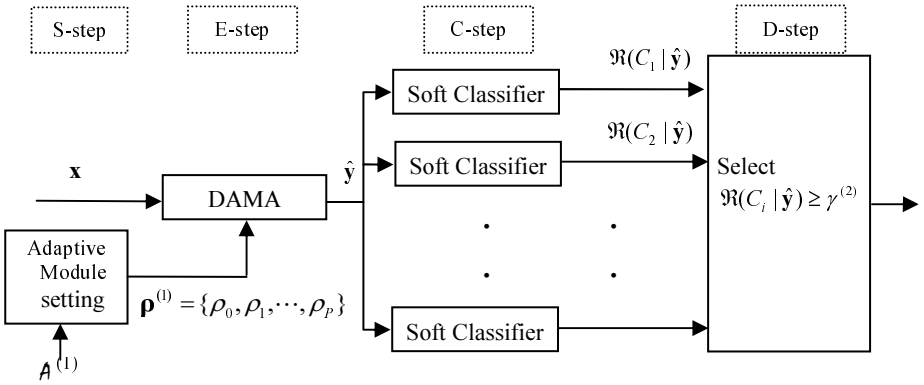


Fig. 7b. Block diagram of *S-D Steps* in the second iteration of iterative BECC structure

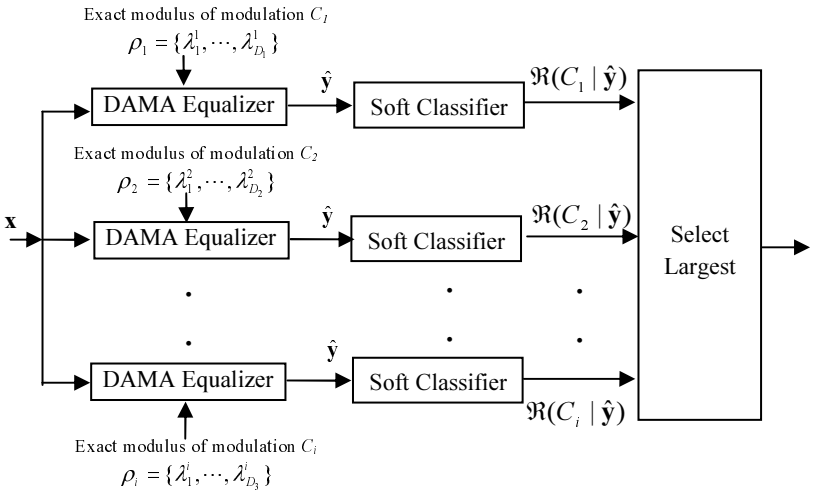


Fig. 8. Block diagram of *E-D Steps* in the second iteration of hybrid (parallel) BECC structure

