

A Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends

Octavia A. Dobre¹, Ali Abdi², Yeheskel Bar-Ness² and Wei Su³

¹Faculty of Engineering and Applied
Science,
Memorial University of Newfoundland
St. John's, NL A1B 3X5, Canada
dobre@engr.mun.ca

²Dept. of Electrical and Computer
Engineering,
New Jersey Institute of Technology,
Newark, NJ 07102, US
abdi@adm.njit.edu, barness@yegal.njit.edu

³ US Army RDECOM
Fort Monmouth, NJ, 07703, US
wei.su@mail1.monmouth.army.mil

*Abstract*__ The automatic recognition of the modulation format of a detected signal, the intermediate step between signal detection and demodulation, is a major task of an intelligent receiver, with various civilian and military applications. Obviously, with no knowledge of the transmitted data and many unknown parameters at the receiver, such as the signal power, carrier frequency and phase offsets, timing information, etc., blind identification of the modulation is a difficult task. This becomes even more challenging in real-world scenarios with multipath fading, frequency-selective and time-varying channels. In this paper we provide a comprehensive survey of different modulation recognition techniques, in a systematic way. A unified notation is used to bring in together, under the same umbrella, the vast amount of results and classifiers, developed for different modulations. The two general classes of automatic modulation identification algorithms are discussed in detail, which rely on the likelihood function and features of the received signal, respectively. The contributions of numerous articles are summarized in compact forms. This helps the reader to see the main characteristics of each technique. However, in many cases, the reported results in the literature have been obtained under different conditions. So, we have also simulated some major techniques under the same conditions, which allows a fair comparison among different methodologies. Furthermore, new problems that have appeared as a result of emerging wireless technologies are outlined. At the end, open problems and possible directions for future research are briefly discussed.

Keywords: Automatic modulation classification, Automatic modulation recognition, Classifier performance, Likelihood function, Maximum likelihood, Feature extraction, Preprocessing tasks, Model mismatches, Probability of correct classification.

Part of this work was published in a preliminary form and presented at the Sarnoff Symposium, Princeton, NJ, USA, 18-19 April 2005, under the title “*Blind modulation classification: a concept whose time has come*”.

1. INTRODUCTION

Automatic modulation classification (AMC) is an intermediate step between signal detection and demodulation, and plays a key role in various civilian and military applications. Implementation of advanced information services and systems for military applications, in a crowded electromagnetic spectrum, is a challenging task for communication engineers. Friendly signals should be securely transmitted and received, whereas hostile signals must be located, identified and jammed. The spectrum of these signals may range from high frequency (HF) to millimeter frequency band, and their format can vary from simple narrowband modulations to wideband schemes. Under such conditions, advanced techniques are required for real-time signal interception and processing, which are vital for decisions involving electronic warfare operations and other tactical actions. Furthermore, blind recognition of the modulation format of the received signal is an important problem in commercial systems, especially in software defined radio (SDR), which copes with the variety of communication systems. Usually, supplementary information is transmitted to reconfigure the SDR system. Blind techniques can be used with an intelligent receiver, yielding an increase in the transmission efficiency by reducing the overhead. Such applications have emerged the need for flexible intelligent communication systems, where the automatic recognition of the modulation of a detected signal is a major task. A simplified block diagram of the system model is shown in Fig. 1. The design of a modulation classifier essentially involves two steps: signal preprocessing and proper selection of the classification algorithm. Preprocessing tasks may include, but not limited to perform some or all of, noise reduction, estimation of carrier frequency, symbol period, and signal power, equalization, etc. Depending on the classification algorithm chosen in the second step, preprocessing tasks with different levels of accuracy are required; some classification methods require precise estimates, whereas others are less sensitive to the unknown parameters.

Regarding the second step, two general classes of AMC algorithms can be crystallized, likelihood-based (LB) [1]-[26] and feature-based (FB) [27]-[88] methods, respectively. The former is based on the likelihood function of the received signal and the decision is made comparing the likelihood ratio against a threshold. A solution offered by the LB algorithms is optimal in the Bayesian sense, viz., it minimizes the probability of false classification. The optimal solution suffers from computational complexity, which in many cases of interest naturally gives rise to suboptimal classifiers. In the FB approach, on the other hand, several features are usually employed and a decision is made based on their observed values. These features are normally chosen in an ad-hoc way. Although a FB-based method may not be optimal, it is usually simple to implement, with near-optimal performance, when designed properly. Once the modulation format is correctly identified, other operations, such as signal demodulation and information extraction, can be subsequently performed. In general, AMC is a challenging task, especially in a non-cooperative environment, where in addition to multipath propagation, frequency-selectivity and time-varying nature of the channel, no prior knowledge of the incoming signal is available.

In recent years, new technologies for wireless communications have emerged. The wireless industry has shown great interest in orthogonal frequency division multiplexing (OFDM) systems, due to the efficiency of OFDM schemes to transmit information in frequency selective fading channels, without complex equalizers

[89]-[90]. Multiple-input multiple-output (MIMO) systems have also received considerable attention, due to the significant capacity increase they offer. Such emerging technologies in wireless communications have raised new challenges for the designers of signal intelligence and SDR systems, such as, discriminating between OFDM and single carrier modulations [91], identification of signals transmitted from multiple antenna systems, and so on.

Research on automatic classification of both digital and analog modulations has been carried out for at least two decades [1]-[88]. Partial surveys of algorithms for identifying digitally modulated signals are given in [92] and [93]. Of course, many techniques have been developed, which are different from each other when it comes to details. However, general structures that connect a variety of apparently different techniques can be identified. In this paper, we provide a unified comprehensive overview of what has been accomplished so far in this area, highlighting the bottlenecks and challenging issues which need to be addressed by further research. A comparison among the performance of different LB and FB algorithms is also carried out, emphasizing the advantages and disadvantages of diverse techniques.

The rest of the paper is organized as follows. In Section II the signal model and classifier performance measures are discussed. Section III and IV are devoted to LB and FB methods, respectively. Numerical performance assessments and comparisons are provided in Section V, and some concluding remarks are given in Section VI.

2. SIGNAL MODEL AND PERFORMANCE MEASURES

AMC algorithms proposed in the literature employ information extracted from either the received baseband waveform [1]-[9], [14]-[28], [41]-[54], [58]-[65] or intermediate frequency [32]-[38], [55], [88]. A general expression for the baseband received complex envelope is given by

$$r(t) = s(t; \mathbf{u}_i) + n(t), \quad (1)$$

where

$$s(t; \mathbf{u}_i) = a_i e^{j2\pi\Delta f t} e^{j\theta} \sum_{k=1}^K e^{j\phi_k} s_k^{(i)} g(t - (k-1)T - \varepsilon T), \quad 0 \leq t \leq KT \quad (2)$$

is the noise-free baseband complex envelope of the received signal. In (2) we have $j^2 = -1$, $a_i = \sqrt{E_s / \sigma_{s^{(i)}}^2 E_p}$, with E_s as the baseband signal energy, $\sigma_{s^{(i)}}^2 = M_i^{-1} \sum_{m=1}^{M_i} |s_m^{(i)}|^2$ as the variance of the i th zero-mean signal constellation, M_i as the number of equi-probable points in the i th signal constellation, $E_p = \int_{-\infty}^{\infty} |p_{TX}(t)|^2 dt$ as the pulse energy, with $p_{TX}(t)$ as the transmitter pulse shape, Δf is the carrier frequency offset, θ is the time-invariant carrier phase, $\{\phi_k\}_{k=1}^K$ represent the phase jitter, $\{s_k^{(i)}\}_{k=1}^K$ are K complex transmitted data symbols taken from the i th finite-alphabet modulation format, T is the symbol period, ε denotes the timing offset with respect to (w.r.t.) the receiver reference clock, such that $0 \leq \varepsilon < 1$, and $g(t) = p_{TX}(t) \otimes h(t)$, with $h(t)$ as the channel impulse response and \otimes as the convolution. For example, for a slowly-varying fading channel with P independent paths, $h(t) = \sum_{p=1}^P \alpha_p e^{j\varphi_p} \delta(t - \tau_p)$, with $\delta(\cdot)$ as the Dirac delta function, such that $\{\alpha_p\}_{p=1}^P$, $\{\varphi_p\}_{p=1}^P$

and $\{\tau_p\}_{p=1}^P$ are the amplitudes, phases and delays of the paths, respectively. We adopt the notation $s(t; \mathbf{u}_i)$ to stress the signal dependence on the unknown quantities, i.e., $\mathbf{u}_i = [a_i \ \Delta f \ \theta \ T \ \varepsilon \ g(t) \ \{\phi_k\}_{k=1}^K \ \{s_k^{(i)}\}_{k=1}^K]^\dagger$, with \dagger as the transpose. The term ‘‘unknown quantities’’ refers to unknown parameters, such as the carrier frequency offset, as well as unknown data symbols. Without loss of generality, unit variance constellations will be considered in the sequel, obtained by normalizing the signal constellations. For example, for M -ary amplitude-shift keying (ASK) the symbols are given by $s_k^{(M\text{-ASK})} = s_{k,I}^{(M\text{-ASK})}$, $s_{k,I}^{(M\text{-ASK})} \in \{(2m-1-M)/\sigma_{s^{(M\text{-ASK})}}\}$, $m=1, \dots, M\}$, $k=1, \dots, K$, for rectangular M -ary quadrature amplitude modulation (QAM) $s_k^{(M\text{-QAM})} = s_{k,I}^{(M\text{-QAM})} + js_{k,Q}^{(M\text{-QAM})}$, $s_{k,I}^{(M\text{-QAM})}, s_{k,Q}^{(M\text{-QAM})} \in \{(2m-1-M^{1/2})/\sigma_{s^{(M\text{-QAM})}}\}$, $m=1, \dots, M^{1/2}\}$, $k=1, \dots, K$, for M -ary phase-shift-keying (PSK) $s_k^{(M\text{-PSK})} = e^{j\theta_m}$, $\theta_m \in \{2\pi m/M, m=0, \dots, M-1\}$, $k=1, \dots, K$ and for M -ary frequency-shift-keying (FSK) $s_k^{(M\text{-FSK})} = e^{j2\pi f_m t}$, $f_m \in \{(2m-1-M)f_d, m=1, \dots, M\}$, where the subscripts I and Q represent the inphase (real) and quadrature (imaginary) parts, respectively, $\sigma_{s^{(i)}}$ is the variance of the constellation before normalization, M is a power of 2, and f_d is the frequency deviation or spacing between any two adjacent FSK constellation points, which for orthogonal signaling is a multiple of $1/2T$ (see, for example, [94] Ch. 4). Note that the data symbols depend on t for FSK. For others, such as ASK, PSK, QAM, $s_k^{(i)}$ is constant for each period $(k-1)T$ to kT . To simplify the notation, subsequently we use $s_k^{(i)}$ without t also for FSK signals, unless otherwise mentioned. Note that $n(t)$ in (1) is the aggregate baseband complex noise, i.e., receiver noise, as well as cochannel interferences and jammers.

A classifier is supposed to correctly choose the modulation format of the incoming signal from a pool of N_{mod} candidate modulations, denoted by the integers $i=1, \dots, N_{\text{mod}}$, or to decide that the modulation format cannot be recognized. The latter case is not discussed here, as not addressed in the literature. Due to the lack of space, we focus on algorithms for ASK, PSK, QAM, and FSK classification.

As a basic performance measure, let $P_c^{(i|i)}$ denote the (classification) probability to declare that the i' th signal format has been sent, when the modulation format of the incoming signal is i . For $i, i'=1, \dots, N_{\text{mod}}$, these probabilities can be arranged as a $N_{\text{mod}} \times N_{\text{mod}}$ confusion matrix, where the diagonal element $P_c^{(i|i)}$ is the probability of correct classification for the i th modulation. In classifying N_{mod} equi-probable modulations, the average probability of correct classification is defined by

$$P_{cc} = N_{\text{mod}}^{-1} \sum_{i=1}^{N_{\text{mod}}} P_c^{(i|i)}. \quad (3)$$

Obviously, one can use the complementary probabilities as a performance measure, i.e., the probability of error for the i th modulation, defined as $P_e^{(i)} = 1 - P_c^{(i|i)}$, and the average probability of error, defined as $P_e = 1 - P_{cc}$. Most of the AMC work used P_{cc} , or equivalently, P_e , as a performance measure. However, by using the confusion matrix, one gains more insight into the classifier behavior.

Clearly, a desirable classifier should provide a high probability of correct classification in a short observation interval, particularly for a large range of signal-to-noise ratio (SNR). In addition, it should satisfy these requirements: capability to recognize many different modulations in environments with diverse propagation characteristics, robustness to model mismatches, real-time functionality, and low computational complexity.

3. LIKELIHOOD-BASED APPROACH TO AMC

Within the LB framework, AMC is a multiple composite hypothesis-testing problem. The idea behind the LB-AMC is that the probability density function (PDF) of the observed waveform, conditioned on the embedded modulated signal, contains all information for classification. Depending on the model chosen for the unknown quantities, three LB-AMC techniques are proposed in the literature: average likelihood ratio test (ALRT) [1]-[13], [21], [23], generalized likelihood ratio test (GLRT) [14], [17], [18] and hybrid likelihood ratio test (HLRT) [14]-[16], [19]-[20]. Quasi ALRT [3]-[5], [7], [9]-[13] and quasi HLRT [20]-[22] are also proposed in the literature.

ALRT

This approach treats the unknown quantities as random variables (r.v.'s) with certain PDFs. So, the likelihood function (LF) under the hypothesis H_i , representative of the i th modulation, $i = 1, \dots, N_{\text{mod}}$, is given by

$$\Lambda_A^{(i)}[r(t)] = \int \Lambda[r(t) | \mathbf{v}_i, H_i] p(\mathbf{v}_i | H_i) d\mathbf{v}_i, \quad (4)$$

where $\Lambda[r(t) | \mathbf{v}_i, H_i]$ is the conditional LF of the noisy received signal $r(t)$ under H_i , conditioned on the unknown vector \mathbf{v}_i , and $p(\mathbf{v}_i | H_i)$ is the a priori PDF of \mathbf{v}_i under H_i . The known PDF of \mathbf{v}_i enabled us to reduce the problem to a simple hypothesis-testing problem by integrating over \mathbf{v}_i .

For a baseband complex additive white Gaussian noise (AWGN) in (1), the conditional LF is given by (see, for example, [94] Ch. 6)

$$\Lambda[r(t) | \mathbf{v}_i, H_i] = \exp \left[2N_0^{-1} \operatorname{Re} \left\{ \int_0^{KT} r(t) s^*(t; \mathbf{u}_i) dt \right\} - N_0^{-1} \int_0^{KT} |s(t; \mathbf{u}_i)|^2 dt \right], \quad (5)$$

where N_0 is the two-sided power spectral density (PSD) of AWGN in W/Hz, with the autocorrelation $E\{n(t)n^*(t+\tau)\} = N_0\delta(\tau)$ such that $E\{\cdot\}$ is expectation and $*$ denotes the complex conjugate. Furthermore, here $\mathbf{v}_i = [\mathbf{u}_i^\dagger N_0]^\dagger$, and $\operatorname{Re}\{\cdot\}$ stands for the real part. If the chosen $p(\mathbf{v}_i | H_i)$ is the same as the true PDF, ALRT results in an optimal classifier in the Bayesian sense.

GLRT

In this approach the unknown parameters are treated as unknown deterministic. The best performance is achieved by the so-called uniformly most powerful (UMP) test [95]. For a necessary and sufficient condition for the existence of an UMP test see, for example, [95] Ch. 2. When an UMP test does not exist or hard to derive, a logical procedure is to estimate the unknown quantities, assuming H_i is true, and then use these estimates in a likelihood ratio test, as if they were correct. If maximum likelihood (ML) is used for estimates, the test is called

GLRT. Obviously, GLRT treats the unknown quantities (including both the parameters and data symbols) as deterministic unknowns, and the LF under H_i is given by

$$\Lambda_G^{(i)}[r(t)] = \max_{\mathbf{v}_i} \Lambda[r(t) | \mathbf{v}_i, H_i]. \quad (6)$$

HLRT

This is a combination of the aforementioned approaches, for which the LF under H_i is given by

$$\Lambda_H^{(i)}[r(t)] = \max_{\mathbf{v}_i} \int \Lambda[r(t) | \mathbf{v}_i, \mathbf{v}_{i_2}, H_i] p(\mathbf{v}_{i_2} | H_i) d\mathbf{v}_{i_2}, \quad (7)$$

where $\mathbf{v}_i = [\mathbf{v}_{i_1}^\dagger \ \mathbf{v}_{i_2}^\dagger]^\dagger$ and, \mathbf{v}_{i_1} and \mathbf{v}_{i_2} are vectors of unknown quantities modeled as unknown deterministic and r.v.'s, respectively. Usually, \mathbf{v}_{i_1} and \mathbf{v}_{i_2} consist of parameters and data symbols, respectively.

Note that ALRT requires a multidimensional integration, whereas GLRT requires a multidimensional maximization. The difficulty of performing a multidimensional integration for a large number of unknown quantities and the need for knowing the prior PDFs may render the ALRT impractical. On the other hand, maximization over the unknown data symbols in GLRT can lead to the same value of the LF for nested signal constellations, e.g., BPSK and QPSK, 16-QAM and 64-QAM, [14], [96] Ch. 6, which in turn yields incorrect classification. Averaging over the unknown data symbols in HLRT, however, removes the nested constellations problem of GLRT. Finally we emphasize that the estimates of the unknown quantities, as by-products of GLRT and HLRT, are of interest for data demodulation.

In a two-hypothesis classification problem, the decision is made according to

$$\Lambda_l^{(1)}[r(t)] / \Lambda_l^{(2)}[r(t)] \underset{H_2}{\overset{H_1}{>}} \eta_l, \quad l = A \text{ (ALRT)}, G \text{ (GLRT)}, H \text{ (HLRT)}, \quad (8)$$

where η_l is a threshold. The left-hand side is referred to as the likelihood ratio and the test is called average likelihood ratio test (ALRT), generalized likelihood ratio test (GLRT) and hybrid likelihood ratio test (HLRT), respectively, depending on the method employed to compute the LF. Extension of (8) to multiple classes is straightforward (see, for example, [95] Ch.2 and [96] Ch. 3 and 6). Equivalently, the log function can be applied to both members of the inequality (8). Accordingly, the terms log-likelihood ratio and log-likelihood ratio test are used.

Table I lists several LB-AMC algorithms proposed in the literature, emphasizing the type of modulations, unknown parameters, and channel used. A simplified signal model of that given in (1) was considered in the literature, as follows. The transmit pulse shape was assumed rectangular, i.e., $p_{TX}(t) = u_T(t)$, where $u_T(t) = 1$ for $0 \leq t < T$ and zero otherwise. So, $E_p = T$ and with $\sigma_{s^{(i)}}^2 = 1$, one obtains $a = \sqrt{S}$, in which $S = E_s / T$ is the signal power. With all this and $h(t) = \alpha e^{j\phi} \delta(t)$, with α and ϕ constant over the K symbol interval, (2) can be written as

$$s(t; \mathbf{u}_i) = \alpha \sqrt{S} e^{j(\theta + \phi)} e^{j2\pi \Delta f t} \sum_{k=1}^K e^{j\phi_k} s_k^{(i)} u_T(t - (k-1)T - \varepsilon T), \quad (9)$$

where $\alpha = 1$ and $\varphi = 0$ when there is no fading. We define the observed per-symbol SNR $\gamma_s = \Omega ST / N_0$, in which $\Omega = 1$ when there is no fading. Otherwise, $\Omega = E\{\alpha^2\}$ is the average fading power, with α as the channel amplitude. Under fading, S is set equal to one.

3.1. ALRT-based Algorithms

In this section, optimal and suboptimal ALRT-based algorithms applicable to identify both linearly modulated and FSK signals under various conditions will be presented¹, as well as ALRT-based classifiers specific to linear modulation and FSK, respectively. Suboptimal classifiers are obtained based on the approximations of the LFs at low SNR. Interestingly, several FB classifiers are shown to be simplified versions of such suboptimal structures [5]. Hence, decision theory can be perceived as a rigorous framework that justifies the selection of features in some FB methods.

3.1.1. ALRT-AMC for linearly modulated and FSK signals

ALRT-based classifiers

With all parameters perfectly known², i.e., $\mathbf{v}_i = [\{s_k^{(i)}\}_{k=1}^K]^\dagger$, ALRT leads to a structure whose performance is better than all the others, which have to deal with some unknown parameters. Therefore, the performance of this classifier can be considered as a benchmark. The data symbols $\{s_k^{(i)}\}_{k=1}^K$ are treated as independent and identically distributed (i.i.d.) r.v.'s. The LF under hypothesis H_i is computed by averaging over the constellation points corresponding to the i th modulation format. This is done by substituting (5) into (4) (see (37), Appendix A)

$$\Lambda_A^{(i)}[r(t)] = \prod_{k=1}^K E_{s_k^{(i)}} \left\{ \exp \left[2\alpha \sqrt{S} N_0^{-1} \operatorname{Re} \left\{ R_k^{(i)} - \alpha^2 S T N_0^{-1} |s_k^{(i)}|^2 \right\} \right] \right\}, \quad (10)$$

where $E_{s_k^{(i)}}\{\cdot\}$ is nothing but a finite summation over all the M_i possible constellation points of the i th modulation, divided by M_i , for the k th interval. Furthermore,

$$R_k^{(i)} = \int_{(k-1)T}^{kT} r(t) s_k^{(i)*}(t) u_T(t - (k-1)T) dt = \int_{(k-1)T}^{kT} r(t) s_k^{(i)*}(t) dt, \quad k=1, \dots, K. \quad (11)$$

Note that for linear modulations $s_k^{(i)}(t)$ is constant over the period $(k-1)T$ to kT and thus, $R_k^{(i)} = s_k^{(i)*} r_k$, with $r_k = \int_{(k-1)T}^{kT} r(t) dt$ the output of the receive matched filter at $t = kT$.