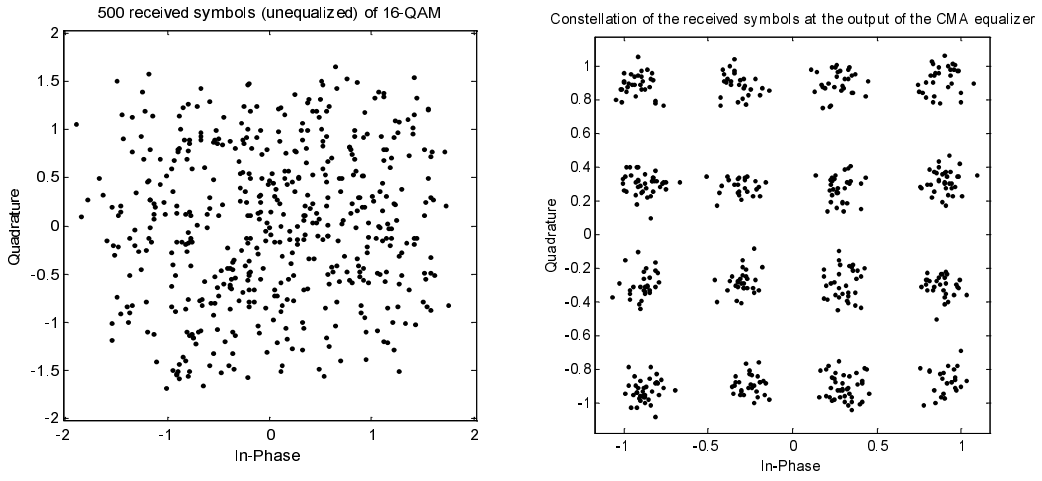
5. SOFT CLUSTERING ALGORITHM FOR MODULATION CLASSIFICATION

We now introduce our soft clustering algorithm for modulation classification, which exploits the constellation shape as a key signature of the modulation type. We consider the sequence of reconstructed symbols $\hat{\mathbf{y}}$ at the output of the blind equalizer as a geometrical pattern in the inphase-quadrature (IQ) scatter diagram in Fig. 9, and develop a soft clustering method that measures the degree of resemblance of the estimated constellation $\hat{\mathbf{y}}$ to constellation C_i in the set of candidate modulations $\mathcal{A} = \{C_1, C_2, ..., C_N\}$. We use this measure as a decision-making parameter introduced in Section III.

In order to deal with the effects of noise, residual channel effects, and unknown phase offsets, we take a soft clustering approach to identify the modulation type of the received bursty signal. In doing so, we note that fuzzy-based clustering methods are useful in such applications (e.g., [37]). Soft clustering means that any given data point $\hat{y}_k$ may belong to several clusters (with cluster centres at $\{S_1^i e^{j\theta}, S_2^i e^{j\theta}, \cdots, S_{M_i}^i e^{j\theta}\}$) each with a different fuzzy membership value $u_{\ell k}^i \in [0,1]$, which is a measure of similarity between the equalized symbol $\hat{y}_k$ and the rotated symbol $S_\ell^i e^{j\theta}$, where S_ℓ^i is a symbol in constellation C_i, i.e., $S_\ell^i \in \{S_1^i, S_2^i, \cdots, S_{M_i}^i\}$. The fuzzy membership matrix $\mathbf{U}_i = [u_{\ell k}^i]$ is a measure of similarity between the equalized symbol vector $\hat{\mathbf{y}}$ and the rotated constellation $C_i e^{j\theta} = \{S_1^i e^{j\theta}, S_2^i e^{j\theta}, \cdots, S_{M_i}^i e^{j\theta}\}$.

We wish to find a membership matrix $\mathbf{U}_i$ that minimizes the following cost function

$$\Omega_i = \sum_{\ell=1}^{M_i} \sum_{k=1}^{K} {u_{\ell k}^i}^2 \left| \hat{y}_k - S_\ell^i e^{j\theta} \right|^2 \tag{16}$$



(a) (b)

Fig. 9a. A typical received signal constellation of 16 QAM modulation with SNR = 30 dB, (b) – Signal constellation of the received symbols at the output of CMA equalizer

where $d_{\ell k}^i = \left| \hat{y}_k - S_\ell^i e^{j\theta} \right|$ is the Euclidian distance between an equalized symbol $\hat{y}_k$ and a rotated symbol $S_\ell^i e^{j\theta}$, where $S_\ell^i \in C_i$.

For fuzzy partitioning, the elements of $\mathbf{U}_i$ are between 0 and 1, and normalization requires that the sum of membership functions for each $\hat{y}_k$ be equal to unity, therefore

$$\sum_{l=1}^{M_i} u_{\ell k}^i = 1, \quad \forall k = 1, \cdots, K \tag{17}$$

The minimum of cost function Ω can be obtained by forming the Lagrangian cost function

$$\tilde{\Omega}(\mathbf{U}_i,\theta,\zeta_1,\cdots,\zeta_k)=\sum_{l=1}^{M_i}\sum_{k=1}^{K}u_{\ell k}^{i\ 2}d_{\ell k}^{i\ 2}-\sum_{k=1}^{K}\zeta_k(\sum_{l=1}^{M_i}u_{\ell k}^{i}-1) \tag{18}$$

where ζ_k for k =1 to K are Lagrangian multipliers. Setting all gradients of $\tilde{\Omega}$ with respect to each input argument equal to zero yields

$$u_{\ell k}^{i}=1\Big/\sum_{s=1}^{M_i}(\frac{d_{\ell k}^{i}}{d_{sk}^{i}})^2 \tag{19}$$

and from (16)

$$\theta=\Lambda(\mathbf{U},C_i,\hat{\mathbf{y}})=\tan^{-1}\left(\frac{\sum_{\ell=1}^{M_i}\sum_{k=1}^{K}u_{\ell k}^{i\ 2}\{\mathrm{Im}(\hat{y}_k)\,\mathrm{Re}(S_\ell^i)-\mathrm{Re}(\hat{y}_k)\,\mathrm{Im}(S_\ell^i)\}}{\sum_{\ell=1}^{M_i}\sum_{k=1}^{K}u_{\ell k}^{i\ 2}\{\mathrm{Re}(\hat{y}_k)\,\mathrm{Re}(S_\ell^i)+\mathrm{Im}(\hat{y}_k)\,\mathrm{Im}(S_\ell^i)\}}\right) \tag{20}$$

Equations (19) and (20) show that membership values $u_{\ell k}^{i}$ and unknown phase θ are nonlinearly inter-related. To obtain θ and $\mathbf{U}_i$, we propose the iterative procedure in Box 2.

Box. 2. Iterative Soft Clustering.

CLU-0: Initialize parameters for $n=0$

Set modulation under test: $C_i=\{S_1^i,S_2^i,\cdots,S_{M_i}^i\}$,

Select a stopping threshold δ

Find the starting point:

- $\mathbf{U}_i(\theta)=[u_{\ell k}(C_i e^{j\theta},\hat{\mathbf{y}})]\big|_{\theta=0,\frac{\pi}{16},\frac{3\pi}{8},\frac{7\pi}{16}}$ (use Eq.19)
- $\hat{\theta}^{(0)}=\underset{\theta\in\{0,\frac{\pi}{16},\frac{3\pi}{8},\frac{7\pi}{16}\}}{\arg\min}\{\Omega(\mathbf{U}^i(\theta),C_i\,e^{j\theta},\hat{\mathbf{y}})\}$ (use Eq. 16)
- $\hat{C}_i^{(0)}=C_i e^{j\hat{\theta}^{(0)}}$
- $\mathbf{U}_i^{(0)}=\mathbf{U}_i(\hat{\theta}^{(0)})$
- $\hat{\Omega}^{(0)}=\Omega(\mathbf{U}_i(\hat{\theta}^{(0)}),C_i e^{j\hat{\theta}^{(0)}},\hat{\mathbf{y}})$

CLU-1: Compute parameters for iteration n+1

- $\theta^{(n+1)}=\theta^{(n)}+\Lambda(\mathbf{U}_i^{(n)},\hat{C}_i^{(n)},\hat{\mathbf{y}})$ (use Eq. 20)
- $\hat{C}_i^{(n+1)}=\hat{C}_i^{(n)}e^{j\hat{\theta}^{(n+1)}}$
- $\hat{\mathbf{U}}_i^{(n+1)}=[u_{\ell k}(\hat{C}_i^{(n+1)},\hat{\mathbf{y}})]$

CLU-2: Convergence checking

- $\hat{\Omega}_i^{(n+1)}=\Omega(\hat{\mathbf{U}}_i^{(n+1)},\hat{C}_i^{(n+1)},\hat{\theta}^{(n+1)})$
- If $\left|\Omega_i^{(n+1)}-\Omega_i^{(n)}\right|\le\delta$

  Break (%END OF CLUSTERING)

  End
- Go to step **CLU-1**

We use the elements of $\mathbf{U}_i$ to examine how data points correspond to modulation type C_i. In doing so, we employ the partition coefficient (*PC*) in [38] as the degree of resemblance $\Re(C_i\,|\,\hat{\mathbf{y}})$. Given a $M_i\times K$ membership matrix $\mathbf{U}_i$, the *PC* of $\mathbf{U}_i$ is

$$\Re(C_i\,|\,\hat{\mathbf{y}})=PC(\mathbf{U}_i\,|\,\hat{\mathbf{y}})=\frac{1}{K}\sum_{\ell=1}^{M_i}\sum_{k=1}^{K}u_{\ell k}^{i\ 2} \tag{21}$$

A value of $PC(\mathbf{U}_i)$ close to 1 indicates a good classification, and a value equal to $1/M_i$ indicates no cluster. In practice, the use of *PC* is not suitable when *K* and $\{M_i\}_{i=1}^{N}$ have a wide range, as *PC* is sensitive to the number of clusters *N* and the length of data *K*. It has been shown that *PC* identifies constellations with smaller values of M_i [37].

We use this soft clustering algorithm as our classification method to compute the degree of resemblance $\Re(C_i \mid \hat{\mathbf{y}})$ in the conventional (Fig. 1), in the iterative (Fig. 7), and in the parallel (Fig. 8) BECC structures in Section III. We compare the asymptotic performances of these structures for classification of QAM signals in AWGN and multipath fading channels in Section VI.

6. ASYMPTOTIC PERFORMANCE ANALYSIS

In this section, we show that the performance of the proposed fuzzy-based clustering scheme approaches that of the optimum classifier, i.e., the generalized likelihood ratio classifier (GLR). However, our proposed scheme is non-parametric, meaning that it is more robust compared to parametric (e.g., the maximum likelihood) classifiers. This is an important point for fading channels.