With multiple antennas at the receiver, AWGN and block fading, and all parameters perfectly known, i.e., $\mathbf{v}_i = [\{s_k^{(i)}\}_{k=1}^K]^\dagger$, the LF is given by (see, e.g., [95] Ch.3)

¹ Due to the lack of space, in the sequel we give details of some of the algorithms, especially those used in the comparative study of the AMC algorithms in Section V, whereas we only mention others.

² From now on, we set the known parameters to some fixed numerical values. In AWGN channel, with all parameters perfectly known, $\theta = \Delta f = \varepsilon = \varphi = \{\phi_k\}_{k=1}^K = 0$ and $\alpha = 1$, whereas in a block fading channel $\Delta f = \varepsilon = \{\phi_k\}_{k=1}^K = 0$, $S = 1$, and θ is included into φ . Of course, the unknowns will be put into the vector \mathbf{v}_i .

$$\Lambda_A^{(i)}[\mathbf{r}(t)] = \prod_{\ell=1}^L \prod_{k=1}^K E_{s_k^{(i)}} \left\{ \exp \left[2\alpha_\ell \sqrt{S} N_0^{-1} \operatorname{Re} \{ e^{-j\varphi_\ell} R_{k,\ell}^{(i)} \} - \alpha_\ell^2 S T N_0^{-1} |s_k^{(i)}|^2 \right] \right\}, \quad (12)$$

where L is the number of antennas at the receiver side, $\alpha_\ell e^{j\varphi_\ell}$ is the fading process on each branch, $\ell = 1, \dots, L$, and $R_{k,\ell}^{(i)} = \int_{(k-1)T}^{kT} r_\ell(t) s_k^{(i)*}(t) dt$, $k = 1, \dots, K$, $\ell = 1, \dots, L$, with $r_\ell(t) = s_\ell(t; \mathbf{u}_i) + n_\ell(t)$ as the received signal on the ℓ th branch, $s_\ell(t; \mathbf{u}_i)$ as the noise free envelope, and $n_\ell(t)$ as the zero-mean AWGN, with the PSD N_0 . The expression for the noise free envelope on the ℓ th branch, $s_\ell(t; \mathbf{u}_i)$, can be easily written similar to (9)³. Note that both the noise $\{n_\ell(t)\}_{\ell=1}^L$ and fading processes $\{\alpha_\ell e^{j\varphi_\ell}\}_{\ell=1}^L$ among the L diversity branches are assumed to be independent. Actually, a maximal ratio combining (MRC) was used here to combine the received signals. In fading channels such a structure takes advantage of the array gain, as well as diversity gain (see, for example, [97] Ch. 5), and thus, performance improvement is expected when compared with a single antenna classifier. However, as one can easily notice, when in addition to the unknown data symbols, there are other unknown parameters, e.g., $\{\alpha_\ell\}_{\ell=1}^L$ and $\{\varphi_\ell\}_{\ell=1}^L$, integration over these parameters becomes more difficult and, the implementation of a multi-antenna ALRT-based classifier turns out to be even more complex.

In AWGN, with $\mathbf{v}_i = [\theta \{s_k^{(i)}\}_{k=1}^K]^\dagger$ and uniform distribution for θ over $[-\pi, \pi)$, representing no prior knowledge of the time-invariant phase, the LF can be shown to be [21] (see (38), Appendix A)²

$$\Lambda_A^{(i)}[\mathbf{r}(t)] = E_{\{s_k^{(i)}\}_{k=1}^K} \left\{ e^{-S T N_0^{-1} \eta_K^{(i)}} I_0(2\sqrt{S} N_0^{-1} | \xi_K^{(i)} |) \right\}, \quad (13)$$

where the notation $E_{\{s_k^{(i)}\}_{k=1}^K} \{ \cdot \}$ emphasizes that the averaging is performed over K data symbols, $\eta_K^{(i)} = \sum_{k=1}^K |s_k^{(i)}|^2$, $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind, and $\xi_K^{(i)} = \sum_{k=1}^K R_k^{(i)}$. Obviously, such a classifier is difficult to implement, as requires M_i^K data sequences to compute the LF under the hypothesis H_i .

In AWGN, under the assumption of per-symbol phase-incoherence due to phase jitter, i.e., $\theta = 0$ and $\mathbf{v}_i = [\{\phi_k\}_{k=1}^K \{s_k^{(i)}\}_{k=1}^K]^\dagger$, with $\{\phi_k\}_{k=1}^K$ as i.i.d. uniform r.v.'s, it can be easily shown that the LF is given by (see (39), Appendix A)²

$$\Lambda_A^{(i)}[\mathbf{r}(t)] = \prod_{k=1}^K E_{s_k^{(i)}} \left\{ e^{-S T N_0^{-1} |s_k^{(i)}|^2} I_0(2\sqrt{S} N_0^{-1} |R_k^{(i)}|) \right\}. \quad (14)$$

In a slowly-varying flat Rayleigh fading channel², characterized via a Rayleigh-distributed α and uniform φ , such that $\mathbf{v}_i = [\alpha \varphi \{s_k^{(i)}\}_{k=1}^K]^\dagger$, the LF is given by [21] (see (42), Appendix A)

³ With a multi-antenna classifier, in AWGN channel and all parameters perfectly known, we set $\theta = \Delta f = \{\varepsilon_\ell\}_{\ell=1}^L = \{\varphi_\ell\}_{\ell=1}^L = \{\phi_k\}_{k=1}^K = 0$ and $\{\alpha_\ell\}_{\ell=1}^L = 1$, whereas with a flat block-fading channel $\Delta f = \{\varepsilon_\ell\}_{\ell=1}^L = \{\phi_k\}_{k=1}^K = 0$, $S = 1$ and θ is included into φ_ℓ , $\ell = 1, \dots, L$.

$$\Lambda_A^{(i)}[r(t)] = \mathbb{E}_{\{s_k^{(i)}\}_{k=1}^K} \left\{ \frac{1}{1 + \Omega TN_0^{-1} \eta_K^{(i)}} \exp \left[\frac{\Omega N_0^{-2} |\xi_K^{(i)}|^2}{1 + \Omega TN_0^{-1} \eta_K^{(i)}} \right] \right\}. \quad (15)$$

One can notice that the LF depends on the average fading power Ω , assumed perfectly known.

The ALRT-based classifiers are implemented by replacing the expression of the LF given in (10), (12), (13), (14) and (15), respectively, in (8), with $\eta_A = 1$; such a classifier is called the ML classifier.

Performance analysis of the ML classifier

Theoretical performance analysis of the ML classifier was performed in [2] for no unknown parameters, when identifying linearly modulated signals in AWGN, with the LF given in (10)². The probability of error under H_i is shown to be [2]

$$P_e^{(i)} = 1 - \underbrace{\int_{-\sqrt{K}m_{i,1}}^{\infty} \dots \int_{-\sqrt{K}m_{i,j-1}}^{\infty} \int_{-\sqrt{K}m_{i,j+1}}^{\infty} \dots \int_{-\sqrt{K}m_{i,N_{\text{mod}}}}^{\infty}}_{N_{\text{mod}}-1} p(\mathbf{b}_i | H_i) d\mathbf{b}_i, \quad i = 1, \dots, N_{\text{mod}}, \quad (16)$$

where

$$\mathbf{b}_i = (\mathbf{a}_i - K\mathbf{m}_i) / \sqrt{K}, \quad \mathbf{a}_i = [a_{i,1} \dots a_{i,i-1} a_{i,i+1} \dots a_{i,N_{\text{mod}}}]^\dagger, \quad a_{i,j} = \sum_{k=1}^K [\ln(\Lambda_A^{(i)}(r_k)) - \ln(\Lambda_A^{(j)}(r_k))], \quad j \neq i,$$

$$\Lambda_A^{(i)}(r_k) = \mathbb{E}_{s_k^{(i)}} \{ \exp[2\sqrt{S}N_0^{-1} \text{Re}\{r_k s_k^{(i)*}\} - STN_0^{-1} |s_k^{(i)}|^2] \}, \quad \mathbf{m}_i = [m_{i,1} \dots m_{i,i-1} m_{i,i+1} \dots m_{i,N_{\text{mod}}}]^\dagger, \quad \text{and}$$

$m_{i,j} = \mathbb{E}\{\ln(\Lambda_A^{(i)}(r_k)) - \ln(\Lambda_A^{(j)}(r_k)) | H_j\}, \quad j \neq i$. The $(N_{\text{mod}} - 1) \times 1$ vector \mathbf{b}_i is shown to be the sum of K i.i.d. random vectors which satisfies the multivariate central limit theorem if K is large. Therefore, $p(\mathbf{b}_i | H_i)$ was considered as a multivariate Gaussian density, with zero mean and covariance matrix, $\mathbb{E}\{\mathbf{b}_i \mathbf{b}_i^\dagger | H_i\}$, depending on the first and second order statistics of $\ln(\Lambda_A^{(i)}(r_k)) - \ln(\Lambda_A^{(j)}(r_k))$ [2]. The integral in (16) was numerically calculated for V. 29, 16-QAM, 32-QAM and 64-QAM, for $K = 100, 200$ and 1000 [2]. Note that V.29 is a special QAM modulation, with 16 points in the signal constellation [2].

Quasi ALRT-based classifier

A synchronous classifier ($\varepsilon = 0$) can be simply transformed into an asynchronous one, with the timing offset $\underline{\varepsilon}$ as a uniformly distributed r.v. over $[0,1)$, using the following approximation of the LF [9], [10], [13]

$$\Lambda_A^{(i)}[r(t)] \approx D^{-1} \sum_{d=0}^{D-1} \Lambda[r(t) | \varepsilon_d, H_i], \quad (17)$$

where D is the number of levels to which the timing offset is quantized and $\varepsilon_d = d/D, \quad d = 0, \dots, D-1$. For such a scenario, $R_k^{(i)}$, as defined in (11), needs to be replaced by $R_k(\varepsilon_d) = \int_{(k-1)T+\varepsilon_d T}^{kT+\varepsilon_d T} r(t) s_k^{(i)*}(t) u_T(t - (k-1)T - \varepsilon_d T) dt$. This approximation improves as $D \rightarrow \infty$, since the summation converges to an integral. The value of D directly determines the classifier complexity, as introduces more terms in (17). Note that a similar approximation can be also used when the carrier phase θ is unknown, by discretizing the range of its values.