Due to the complexity associated with fuzzy parameters, it is not easy to analyze the performance of our proposed scheme in general. As an alternative, we can consider two asymptotic conditions, namely long bursts ($K \to \infty$), and high SNR ($SNR \to \infty$). Since our focus is on the burst mode, we present the asymptotic analysis for $SNR \to \infty$.

We use the average probability of correct classification P_{CC} as a measure of performance

$$P_{CC} = \frac{1}{N}\sum_{i=1}^{N} \Pr(\text{success} \mid C_i) \tag{22}$$

where $\Pr(\text{success} \mid C_i)$ means the probability of identifying C_i when C_i was transmitted. For the proposed soft clustering classifier, $\Pr(\text{success} \mid C_i)$ is

$$\Pr(\text{success} \mid C_i) = \Pr\{\Re(C_i \mid \hat{\mathbf{y}}) > \Re(C_j \mid \hat{\mathbf{y}}), \forall j = 1,\ldots,N, j \neq i \mid C_i\} \tag{23}$$

Since the proposed soft clustering algorithm exploits the constellation shape as its discriminating feature, increasing the burst length above a threshold $K > K_0$ does not increase P_{CC}. This is confirmed by simulations.

We study $\Re(C_i \mid \hat{\mathbf{y}})$ for high SNR and fixed *K*, which is a possible and important case in our problem. From (21), we note that the membership values $\{u_{\ell k}^i\}$ play an important role in the analysis of $\Re(C_i \mid \hat{\mathbf{y}})$.

We now show that at high enough SNR ($SNR \to \infty$), $\Re(C_i \mid \hat{\mathbf{y}}, C_i) \to 1$. Without loss of generality, for $\theta = 0$ and high SNR e.g., $SNR \to \infty$, the received symbol after applying the equalizer, i.e., $\hat{y}_k$, gets closer to one of the symbols $S_{\ell'_k}^i$ of the transmitted constellation C_i (e.g., Fig. 9b). Therefore, $d_{\ell'_k k}^i = \left|\hat{y}_k - S_{\ell'_k}^i\right| \to 0$. Considering Eq. (19), we then conclude by simple mathematical manipulations that

$$u_{\ell k}^i = \begin{cases} \to 0, & (\ell \neq \ell'_k) \\ \to 1, & (\ell = \ell'_k) \end{cases} \tag{24}$$

In other words, at high SNRs, the membership vector $\mathbf{u}_k^i = [u_{\ell k}]_{M_i \times 1}$ tends to a fuzzy singleton vector, i.e., $\delta(\ell - \ell'_k)$. Therefore, using Eq. (21) yields $\Re(C_i \mid \hat{\mathbf{y}}, C_i) \to 1$.

In what follows, we also show that in high SNRs, the proposed soft-clustering technique and the generalized maximum likelihood classifier are equivalent. Under such assumptions, we write the soft-clustering cost function (16) as

$$\Omega_i = \sum_{k=1}^{K} \left|\hat{y}_k - S_{\ell'_k}^i e^{j\theta}\right|^2 \tag{25}$$

Similarly, the distance between the equalized symbol $\hat{y}_k$ and the rotated symbol $S_{\ell'_k} e^{j\hat{\theta}}$ (i.e., $\left|\hat{y}_k - S_{\ell'_k} e^{j\hat{\theta}}\right|$) is minimized

$$\left|\hat{y}_k - S^i_{\ell'_k} e^{j\hat{\theta}}\right| \le \left|\hat{y}_k - S^i_{\ell} e^{j\hat{\theta}}\right| \quad \forall \ell \in [1, M_i] \tag{26}$$

Then, we have

$$\Omega_i = \sum_{k=1}^{K} \min_{\ell \in [0, M_i]} \left|\hat{y}_k - S^i_{\ell} e^{j\theta}\right|^2 \tag{27}$$

which is the same as that of the generalized likelihood ratio classifier (GLR) developed in [12, 7]. This implies that our soft clustering method in high SNR approaches the performance of the optimum classifier in [7].

6. SIMULATION RESULTS

We use simulations to observe the performance of our proposed method in additive white Gaussian noise and fading channels. The probability of correct classification P_{CC} is utilized to compare our proposed approach with the cumulant based methods in [3]. This was done by 1000 Monte Carlo simulations for each modulation in the set $\mathcal{A}$. We consider the following two examples. In Example 1, we consider an AWGN channel and short burst transmissions, and in Example 2, we consider a fading channel.