3.1.2. ALRT-AMC for linearly modulated signals

In the sequel we present various classifiers for linear modulation classification under different conditions, such as a differential ALRT algorithm designed for unknown carrier phase, quasi ALRT classifiers designed also for unknown carrier phase, as well as unknown carrier phase/ timing offset, etc.

Differential ALRT with unknown carrier phase

A differential data solution was proposed in [1] to classify linearly modulated signals in AWGN channel, with the unknown carrier phase uniformly distributed over $[-\pi, \pi)$, i.e., $\mathbf{v}_i = [\theta \{s_k^{(i)}\}_{k=1}^K]^\dagger$. The joint PDF of the magnitude $|r_k|$ and phase difference $\Delta\psi_k = |\psi_{k+1} - \psi_k|_{\text{mod } 2\pi}$, such that $\tan \psi_k = r_{k,Q} / r_{k,I}$ and $r_k = r_{k,I} + jr_{k,Q}$, was used to derive the LF. Under the assumption that the signal amplitude and phase difference are approximately independent r.v.'s at large SNRs, the LF is given by [1]

$$\Lambda_A^{(i)}[r(t)] = \prod_{k=1}^{K-1} E_{\{s_k^{(i)}, s_{k+1}^{(i)}\}} \left\{ p(|r_k| | s_k^{(i)}, H_i) p(\Delta\psi_k | s_k^{(i)}, s_{k+1}^{(i)}, H_i) \right\}, \quad (18)$$

in which $|r_k|$ was taken as Ricean-distributed and, for large SNRs, $\Delta\psi_k$ was approximated by a Gaussian PDF. Due to the lack of space we omit here the expressions for the PDFs of $|r_k|$ and $\Delta\psi_k$. For details see [1], p. 219. The advantage of using the phase difference instead of phase itself is that the effect of a time-invariant phase offset will be mitigated. However, the classifier performance can still be degraded due to the phase jitter. The differential ALRT-based classifier was implemented with the decision rule given in (8) ($\eta_A = 0$) and the expression of the LF given in (18).

Quasi ALRT with unknown carrier phase⁴

With $\mathbf{v}_i = [\theta \{s_k^{(i)}\}_{k=1}^K]^\dagger$, where θ is uniformly distributed over $[-\pi, \pi)$, approximations of the LF were developed in [3]-[5], [7] for AWGN channel, leading to suboptimal classification structures. A low SNR approximation of the LF for PSK and QAM signals is given by⁵

$$\Lambda_A^{(i)}[r(t)] \approx \exp \left\{ \sum_{n=1}^{\infty} (\sqrt{S} N_0^{-1})^n K \sum_{q=0}^{\lfloor n/2 \rfloor} \left[\upsilon_{n-2q} (q!(n-q)!)^{-1} |m_{s^{(i)}, n, q} | \hat{m}_{r, n, q}(\mathbf{0}_{n-1})| \right] \right\}, \quad (19)$$

where $m_{s^{(i)}, n, q} = E_{s^{(i)}} \{ (s^{(i)})^{n-q} (s^{(i)*})^q \}$ is the n th-order/ q -conjugate moment of the i th constellation, $\hat{m}_{r, n, q}(\mathbf{0}_{n-1}) = K^{-1} \sum_{k=1}^K r_k^{n-q} (r_k^*)^q$ is the sample estimate of the n th-order/ q -conjugate moment at zero-delay vector $\mathbf{0}_{n-1}$ ⁶, $\lfloor \cdot \rfloor$ denotes rounding up to the nearest integer, and finally υ_{n-2q} is 1 if $q = n/2$ and 2 if $q < n/2$. Eq. (19)

⁴ The quasi ALRT classifier was originally derived as an LB method. However, it can be considered as an FB technique as well.

⁵ Eq. (19) can be easily obtained from eq. (5) given in [7], by using the signal moments.

⁶ For the definition of the n th-order moment/ q -conjugate, $m_{r, n, q}(\boldsymbol{\tau}_{n-1})$ and the cumulant $c_{r, n, q}(\boldsymbol{\tau}_{n-1})$ of a stationary random process, as well the relations between the moments and cumulants, see, e.g., [98] and [99] Ch.2. A $\mathbf{0}_{n-1}$ delay vector is an $(n-1) \times 1$ vector, with all the elements equal to zero.

can be further simplified, as for symmetric constellations the odd order moments are equal to zero [63]. The right-hand side of (19) is actually a measure of the correlation between the theoretical moments of the i th constellation and sample estimates. Theoretical values of the n th order/ q -conjugate moment $m_{s^{(i)},n,q}$, with $n = 2, 4, 6, 8$ and $q = 0, \dots, \lfloor n/2 \rfloor$, are given in Table II for different signal constellations. These values were computed as ensemble average over the ideal-noise free constellation under the constraint of unity variance and the assumption of equiprobable symbols. As one can notice from Table II, the lowest order statistic to distinguish between M -PSK and M' -PSK ($M' > M$) is the M th-order / zero-conjugate moment, $m_{s^{(i)},M,0}$. On the other hand, also from Table II, this property does not hold for QAM signals. A general rule for QAM signals is that $m_{s^{(i)},n,q} = 0$ when n is a multiple of four ($n = 4$ and 8 in Table II) and q is odd or n is not a multiple of four ($n = 6$ in Table II) and q is even.

By resorting to only the lowest order statistics, small n 's, suboptimal but implementationally manageable classifiers were proposed to discriminate PSK and QAM in [3]-[5] and [7], respectively. Based on the aforementioned property of moments for PSK signals, Polydoros et al. proposed a binary decision tree classifier for PSK signals [3]-[5], where the decision at each node was made by comparing the following metric

$$v_{M,PSK} = \left| \sum_{k=1}^K r_k^M \right| \quad (20)$$

against a threshold, denoted here by $\eta_{v_M}^{(M-PSK, M'-PSK)}$. As an example, see Appendix B for the derivation of (20) for BPSK/QPSK classification. The decision rules involving the approximations of the LFs require appropriate thresholds. In order to maximize the probability of correct classification when discriminating between the M -PSK and M' -PSK modulations, the threshold $\eta_{v_M}^{(M-PSK, M'-PSK)}$ was chosen to satisfy

$$p^{(M-PSK)}(\eta_{v_M}^{(M-PSK, M'-PSK)}) = p^{(M'-PSK)}(\eta_{v_M}^{(M-PSK, M'-PSK)}), \quad (21)$$

where $p^{(M-PSK)}(v_M)$ and $p^{(M'-PSK)}(v_M)$ are the PDFs of the metric v_M under the hypothesis that M -PSK and M' -PSK are the modulation formats of the incoming signal, respectively. A closed-form solution was derived in [5] for classifying PSK signals with unknown carrier phase. Under the assumption of a large number of available symbols K , this threshold was approximated by⁷

$$\eta_{v_M}^{(M-PSK, M'-PSK)} = KS^{M/2}T^M / 2. \quad (22)$$

As one can notice, the threshold depends on the signal power, S . In the sequel we denote it by $\eta_{QA,TE}$, where QA and TE stand for quasi-ALRT and theoretical, respectively. An example of a binary decision tree classifier used for PSK signals is shown in Fig. 2.

⁷ The original presentation in [4] used $2r_k/\sqrt{N_0T}$ instead of r_k in the metric in (20). To account for this difference, the threshold we give in (22) is indeed $(\sqrt{N_0T}/2)^M$ times the threshold of [4].

As previously explained, the n th order/ q -conjugate moments for QAM signals do not have the attractive property used to devise a binary decision tree for PSK signal identification. Using (19) and the results given in Table II, it can be easily noticed that the lowest order statistics which can be used for QAM signals classification is of order $n = 4$ ($q = 0, 2$). Following the same procedure as in the example given in Appendix B for BPSK and QPSK signals, it can be easily shown that the lowest order metric which can be used to distinguish between any two QAM signals is given by

$$v_{4,QAM} = A \left| \sum_{k=1}^K r_k^4 \right| + B \left| \sum_{k=1}^K |r_k|^4 \right|, \quad (23)$$

where the coefficients A and B depend on the theoretical values of the fourth-order/ zero- and two-conjugate moments of the QAM signals, respectively. Such a metric was used in [7] to discriminate between 16-QAM and V.29, with $A = 0.0135$ and $B = -0.0246$. The decision was made by comparing the metric against a threshold, which was empirically set⁸. We denote this threshold by $\eta_{QA,E}$, where QA and E stand for quasi-ALRT and empirical, respectively.

By comparing (20) and (23) with (13), one can say that the complexity of a quasi-ALRT classifier is much less than that of the ALRT classifier, as it needs neither an averaging operation nor the computation of the Bessel function.

Quasi ALRT with unknown carrier phase and timing offset

In AWGN, with $\mathbf{v}_i = [\theta \ \varepsilon \ \{s_k^{(i)}\}_{k=1}^K]^\dagger$, where the carrier phase θ and timing offset ε are uniformly distributed over $[-\pi, \pi)$ and $[0, 1)$, respectively, the following statistic was used to distinguish between M -PSK and M' -PSK [5]

$$v_{M,PSK} = \sum_{d=0}^{D-1} \left| \sum_{k=1}^K r_k^M (d/D) \right|, \quad (24)$$

with $r_k(\varepsilon_d) = \int_{(k-1)T+\varepsilon_d}^{kT+\varepsilon_d} r(t) u_T(t - (k-1)T - \varepsilon_d T) dt$. Similar to Fig. 2, a binary decision tree classifier was employed for PSK signal classification, with (24) compared against a threshold. The threshold was empirically chosen, following the ‘‘histogram method’’⁸.

ALRT with unknown signal level

An ALRT algorithm was developed in [8] to identify PSK signals in AWGN, with $\mathbf{v}_i = [\alpha \ \{s_k^{(i)}\}_{k=1}^K]^\dagger$, where the signal level α is a Rayleigh-distributed r.v. The decision rule given in (8) was employed, with the threshold η_A set to one.

Miscellaneous classifiers

Numerical calculation of the integrals in ALRT using a Markov chain Monte Carlo method was performed in [24]. In the algorithms previously described, AMC was treated as a hypothesis testing problem with a fixed

⁸ This threshold is set to maximize the average probability of correct classification over a large number of data and noise realizations. It is assumed that such simulations can be run off-line and the threshold can be stored as a function of the noise and signal parameters. In a practical implementation, the threshold is therefore obtained from a look-up table.

number of received symbols. By formulating the AMC problem as a variable sample size hypothesis testing problem, an algorithm based on the sequential probability ratio test was proposed in [25], [26].

3.1.3. ALRT-AMC for FSK signals

ALRT-based classifier under the assumption of per-symbol phase-incoherence

In AWGN, under the assumption of per-symbol phase-incoherence, i.e., $\mathbf{v}_i = [\{\phi_k\}_{k=1}^K \{s_k^{(i)}\}_{k=1}^K]^\dagger$, with $\{\phi_k\}_{k=1}^K$ as i.i.d. uniform r.v.'s, the LF of (14), with $|s_k^{(i)}| = 1$, was derived in [9]. As one can easily notice from (11) and (14), the implementation of such classifier requires the explicit calculation of the Fourier spectrum of the received waveform at a set of M_i candidate frequencies [9]. The signal bandwidth of an M_i -FSK signal is defined as $BW_i = M_i T_i^{-1}$, with T_i as the symbol period under the hypothesis H_i . Optimal structures were developed for the three possible cases: the same signal bandwidth and distinct symbol period (equal $M_i T_i^{-1}$, with different T_i), the same symbol period and distinct bandwidth (different $M_i T_i^{-1}$, with equal T_i), and both the symbol period and signal bandwidth different (different $M_i T_i^{-1}$, with different T_i) [9]. The decision was made based on the rule given in (8), with $\eta_A = 0$.

Quasi ALRT-based classifier under the assumption of per-symbol phase-incoherence⁴

On the other hand, with $\mathbf{v}_i = [\{\phi_k\}_{k=1}^K \{s_k^{(i)}\}_{k=1}^K]^\dagger$, by using a power series expansion of the modified Bessel function in (14), an approximation of the LF based on higher-order correlations (HOCs) was given in [9]-[12]. For a single symbol interval this is given by

$$\Lambda_A^{(i)}[r(t)] \approx 1 + C_1 c_{1,r}(0) + C_2 \int_{-T}^T |c_{1,r}(\tau)|^2 d\tau + C_3 \int_{-T}^T c_{1,r}(\tau) c_{2,r}^*(\tau) d\tau + C_4 \int_{-T}^T |c_{2,r}(\tau)|^2 d\tau + C_5 \int_{-T}^T c_{1,r}(\tau) c_{3,r}^*(\tau) d\tau + \dots \quad (25)$$

where $c_{r,1}(\tau)$ is the signal autocorrelation, defined as

$$c_{r,1}(\tau) = \begin{cases} \int_0^{T-\tau} r^*(t) r(t+\tau) dt, & 0 \leq \tau < T, \\ \int_0^{T-|\tau|} r^*(t+|\tau|) r(t) dt, & -T \leq \tau < 0. \end{cases}$$

and $c_{r,n}(\tau)$ is the n th-order ($n \geq 2$) correlation, defined as the autocorrelation of $c_{r,n-1}(\tau)$,

$$c_{r,n}(\tau) = \begin{cases} \int_{-2^{n-2}T}^{2^{n-2}T-\tau} c_{r,n-1}^*(t) c_{r,n-1}(t+\tau) dt, & 0 \leq \tau < 2^{n-1}T, \\ \int_{-2^{n-2}T}^{2^{n-2}T-|\tau|} c_{r,n-1}^*(t+|\tau|) c_{r,n-1}(t) dt, & -2^{n-1}T \leq \tau < 0. \end{cases}$$

The coefficients C_1, C_2, C_3, C_4 , and C_5 are given by $\gamma_s(2N_0T \times BW)^{-1}$, $\gamma_s(16N_0T \times BW)^{-1}$, $\gamma_s(288N_0T \times BW)^{-1}$, $\gamma_s(9216N_0T \times BW)^{-1}$ and $\gamma_s(460800N_0T \times BW)^{-1}$, respectively. When the received signal includes K i.i.d. symbols, the LF is simply the product of K LFs, each one given by (25). To increase the accuracy of the LF approximation in (25), one needs to use higher-order correlations. This results in a better but

more complex classifier. The decision was made by comparing the metric against an empirical threshold, $\eta_{HOC,E}$, where HOC and E denotes higher-order correlation and empirical, respectively.

An FSK signal classifier based also on an approximation of the LF was proposed in [13] for flat Rayleigh fading channels. Usually, in the literature only one incoming signal is assumed at the receiver. The case of multiple FSK signals at the receiver was discussed in [13]. Using (17), the previously discussed ALRT- based and quasi-ALRT synchronous algorithms were transformed into asynchronous classifiers, where the timing offset ε was assumed uniformly distributed over $[0,1)$ [10], [12], [13].

3.2. GLRT- and HLRT-based Algorithms

As the ALRT algorithm suffers from high computational complexity in most practical cases, GLRT and HLRT algorithms have been investigated as possible solutions to identify linear modulations [14]-[22]. In AWGN and with $\mathbf{v}_i = [\theta \{s_k^{(i)}\}_{k=1}^K]^\dagger$, the LF for GLRT and HLRT are respectively given by [14]

$$\Lambda_G^{(i)}[r(t)] = \max_{\theta} \left\{ \sum_{k=1}^K \max_{s_k^{(i)}} \left(\text{Re}[s_k^{(i)*} r_k e^{-j\theta}] - 2^{-1} \sqrt{ST} |s_k^{(i)}|^2 \right) \right\}, \quad (26)$$

$$\Lambda_H^{(i)}[r(t)] = \max_{\theta} \left\{ \prod_{k=1}^K E_{s_k^{(i)}} \left\{ \exp \left[2\sqrt{S} N_0^{-1} \text{Re}[s_k^{(i)*} r_k e^{-j\theta}] - S T N_0^{-1} |s_k^{(i)}|^2 \right] \right\} \right\}. \quad (27)$$

These relations can be easily derived according to (6) and (7), respectively, after substituting (9)² into (5). The empirical histogram method⁸ was used to set the thresholds with the GLRT and HLRT tests [14]. Here we denote these thresholds as $\eta_{G,E}$ and $\eta_{H,E}$, respectively.

GLRT displays some implementation advantages over ALRT and HLRT, as it avoids the calculation of exponential functions and does not require the knowledge of noise power to compute the LF. However, it suffers from the nested constellation problem discussed earlier. Note that HLRT does not have this problem.

Other GLRT- and HLRT-based classifiers investigated in the literature are as follows. The AMC problem was examined in an intersymbol interference (ISI) environment, where the signal was considered to be degraded by AWGN and ISI [17], [18]. The LF was computed using the ML estimates of the data sequence and channel coefficients $\{g_p\}_{p=1}^P$, with the per-survivor processing technique employed for estimation [18]. Obviously, this is GLRT with $\mathbf{v}_i = [\{g_p\}_{p=1}^P \{s_k^{(i)}\}_{k=1}^K]^\dagger$. The threshold used for decision was empirically set⁸. With $\mathbf{u}_i = [\theta S \{s_k^{(i)}\}_{k=1}^K]^\dagger$ and unknown PSD N_0 , i.e., $\mathbf{v}_i = [\theta S \{s_k^{(i)}\}_{k=1}^K N_0]^\dagger$, an HLRT classifier was explored in [15], where ML estimates of S and N_0 were used, together with the approximate LF in (19), obtained by averaging over θ and $\{s_k^{(i)}\}_{k=1}^K$. In other words, the LF is computed based on (7), with $\mathbf{v}_{i_1} = [\{s_k^{(i)}\}_{k=1}^K]^\dagger$ and $\mathbf{v}_{i_2} = [S N_0]^\dagger$. Both empirical and theoretical thresholds were used. An HLRT-based multi-antenna classifier was developed for BPSK/QPSK in AWGN [16], with $\mathbf{v}_i = [\vartheta \{s_k^{(i)}\}_{k=1}^K]^\dagger$, where ϑ is an unknown phase shift between two adjacent antenna elements, which appears due to their spatial separation. The decision rule given in (8) was employed, with the threshold set to one. HLRT-based classifiers were developed for linear modulation identification in flat block fading channels in

[19]-[20], with $\underline{\mathbf{v}}_i = [\alpha \varphi \{s_k^{(i)}\}_{k=1}^K]^\dagger$ and $\underline{\mathbf{v}}_i = [\alpha \varphi \{s_k^{(i)}\}_{k=1}^K N_0]^\dagger$, respectively. The LF was computed by averaging over the data symbols and using the ML estimates of the unknown parameters. The threshold η_H used for decision was set to one. HLRT has the advantage over ALRT that no prior PDFs of the channel parameters are needed, and therefore, it is applicable to different environments, e.g., Rician and Rayleigh fading [19]-[20].

Quasi HLRT classifier

HLRT does not seem to be a good solution with an increased number of unknown parameters, as finding their ML estimates can be very time consuming. Quasi-HLRT classifiers, which use low-complexity yet accurate estimates, can be used instead. Such classifiers were proposed in [20]-[21] to identify linear digital modulations in block fading, with $\underline{\mathbf{v}}_i = [\alpha \varphi \{s_k^{(i)}\}_{k=1}^K]^\dagger$ and $\underline{\mathbf{v}}_i = [\alpha \varphi \{s_k^{(i)}\}_{k=1}^K N_0]^\dagger$, respectively. A quasi-HLRT classifier was also proposed in [22] to discriminate QAM signals in AWGN channels, with $\underline{\mathbf{v}}_i = [\Delta f \{s_k^{(i)}\}_{k=1}^K]^\dagger$. As an example, to discriminate QAM signals in block fading channels, the estimators used in [21] for the channel amplitude and phase are given, respectively, by

$$\hat{\alpha}_i^4 T^4 = [2\hat{m}_{r,2,1}^2(0) - \hat{m}_{r,4,2}(\mathbf{0}_3)](2m_{s^{(i)},2,1}^2 - m_{s^{(i)},4,2})^{-1}, \quad (28)$$

and

$$\hat{\phi} = 4^{-1} \text{angle}(\sum_{k=1}^K r_k^4). \quad (29)$$

Then, using these estimators, the LF was calculated according to²

$$\Lambda_{quasi-H}^{(i)}[r(t)] = \prod_{k=1}^K E_{s_k^{(i)}} \left\{ \exp \left[2\hat{\alpha}_i N_0^{-1} \text{Re}[s_k^{(i)*} r_k e^{-j\hat{\phi}}] - \hat{\alpha}_i^2 T N_0^{-1} |s_k^{(i)}|^2 \right] \right\}. \quad (30)$$

The decision rule is given in (8), with the threshold set to one. A multi-antenna quasi-HLRT classifier was also proposed in [21], with (28) and (29) used to estimate the channel phase and amplitude on each branch, respectively. The threshold used for decision was set to one.