Example 1- *AWGN Channel:* We assume the channel is completely equalized by the blind equalizer for the received burst of K=256 symbols, and consider the performance of our soft clustering algorithm for classifying the modulation type from $\mathcal{A}1$= {8 PSK, 8 QAM}, $\mathcal{A}2$={V.29c, 16 QAM}, $\mathcal{A}3$={8 QAM, 8 PSK, 16QAM, V.29c}, and $\mathcal{A}4$= {128QAM, 256QAM}. Simulation results are shown in Figs. 10-12. Fig. 10 shows the performances of the soft-clustering method for the two-class problems $\mathcal{A}1$ and $\mathcal{A}2$ and for the four class problem $\mathcal{A}3$. Figures 11-12 compare the performances of soft-clustering and cumulant-based methods for the two class problems $\mathcal{A}2$ and $\mathcal{A}4$, respectively. We note that for short burst K=256, the value of P_{CC} for the soft-clustering method approaches unity (accurate classification) when SNR≥10dB, whereas P_{CC} for cumulant based method in [3] is much less.

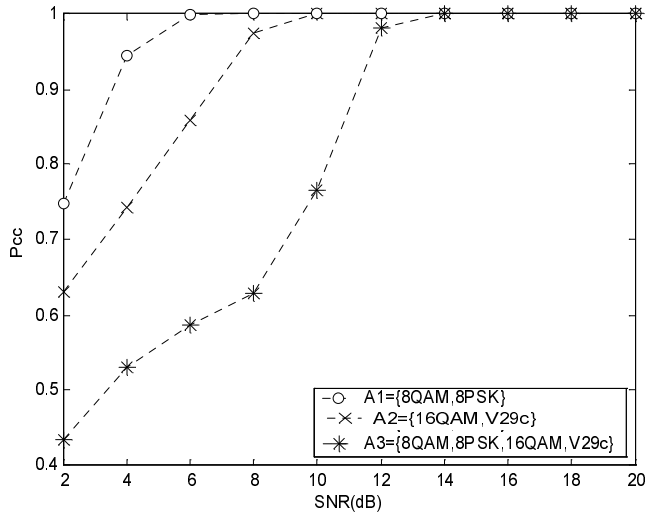


Fig 10. Probability of correct classification P_{CC} of soft-clustering classifier versus SNR for $\mathcal{A}1$, $\mathcal{A}2$, and $\mathcal{A}3$ modulation sets in AWGN and K=256

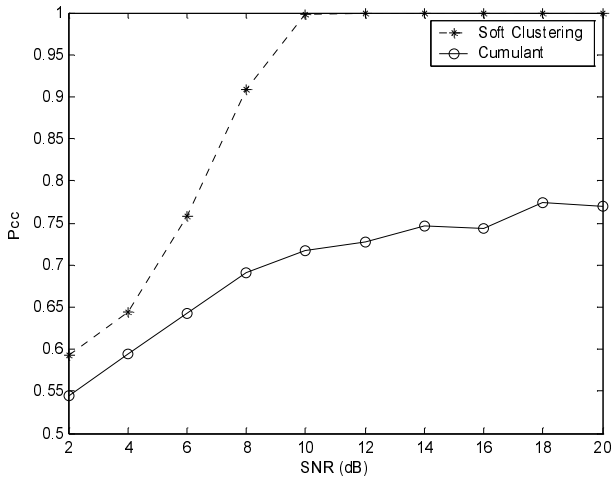


Fig 11. Probability of correct classification P_{CC} of soft-clustering and cumulant-based classifiers versus SNR for $A2$={16QAM,V29c} modulation set in AWGN and K=256

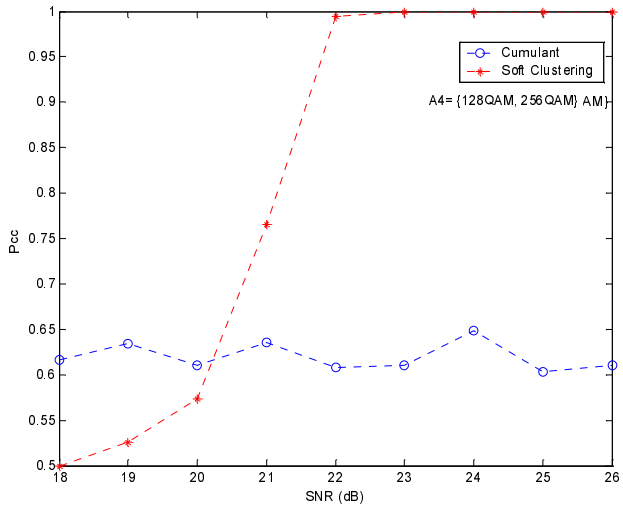


Fig 12. Probability of correct classification P_{CC} of soft-clustering and cumulant-based classifiers versus SNR for $A4$={128QAM,256QAM} modulation set in AWGN and K=256