4. FEATURE-BASED APPROACH TO AMC

The design of a FB algorithm first needs some features for data representation and then decision making [100]. Examples of features are the correlation between the in-phase and quadrature signal components [27], the variance of the centered normalized signal amplitude, phase and frequency [28], the variance of the zero-crossing interval [32], [33], the variance of the magnitude of the signal wavelet transform (WT) after peak removal [36]-[38], the phase PDF [44]-[46] and its statistical moments [47]-[49], moments, cumulants, and cyclic cumulants of the signal itself [41]-[43], [53], [54], [58]-[66], etc. The entropy [67], [68], fuzzy logic [69], [70], a moment matrix technique [71], [72] and a constellation shape recovery method [73] were also used for AMC. Different methods were employed for decision making, such as PDF-based [41]-[53], the Hellinger distance [74], [75], the Euclidian distance [60]-[65] and unsupervised clustering techniques [76], [77]. Table III summarizes most of the FB-AMC work, emphasizing the selected features, type of modulations, channel and the unknown parameters. In its first part, algorithms which employ information extracted from the instantaneous amplitude,

phase and frequency of the received signal are presented. The second and third parts include classifiers based on the wavelet transform and signal statistics, respectively. Finally, a classifier based on spectral properties of FSK signals is mentioned. In what follows, these FB-AMC algorithms are presented¹ with a hierarchical approach in mind, i.e., the modulation class of the incoming signal is first identified (e.g., ASK, PSK, QAM, FSK), and then the modulation order M within the recognized class. A single signal in AWGN, with the parameters perfectly known, and a rectangular pulse shape $u_r(t)$ were assumed, except otherwise mentioned. In addition to ASK, PSK, QAM and FSK, the identification of other modulations was examined in the literature, e.g., MSK [66], [78]-[80], OQPSK [79], continuous-phase FSK (CPFSK) [81].

4.1. FB Algorithms to Distinguish between Different Classes

Instantaneous amplitude, phase and frequency-based algorithms

The most intuitive way to identify the modulation class of the incoming signal is to use the information contained in its instantaneous amplitude, phase and frequency. To extract such information, different methods were applied in the literature [28]-[40]. The following differences between signal classes were employed for classification in [28]-[31]: •FSK signals are characterized by constant instantaneous amplitude, whereas ASK signals have amplitude fluctuations, and PSK signals have information in the phase. The maximum of the discrete Fourier transform (DFT) of centered⁹ normalized instantaneous amplitude was used as a feature to distinguish between FSK and ASK/ PSK classes, •ASK and BPSK signals have no information in the absolute phase, whereas M -PSK ($M > 2$) has. The variance of absolute centered⁹ normalized phase was used to distinguish between M -PSK ($M > 2$) and real-valued constellation, BPSK and ASK, •ASK signals have no phase information by their nature, whereas BPSK has. Variance of direct (not absolute) centered⁹ normalized phase was used to distinguish between BPSK and ASK classes. A binary decision tree structure was employed to discriminate between classes, and furthermore, within each class, as we will briefly mention in Sections 4.2 and 4.3. At each node of the tree, the decision was made by comparing a statistic against a threshold⁸.

In [32] and [33], the variance of the zero-crossing interval was used as a feature to distinguish FSK from PSK and the unmodulated waveform (UW). The zero-crossing interval is a measure of the instantaneous frequency, and it is a staircase function for FSK signals, whereas a constant for UW and PSK signals. The AMC is treated as a two hypothesis testing problem: H_1 for FSK and H_2 for UW and PSK. The hypotheses are formulated based on the Gaussian assumption for the estimated feature, i.e, $\mathcal{N}(\mu_{H_i}, \sigma_{H_i}^2)$, $i = 1, 2$, with the hypothesis-dependent mean μ_{H_i} ¹⁰ and variance $\sigma_{H_i}^2$. An LRT is used for decision, which due to the Gaussian assumption is simplified to the comparison of the feature against a threshold η , derived from the LRT. For any two class problem, assuming equal priors, the average probability of error is then given by

⁹ The term “centered” specifies that the average is removed from the data set.

¹⁰ The mean is actually the theoretical value of the feature under H_i , whereas the variance is estimated under each hypothesis.

$$P_e = [\text{erfc}((\eta - \mu_{H_1})/\sigma_{H_1}) + \text{erfc}((\mu_{H_2} - \eta)/\sigma_{H_2})]/2, \quad (31)$$

where $\text{erfc}(\cdot)$ is the complementary error function, defined as $\text{erfc}(x) = (2\pi)^{-1/2} \int_x^\infty \exp(-u^2/2) du$.

The variance of the instantaneous frequency was also employed in [34], [35] to discriminate FSK from UW and PSK. In fact, the autoregressive spectrum modeling was used to extract the instantaneous frequency. The decision was made by comparing the feature against a threshold⁸.

Wavelet transform-based algorithms

The utility of the wavelet transform to localize the changes in the instantaneous frequency, amplitude and phase of the received signal was also studied for AMC. The distinct behavior of the Haar WT (HWT) magnitude for PSK, QAM and FSK signals was employed for class identification in [36]-[38]. For a PSK signal this is a constant, with peaks occurring at phase changes. On the other hand, because of the frequency and amplitude variations in FSK and QAM, respectively, the HWT magnitude is a staircase function with peaks at phase changes. These peaks do not provide useful information for non-continuous phase FSK signals. If only the phase is retained for a QAM signal, it behaves like a PSK signal and thus, the HWT magnitude is constant. On the other hand, as PSK and FSK signals are of constant amplitude, amplitude normalization has no effect on their HWT magnitude. After peak removal, the variance of the HWT magnitude with amplitude normalization was used to discriminate FSK from PSK and QAM. Furthermore, the variance of the HWT magnitude without amplitude normalization was employed to distinguish between QAM and PSK. The decisions were made by comparing the features against some thresholds, chosen based on the statistical analysis of the features, to minimize the probability of error for PSK signals [36]-[38].

Neural networks (NNs) were also used for classification in [28]-[31], [35]. The Wigner-Ville distribution was used in [82] to distinguish between PSK and FSK signals.

Signal statistics-based algorithms

To discriminate among BPSK, ASK, M -PSK ($M > 2$) and QAM, the cumulant-based feature $c_{r,4,0}(\mathbf{0}_3)/c_{r,2,1}^2(0)$ was proposed in [41], where $c_{r,n,q}(\mathbf{0}_{n-1})$ ⁶ is the n th-order/ q -conjugate cumulant of the output of the matched filter $\{r_k\}_{k=1}^K$, at the zero delay vector. For decision, an LRT based on the PDF of the sample estimate of the feature was formulated to achieve minimum probability of error. The moment-based feature $m_{r,6,3}(\mathbf{0}_5)/m_{r,2,1}^3(0)$ was used in [42], where $m_{r,n,q}(\mathbf{0}_{n-1})$ is the n th-order/ q -conjugate moment of the output of the matched filter $\{r_k\}_{k=1}^K$, at the zero delay vector. The goal was to distinguish between PSK and QAM. A joint power estimation and classification was performed in [42]. The decision was made based on the minimum absolute value of the difference between the sample estimate and prescribed values of the feature. Reference [43] combined several normalized moments and cumulants for training a NN, to identify FSK, PSK and QAM in multipath environments.

4.2. FB Algorithms for Linearly Modulated Signals

Classifiers summarized in Table III, which can be applied to identify the modulation order M of linear modulations, are discussed in the sequel.

Instantaneous amplitude and phase-based algorithms

Information extracted from the instantaneous amplitude and phase of the received signal was exploited for linear modulation recognition, as follows. The variance of the absolute value of the normalized centered⁹ instantaneous amplitude was used to distinguish between 2-ASK and 4-ASK, as for the former the amplitude changes between two levels, equal in magnitude and opposite in sign, so, it has no information in the absolute amplitude, whereas it has for the latter [28]-[31]. The statistic was compared against a threshold for decision making⁸ at a tree node, as part of the binary decision tree classifier mentioned in Section 4.1. The phase PDF and its statistical moments were investigated for PSK signal recognition in [44]-[50]. The phase PDF is multimodal, and the number of modes provides information for the PSK order identification. In the high-SNR region, M -PSK exhibits M distinct modes, while when the SNR decreases or M increases, the peaks smear off and finally the PDF converges to a uniform PDF [48]. Specifically for PSK signal classification, an approximation using the Tikhonov PDF and a Fourier series expansion of the phase PDF were employed in [44]-[46], with a log-likelihood ratio test for decision. By using these methods to compute the phase PDF, closed-form expressions for the phase statistical moments were derived, and the PDFs of the sample estimates of the moments were used for decision making [47]-[50]. The distribution of the sample estimate of the n th-order moment was assumed to be Gaussian, $\mathcal{N}(\mu_{n,H_i}, \sigma_{n,H_i}^2)$, where the mean μ_{n,H_i} and variance σ_{n,H_i}^2 depend on the hypothesis H_i and n . The decision criterion was further reduced to comparing the sample estimates of the phase moments with a threshold. The histogram of the phase difference between two adjacent symbols was used in [32], [33], [39] for PSK order identification, with the decision made based on the comparison of the histogram against particular patterns. The periodic components of the phase PDF were analyzed for PSK order identification in [51], using the DFT of the phase histogram. In other words, the empirical characteristic function of the phase was exploited for classification in this work. Furthermore, in [52] the algorithm was extended to QAM signal classification, by exploiting the additional information provided by the magnitude of the received signal. Other features extracted from the instantaneous amplitude and phase were investigated for PSK and QAM identification in [40], [78], [83], [84], such as the kurtosis of the amplitude.

Wavelet transform-based algorithm

Different PSK signals give rise to different sets of peak values in the magnitude of the Haar wavelet transform. The histogram of the peak magnitudes was employed to identify the order of a PSK signal in [37], with the decision made by comparing the histogram with the theoretical PDFs corresponding to different orders.

Signal statistics-based algorithms

Cumulant-based features were proposed in [41] to identify the order of ASK, PSK, and QAM modulations, as follows: the normalized cumulant of fourth-order/ two-conjugate, $c_{r,4,2}(\mathbf{0}_3)/c_{r,2,1}^2(0)$, for ASK, the magnitude of the normalized cumulant of fourth-order/zero-conjugate, $|c_{r,4,0}(\mathbf{0}_3)/c_{r,2,1}^2(0)|$, for PSK ($M > 2$), and the normalized cumulant of fourth-order/zero-conjugate, $c_{r,4,0}(\mathbf{0}_3)/c_{r,2,1}^2(0)$, for QAM. The theoretical values of the n th-order/ q -conjugate cumulant $c_{s^{(i)},n,q}$, $q = 0, \dots, n/2$, n even, for several linear modulations are given in Table IV. These

values were computed using the moment to cumulant formula⁶, in which the n th-order moments were calculated as ensemble averages over the noise-free unit-variance constellations with equiprobable symbols. Note that due to the symmetry of the signal constellations considered, the n th-order moments for n odd are zero and hence, using the moment to cumulant formula, it is easy to show that the n th-order cumulants for n odd are also zero. On the other hand, for n even we have $c_{s^{(i)},n,q} = c_{s^{(i)},n,n-q}$. An LRT was formulated based on the PDFs of the sample estimates of features, which are Gaussian, i.e., $\mathcal{N}(\mu_{H_i}, \sigma_{H_i}^2)$ ¹⁰. With a simplifying approximation, i.e., equal variances under all the hypotheses, the decision was further reduced to comparing the sample estimate of the chosen feature $\hat{\omega}$ against a threshold, with ω as any of the cumulant-based features previously mentioned. For an N_{mod} hypothesis testing problem, with the hypotheses ordered such that $\mu_{H_1} < \mu_{H_2} < \dots < \mu_{H_{N_{\text{mod}}}}$, the decision rule is to choose H_i if

$$(\mu_{H_{i-1}} + \mu_{H_i})/2 < \hat{\omega} < (\mu_{H_i} + \mu_{H_{i+1}})/2, \quad (32)$$

where $\mu_{H_0} = -\infty$ and $\mu_{H_{N_{\text{mod}}+1}} = \infty$.

Note that the cumulant-based features $c_{r,4,2}(\mathbf{0}_3)/c_{r,2,1}^2(0)$ and $|c_{r,4,0}(\mathbf{0}_3)/c_{r,2,1}^2(0)|$ do not depend on a fixed carrier phase θ , as for $q = n/2$ the exponential factors which depend on θ cancel each other, whereas for $q \neq n/2$ the phase dependency is dropped by taking the magnitude. This work was extended in [53] to classify linear modulations in frequency-selective channels. The blind alphabet-matched equalization algorithm (AMA) [101], which was used for equalization, was also employed for classification. Some other cumulant-based features were added [30] to the set of features extracted from the instantaneous amplitude, phase and frequency [28]-[29], to include QAM signals in the set of candidate modulations to be recognized.

Signal moments were applied to distinguish between QPSK and 16-QAM in [54]. Specifically, a linear combination of the fourth-order/two-conjugate moment and the squared second-order/one-conjugate moment were employed, with the coefficients and the delay vector optimized to maximize the probability of correct classification. A set of features was chosen for certain values of the delay vector, and classification was made based on the correlation between the sample estimate and theoretical feature vectors. The signal-moment feature $m_{r,6,3}(\mathbf{0}_5)/m_{r,2,1}^3(0)$ was employed to identify the order of QAM signals in [42], with the decision made based on the minimum absolute value of the difference between the sample estimate and prescribed values of the feature.

Signal cyclostationarity was also exploited for linear modulation identification [55]-[65], via two approaches: spectral line generation when passing the signal through different nonlinearities [55]-[57], and periodic fluctuations with time of cumulants up to the n th-order [58]-[65]. We note that the n th-order cycle frequencies (CFs) are given by $(n-2q)\Delta f + m/T$, with m an integer [60], [63]. The n th-order CF formula also holds for an IF signal, where Δf is replaced by the IF frequency, f_{IF} . With this property, the cyclostationarity of the received signal was exploited for AMC through a pattern of sine-wave frequencies in signal polynomial transformations. For example, the $2f_{IF}$ and $4f_{IF}$ sinusoids that appear in the second and fourth powers of the received signal, respectively, were used in [55] to distinguish between BPSK and QPSK. In [56], [57] the same property was

explored for a baseband signal. By increasing the order of the nonlinear signal transformation beyond fourth powers, this argument can be extended to identify modulations of order higher than QPSK. Note that the quasi-optimal algorithm derived within the LB framework for PSK signal classification also exploits such a property, by using the information extracted in time domain [3]-[5]. However, the signal cyclostationarity is not exploited in this work, as the sampling is performed at the symbol rate T^{-1} .

Cyclic-cumulant (CC) based features of different orders were investigated for modulation classification in [58]-[65]. A feature based on fourth-order/two-conjugate and second-order/one conjugate CCs at the CF equal to the symbol rate, similar to the one that used moments [54], was proposed in [58] and [59], to identify the order of QAM modulations. The same decision criterion as in [54] was employed. In [61] a generic algorithm was proposed to exploit signal cyclostationarity for classification. A feature vector was proposed, whose components were the magnitudes of the CCs up to the n th-order, raised to the power of $2/n$, when n goes to infinity, and computed at all possible CFs and delay vectors. Apparently, such a classifier is hard to implement. Note that raising the n th-order CC magnitudes to the power of $2/n$ forces the features to take values within the same order of magnitude. Therefore, the classical Euclidian distance can be used for decision.

For linear modulations, the n th order/ q -conjugate CC of $r(t)$, where $r(t)$ is given in (1), with $\phi_k = 0$, $k = 1, \dots, K$, $g(t)$ as a raised cosine pulse shape, and $n(t)$ as the AWGN, and the set of CFs are given by [60], [63]

$$c_{r,n,q}^{(i)}(\gamma; \boldsymbol{\tau}_{n-1}) = a^n c_{s^{(i)},n,q} T^{-1} e^{-j2\pi\beta\epsilon T} e^{j(n-2q)\theta} e^{j2\pi\Delta f \sum_{u=1}^{n-1} (-)_{\tau_u}} \int_{-\infty}^{\infty} g^{(*)_n}(t) \prod_{u=1}^{n-1} g^{(*)_u}(t + \tau_u) e^{-j2\pi t \beta} dt, \quad q = 0, \dots, n, \quad (33)$$

$$\kappa_{n,q} = \left\{ \gamma : \gamma = \beta + (n-2q)\Delta f, \beta = k/T, k \text{ integer}, c_{r,n,q}^{(i)}(\gamma; \boldsymbol{\tau}_{n-1}) \neq 0 \right\}, \quad (34)$$

where γ is a CF, $(*)_u$ represents a possible conjugation of the u th term, $u = 1, \dots, n$, such that the total number of conjugations is q , and $(-)_{\tau_u}$ is the minus sign associated with the possible conjugation $(*)_{\tau_u}$, $u = 1, \dots, n-1$. Since $n(t)$ is a stationary, zero-mean Gaussian process, its cumulants are time independent and non-zero only for the second order. Therefore, AWGN does not have any contribution to the higher-order ($n \geq 3$) CCs of $r(t)$. One can easily notice that by taking the magnitude of the n th order CC, a feature robust to the carrier phase and timing offset is obtained.

In [60]-[61] and [62], the magnitudes of the CCs up to the fourth- and sixth-order, at a CF equals to $\gamma = (n-2q)\Delta f + 1/T$ and a delay vector for which a maximum is reached (i.e., $\boldsymbol{\tau}_{n-1} = \mathbf{0}_{n-1}$ [63]), were investigated as features, respectively. Based on these features, a feature vector was proposed in [62], as

$$\mathbf{F}^{(i)} = [|c_{r,2,0}^{(i)}(\gamma; \mathbf{0})| \dots |c_{r,2,2}^{(i)}(\gamma; \mathbf{0})| |c_{r,4,0}^{(i)}(\gamma; \mathbf{0}_3)| \dots |c_{r,4,4}^{(i)}(\gamma; \mathbf{0}_3)| |c_{r,6,0}^{(i)}(\gamma; \mathbf{0}_5)| \dots |c_{r,6,6}^{(i)}(\gamma; \mathbf{0}_5)|]^{\dagger}, \quad i = 1, \dots, N_{\text{mod}}. \quad (35)$$

The CC-based features are estimated from $K\rho$ samples, taken over the observed K symbol interval [102]. Note that the received signal is oversampled in order to exploit signal cyclostationarity. The sampling frequency is equal to ρ/T , with ρ a positive integer, called the oversampling factor. The decision is made by comparing the sample estimate with prescribed feature vectors from a look-up table,

$$\hat{i} = \arg \min_{1 \leq i \leq N_{\text{mod}}} d(\mathbf{F}^{(i)}, \hat{\mathbf{F}}), \quad (36)$$

where \hat{i} represents the decision on the modulation type of the intercepted signal, $\mathbf{F}^{(i)}$ is the theoretical feature-vector which corresponds to the i th modulation in the look-up table, $\hat{\mathbf{F}}$ is the estimated feature vector, and $d(.,.)$ is the Euclidean distance. Joint detection and classification was investigated in [62], with single and multiple incoming signals present at the receiver, respectively. Multiple signals, which overlap in time and frequency but have distinct symbol rates and hence, different cycle frequencies, can be distinguished using CCs (selectivity property of CCs) [103]. Moreover, it is claimed that all the parameters necessary for identification, i.e., symbol period, carrier frequency offset, excess bandwidth and signal amplitude are estimated. However, no details are given about the estimation methods. Eight-order CC-based features, i.e., the magnitude of the eighth-order ($q=0, \dots, 8$) CCs at the CF $\gamma = 1/T + (n - 2q)\Delta f$ and zero delay vector were investigated in [63] for classifying real- and complex-valued constellations, respectively. The n th-order ($n = 4, 6, 8$) CC-based features were shown to be robust to a carrier frequency offset and phase jitter for $q = n/2$ and used for QAM classification in [64]. Features similar to those proposed in [63] were investigated in [65] for classifying linear modulations in block fading channels. In this case, the features were extracted from the signal at the output of a selection combiner. Such a CC-based multi-antenna classifier takes advantage of the robustness of the CC-based features to phase, timing errors and stationary noise, as well as of the ability of the selection combiner to mitigate the impact of fading via spatial diversity. Furthermore, the classifier is robust to the variations of the Ricean factor and a possible correlation among the antennas. By increasing the number of antenna elements, a smaller number of symbols is needed to attain a specific performance. In [63]-[65] the minimum Euclidian distance between the sample estimate and prescribed feature-vectors was also used for decision making. A raised-cosine pulse shape was considered in [60]-[65].

Miscellaneous classifiers

Finally, the Radon transform was investigated for QAM classification in [85]-[87].

4.3. FB Algorithms for FSK Signals

Similar to using the information contained in the instantaneous phase to identify the order of the PSK modulation, the information extracted from the instantaneous frequency is exploited to recognize the order of the FSK modulation. In [28]-[31], the variance of the absolute value of the normalized centered⁹ instantaneous frequency was used to distinguish between 2-FSK and 4-FSK. The feature was compared against a threshold⁸ for decision, at a tree node, as part of the binary decision tree classifier mentioned in Section 4.1. In [34], the instantaneous frequency derivative was used to distinguish between 2-FSK and 4-FSK, under the assumption of the same bandwidth of the signals. The height of the peaks which occur in the differentiated instantaneous frequency is proportional to the frequency deviation, and thus, for 4-FSK this is expected to be two times lower than for 2-FSK. If the peak average falls below a certain threshold, 4-FSK is chosen, otherwise 2-FSK. As for PSK order identification, the number of modes in the instantaneous frequency histogram was employed to determine the order of the FSK modulation in [32] and [33]. The number of modes in the histogram of the Haar wavelet transform magnitude was investigated for the FSK order identification in [36], [37]. If $M/2 + 1$ to M

modes appear in the histogram, the input is identified as M -FSK. Finally, spectral properties of FSK signals were explored for classification in [88].