Note that as Fig. 11 shows, our proposed classifier has a significantly higher P_{CC} compared to the cumulant-based classifier for short bursts, particularly for higher values of SNR.

Example 2- *Fading Channel*: We evaluate the performance of our method for $A5$={16 QAM, V.29c, 32 QAM} in multipath fading channels. The fading channel is assumed to be a slow fading wireless mobile channel with $f_D T = 5\times10^{-4}$. We use a method proposed in [39] to simulate the discrete time model of the fading channel. We also use the six-path channel model described in [27] for the typical urban (TU) wireless system. The channel intensity profile is shown in Table 2 and the sampling time T_S is T/4.

Table 2. TU Six-Path Channel Model (T=3.7 μsec) [27]

Path delay μsec	Fractional power	Path delay kT_S	Fractional power
0.0	0.189	0	0.189
0.2	0.379	1	0.618
0.5	0.239	2	0.095
1.6	0.095	3	0.061
2.3	0.061	4	0.000
5.0	0.037	5	0.037

We compare the performances of the following three different classifiers in the above mentioned fading channel using 1000 Monte Carlo trials for each modulation type, i.e., a total of 3000 trials for the three class problem $\mathcal{A}5$ for each value of SNR.

Classifier 1: CMA + cumulant-based modulation classifier as a benchmark (shown in Fig. 1 in which the CMA is used as the equalizer, and the cumulant-based method [3] is used as its classification algorithm).
Classifier 2: CMA + soft clustering algorithm (SCA) modulation classifier (shown in Fig. 7a in which soft-clustering is used for classification).
Classifier 3: Iterative DAMA + soft clustering algorithm modulation classifier (with the iterative structure shown in Figs. 7a-b and soft-clustering used as its classification algorithm).

As we are focusing on fading channels, short bursts are inadequate for modulation classification. In order to alleviate this, each Monte Carlo trial uses K=800 independent symbols in each burst (as compared to 256 independent symbols used in Example 1). As per assumption AS7, the transmitter switches to higher order modulations when the channel is relatively stable during a burst. Thus, it is reasonable to assume that for $\mathcal{A}5$={16 QAM, V.29c, 32 QAM}, bursts of 800 symbols length are available. Each burst was reused 5 times in forward-reverse to generate a long burst. We observed that longer bursts would not have a significant impact on the classifier's performance. The results are shown in Fig. 13.

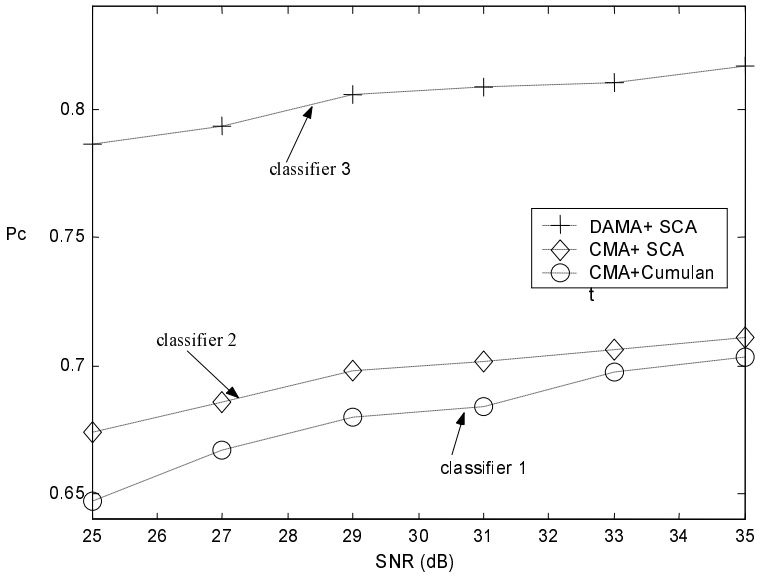


Fig. 13. Probability of classification *Pcc* of three classifiers of Example 2, for three class problem $\mathcal{A}5$ ={16QAM, V.29c, 32QAM}in a slow fading channel

The proposed iterative method (classifier 3) outperforms one-pass CMA+Cumulant (classifier 1) and CMA+SCA (classifier 2). The iterative classifier has a significant gain in SNR as compared to classifiers 1 and 2. The SNR gain is achieved at a higher computational cost. The complexity of the proposed method in this paper is $O(K\sum_{i=1}^{N} M_i)$ compared to that of the cumulant-based method which is $O(K)$. In a similar setting, the cumulant-based classifier in [3] requires significantly more symbols to achieve the same performance as our proposed iterative classifier. Although the performances of both classifiers 1 and 2 are not satisfactory, the CMA+Cumulant-based classifier (Classifier 1) requires 7-8 dB higher SNR to match the performance of the CMA+SCA (classifier 2).

7. CONCLUSION

We proposed a novel scheme for identifying the modulation types of bursty QAM signals. This structure is based on iterative adjusted-modulus blind equalization and soft clustering. As the number of available symbols in the received burst may not be sufficient for equalizer training, we repeatedly use the available symbols in a forward-reverse manner. Moreover, for fast processing, we employ parallel structures. Our soft clustering algorithm does not require a precise statistical model of the input signal and noise parameters, as compared to the decision theoretic approach in [1].

In order to show the effectiveness of our approach, we compared our results with those of the cumulant method in [3]. For a multipath fading channel, our algorithm requires significantly fewer symbols to achieve the same classification as compared to the use of cumulants. This is particularly important in bursty transmissions where high number of symbols are not available.

REFERENCES

1. Wei, W. & Mendel, J. M. (2000). Maximum-likelihood classification for digital amplitude-phase modulations. *IEEE Trans. Comm*., Vol. 48, No. 2, pp. 189-193.
2. Webb, W. T. & Hanzo, L. (1994). *Modern Quadrature Amplitude Modulation: Principles and Applications for Fixed and Wireless Channels.* IEEE Press.
3. Swami, A. & Sadler, B. M. (2000). Hierarchical digital modulation classification using cumulants. *IEEE Trans. Comm.*, Vol. 48, No. 3, pp. 416-429.
4. Azzouz, E. E. & Nandi, A. K. (1996). *Automatic modulation recognition of communication signals*. Springer.
5. Hong, L. (2002). *Advanced techniques for automatic classification of digitally modulated communication signals*. PhD Dissertation, University of Missouri, Columbia.
6. Attar, A. R. & Sheikhi, A., Abiri, H. & Mallahzadeh, A. (2006). A new method for modulation system recognition. *Iranian Journal of Science & Technology, Transaction B: Engineering,* Vol. 30, No. B6, pp 775-788.
7. Dobre, O. A., Abdi, A., Bar-Ness, Y. & Su, W. (2007). Survey of automatic modulation classification techniques: Classical approaches and new trends. *IET Commun.,* Vol. 1, No. 2, pp. 137-156.
8. Polydoros, A. & Kim, K. (1990). On the detection and classification of quadrature digital modulation in broad-band noise. *IEEE Trans. Comm.*, Vol. 38, pp. 1199-1211.
9. Huang, C. Y. & Polydoros, A. (1995). Likelihood methods for MPSK modulation classification. *IEEE Trans. Comm.*, Vol. 43, No. 2/3/4, pp.1493-1504.
10. Long, C. S. (1996). *On likelihood digital modulation classification.* PhD Dissertation, University of Southern California, Los Angeles, CA.
11. Lin, Y. C. & Kuo, C. C. J. (1997). Classification of quadrature amplitude modulation (qam) signals via sequential probability ratio test (sqrt). *Signal Processing*, Vol. 60, No. 3, pp. pp. 263-280.
12. Wei, W. (1998). Classification of digital modulations using constellation analysis. PhD Dissertation, University of Southern California, Los Angeles, CA.
13. Paris, B. P. *et al.* (1997). Modulation classification in unknown dispersive environnants. *Proc. IEEE ICASSP'97*, Vol. 5, pp. 3853-3856, Munich, Germany.
14. Benvenuto, N. & Goeddel, T. W. Classification of voice band data signals using constellation magnitude. *IEEE Trans. Comm.*, Vol. 43, No. 11, pp. 2759-2770.
15. Delgosha, F. & Nayebi, M. M. (1998). Cyclostationarity based classification of digital quadrature and staggered modulation. *Proc. of the International Conference on Telecommunications, ICT'98,* Vol. 2, pp. 180-184, Porto Carras, Greece.

16. Schreyogg, C. & Reichert, J. (1997). Modulation classification of QAM Schemes using the DFT of phase histogram combined with modulus information. *Proc. IEEE MILCOM'97*, Vol. 3, pp. 1372-1376, Monterey, California.
17. Wood, S. L., Ready, M. J. & Treichler, J. R. (1988). Constellation identification using radon transform. *Proc. IEEE ICASSP'98*, Vol. 3, pp. 1878-1881, New York.
18. Wood, S. L. & Treichler, J. R. (1994). Computational and performance analysis of radon transform based constellation identification. *Proc. IEEE ICASSP'94*, Vol. 3, pp. 241-244, Adelaide, Australia.
19. Zhu, Q., Kam, M. & Yeager, R. (1992). Non-parametric identification of QAM constellations in noise. *Proc. IEEE ICASSP'92*, Vol. 4, pp. 184-187, San Francisco, California.
20. Mobasseri, B. G. (1999). Constellation shape as a robust signature for modulation recognition. *Proc. IEEE MILCOM '99*, Vol. 1, pp.442-446, Atlantic City, New Jersey.
21. Wesel, R. D. (1989). *Adaptive equalization for modem constellation identification*, MSc Thesis, MIT.
22. Duda, R. O. & Hart, P. E. (1973). *Pattern classification and scene analysis*. John Wiley, New York.
23. Swami, A. & Sadler, B. M. (1997). Modulation classification via hierarchical agglomerative cluster analysis. *Proc. IEEE SPWC*, pp. 141-144, Paris, France.
24. Schreyogg, C. (1995). Identification of voice band data signal constellations using divisive cluster algorithm. *Proc. IEEE Digital Signal Processing Workshop*, pp. 474-477, Loen, Norway.
25. Shahmohammadi, M. & Nikoofar, H. R. (2002). Modulation classification for QAM/PSK using a soft-clustering algorithm. *Proc. IEEE ISIT'02*, page 19, EPFL, Lausanne, Switzerland.
26. Proakis, J. (1995). *Digital communications*. 3rd edition, McGraw-Hill.
27. Yiin, L. & Stuber, G. L. (1997). MLSE and soft-output equalization for Trellis coded continuous phase modulation. *IEEE Trans. Comm*., Vol. 45, No. 6, pp. 651-657.
28. Ghogho, M., Swami, A. & Durrani, T. (2000). Blind estimation of frequency offset in the presence of unknown multipath. *Proc. of IEEE International conference on Personal Wireless Communication, ICPWC'2000*, pp.104-108, Heyderabad, India.
29. Wang, Y., Ciblat, P., Serpedin, E. & Loubaton, P. (2002). Performance analysis of a class of non-data aided carrier frequency offset and symbol timing delay estimators for flat-fading channels. *IEEE Trans. Sig. Proc.*, Vol. 50, No. 9.
30. Ciblat, P., Loubaton, P., Serpedin, E. & Giannakis, G. B. (2002). Asymptotic analysis of blind cyclic correlation based symbol rate estimators. *IEEE Trans. Information Theory*, Vol. 48, No. 5.
31. Bello, P. A. (1963). Characterization of randomly time-variant linear channels. *IEEE Trans. Comm.*, Vol. 11, pp. 360-393.
32. Webb, W. & Steele, R. (1995). Variable rate QAM for mobile radio. *IEEE Trans. Comm.*, Vol. 43, pp. 2223–2230.
33. Johnson, C. R. *et al.*, (1998). Blind equalization using the constant modulus criterion. *IEEE Proc.*, Vol. 86, No. 10, pp. 1927-1950.
34. Yang, J. (1997). *Multimodulus algorithms for blind equalization*, Ph.D. Dissertation, University of British Columbia.
35. Axford, R. A., Milstein, L. B. & Zeidler, J. R. (1996). A dual-mode algorithm for blind equalization of QAM signals: CADAMA. *Proc. of 29th Asilomar Conf. Signals, Systems and Computers*, Vol. 1, pp. 172 -176.
36. Kim, B. J. (1998). *Blind equalization for short burst wireless communications*. PhD Dissertation, Stanford University, Stanford, California.
37. Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum.
38. Bezdek, J. C. (1973). Cluster validity for the fuzzy sets. *J. Cybernet.,* Vol. 3, pp. 58-73.
39. Jakes, W. C. (1974). Microwave mobile communications. Wiley.