5. NUMERICAL RESULTS

After scanning the literature to answer the question “which AMC technique is the best in terms of performance under realistic conditions?,” it turned out that performance comparison of published classifiers is not straightforward. There are a number of reasons for this. First, performance of different classifiers cannot be compared, unless the candidate modulations are the same. Second, most of the classifiers are designed to handle specific unknown parameters. So, one cannot really compare their performance, unless the uncertainties the classifiers take into account are the same. Nevertheless, to obtain some insight, we make some comparisons in Sections 5.1, 5.2, and 5.3. We rely on some numerical results reported in the literature. However, we have also simulated some of the algorithms under the same conditions, to make the comparison possible.

5.1. Comparative Study of Classifiers for PSK signals

Performance achieved with several algorithms for classifying PSK signals in AWGN is presented in Table V. We consider here the ideal scenario, i.e., no unknown parameters, as well as the scenarios with unknown carrier phase, and unknown carrier phase/ timing offset, respectively. We have examined the ALRT, quasi-ALRT, HLRT, and cumulant-based algorithms. For illustration, BPSK and QPSK are considered as candidate modulations. Of course, when higher order modulations are included in the modulation pool, higher SNRs and/or a larger number of symbols are needed to achieve the same performance [2], [5], [19]. We have simulated these modulations to draw some basic yet insightful conclusions, which shed some light on major techniques.

Subsequently, the probability of correct classification P_{cc} is used to evaluate the classification performance, with $P_c^{(ii)}$ estimated based on 1000 Monte Carlo trials. Unless otherwise mentioned, the pulse shape is rectangular, the received SNR per symbol is defined as $\gamma = ST/N_0$, the symbol period and signal power are set to $T = 1$ and $S = 1$, respectively, and the number of symbols is $K = 100$.

Under ideal conditions, the ALRT-based classifier in (10) and (8), with $\eta_A = 1$, provides a P_{cc} of 0.975 at the -3 dB SNR (V-1)¹¹. A 3dB SNR improvement is achieved with an extra antenna, i.e., $L = 2$, and maximal ratio combining at the receiver, presented in (12)³ and (8), with $\eta_A = 1$ (V-2). With a carrier phase θ uniformly distributed over $[-\pi, \pi)$, the performance of the ALRT-based classifier designed for no unknown parameters, shown in (10) and (8) with $\eta_A = 1$, drops to 0.63 at -3 dB (V-3). With the quasi-ALRT classifier in (20) and (22), with $M = 2$, a P_{cc} of 0.96 at -2 dB SNR is achieved (V-4), which is close to that under ideal condition (V-1), (V-4). As noticed from (V-5), a slightly better performance is achieved with the HLRT designed for an unknown carrier phase θ , given in (27) and (8), with the threshold η_H set to 1. Using the cumulant-based classifier in (32)

¹¹ When necessary, we refer to the i th row in Table V, Table VI and Table VII as (V- i), (VI- i), and (VII- i), respectively.

with $\hat{\omega} = |\hat{c}_{r,4,0}(\mathbf{0}_3) / \hat{c}_{r,2,1}^2(0)|$ (for the cumulant estimator formula one can see [41], eqs. (5) and (6)), $N_{\text{mod}} = 2$, $\mu_{H_1} = 1$ and $\mu_{H_2} = 2$, 4dB SNR is required to attain a P_{cc} of 0.96 (V-6). However, a timing error degrades the performance of the quasi-ALRT classifiers dramatically, i.e., a 13dB SNR loss (V-4), (V-7), whereas only 2dB loss is observed for the cumulant-based classifier (V-6), (V-8).

As expected, the ALRT-based classifier designed for an ideal scenario fails under an unknown carrier phase offset. Investigated solutions are the HLRT, quasi-ALRT and cumulant-based classifiers. A slightly better performance is achieved with HLRT, when compared with the quasi-ALRT. Also, the quasi-ALRT provides a better performance than the cumulant-based classifier. However, a synchronization error dramatically degrades the performance of the former, whereas it has a smaller effect on the latter. Though, the quasi-ALRT can be modified as in (24) to further account for the unknown timing. Performance enhancement can be achieved by using multiple receive antennas, even with $L = 2$. More examples of modulation recognition with multiple antennas can be found in [21]-[22], [65].

5.2. Comparative Study of Classifiers for QAM Signals

The performance of several algorithms is given in Table VI, when discriminating between 16-QAM and V.29 signals in AWGN. We consider here the ideal scenario, as well as the case with the unknown carrier phase θ . We have investigated the ALRT, quasi-ALRT, HLRT, quasi-HLRT, cumulant- and cyclic cumulant- based algorithms. Under ideal conditions, ALRT in (10) and (8), with $\eta_A = 1$, recognizes 16-QAM and V.29 with a P_{cc} of 0.99 at 7dB SNR (VI-1). With an unknown carrier phase θ , uniformly distributed over $[-\pi, \pi)$, the ALRT classifier designed for ideal conditions fails (VI-2), as expected. For a detailed sensitivity analysis of the ALRT-based classifier to model mismatches one can see [93]. A 2dB SNR degradation occurs when using the cumulant-based algorithm in (32), with $\hat{\omega} = \hat{c}_{r,4,0}(\mathbf{0}_3) / \hat{c}_{r,2,1}^2(0)$, $N_{\text{mod}} = 2$, $\mu_{H_1} = -0.68$, and $\mu_{H_2} = 0.5185$, under ideal conditions (VI-3). This also fails when the unknown carrier phase θ appears (VI-4). On the other hand, even the quasi-ALRT which is designed for an unknown carrier phase scenario (see (23), with $A = 0.0135$, $B = -0.0246$, and the threshold $\eta_{QA,E}$ ¹²), provides a P_{cc} of 0.88 at 30dB SNR (VI-5). An acceptable performance is attained with the HLRT classifier in (27) and (8), with $\eta_{H,E}$ ¹², designed for unknown carrier phase: a P_{cc} of 0.99 at 9dB SNR (VI-6). By using the low complexity quasi-HLRT classifier, with the LF given in (30)¹³, the carrier phase as in (29), and threshold set to 1, 19dB SNR is requested to attain the same performance (VI-7). By increasing the number of symbol to 6000, 11dB SNR is needed to obtain a P_{cc} of 0.91 (VI-7). On the other hand, with 6000 symbols, the CC-based classifier in (35) and (36) provides a P_{cc} of 0.99 at only 9dB (VI-8). Moreover, this classifier is robust to the timing offset, is applicable to a larger pool of modulations, including QAM, PSK and ASK, and benefit from the

¹² The values of these thresholds were not specified in the papers.

¹³ The LF in (30) is simplified for the AWGN channel, with the carrier phase as the only unknown parameter, and $\hat{\alpha}_i$ replaced by S , which is assumed perfectly known.

selectivity property of CCs. Note that with the CC-based classifier, the pulse shape is raised cosine, with a roll-off factor of 0.35 and the SNR is defined at the output of the root raised cosine receive filter. With multiple receive antennas one can enhance the performance, as shown for quasi-HLRT [21]-[22] and CC-based classifiers [65].

We can conclude that the ALRT- and cumulant-based classifiers, designed for an ideal scenario, fail when the unknown carrier phase offset is present. Investigated solutions are the quasi-ALRT, HLRT, quasi-HLRT, and CC-based classifiers. With the quasi-ALRT classifier, QAM signals are not identified accurately enough. As expected, the HLRT classifier provides a better classification performance compared with the quasi-HLRT. By increasing the number of processed symbols, the performance of the quasi-HLRT classifier improves. However, with enough symbols, the CC-based classifier provides superior performance when compared to the quasi-HLRT, taking also advantage of the CC properties.

5.3. Comparative Study of Classifiers for FSK signals

Here we mainly take the numerical results from [9]. Performance of the optimal ALRT and quasi-ALRT algorithms is given in Table VII, to distinguish between 32-FSK and 64-FSK in AWGN, under the assumption of per-symbol phase incoherence and identical bandwidths of the incoming signals. The symbol period of 64-FSK is two times larger than that of 32-FSK, under the assumption of the equal bandwidths. With the same observation interval, $K = 10$ and 5 symbols are considered for 32-FSK and 64-FSK, respectively, to have a fair comparison. The optimal ALRT in (14) and (8), with $\eta_A = 1$, provides a P_{cc} of 0.975 at 6dB SNR (V-1). The quasi-ALRT classifier based on the first-order correlation, i.e., (25) with the first three terms and the threshold $\eta_{HOC,E}$ ¹² for decision making, provides a P_{cc} of 0.74 at the same SNR (V-2). The quasi-ALRT classifier that employs the first- and second-order correlations, i.e., the first five terms in (25), provides a P_{cc} of 0.90 (V-3). The ALRT-based classifier essentially breaks down for a frequency drift of half of the frequency deviation, i.e., $\Delta f = f_d / 2$ (V-4), whereas the quasi-ALRT classifier is insensitive to such a model mismatch [9].

The performance of the quasi-ALRT classifier approaches to the optimal ALRT, by increasing the correlation order. The quasi-ALRT classifier has the advantage of robustness to the carrier frequency offset.

6. CONCLUSION

Based on a comprehensive literature survey, this paper has summarized the two main approaches to automatic modulation classification (AMC), i.e, the likelihood based (LB) and the feature based (FB) methods, and has lightened their advantages and drawbacks. Although the LB approach provides optimal performance, it is difficult to obtain an exact analytical solution for the decision function, when the number of unknown parameters increases. When a closed-form solution exists, the computational complexity can make the classifier impractical. By using a low SNR approximation of the likelihood function (LF), the so-called quasi-ALRT algorithms were proposed in the literature. These algorithms provide near optimal performance for identifying PSK and FSK signals. Nevertheless, they are not accurate for recognizing QAM signals. Using ML estimates of the unknown quantities, GLRT and HLRT techniques were investigated as two alternatives. Although GLRT has some

advantages, it fails in identifying nested constellations. On the other hand, HLRT does not have this problem. However, with several unknown parameters, finding the ML estimates can be very time consuming. The complexity is reduced in quasi-HLRT classifiers, which rely on low-complexity yet accurate estimators. Obviously, there is a trade-off between the complexity and performance, which depends on the estimation method. In the FB approach, some signal features are employed to identify the modulation format. Although suboptimal, FB algorithms can be simpler to implement. Many AMC algorithms which use the instantaneous amplitude, phase, frequency, signal wavelet transform, and signal statistics such as moments, cumulants, and cyclic cumulants, were compactly presented, as well as other miscellaneous techniques. In a hierarchical classification system, some of these features can be used to identify the modulation class and then, within each class, the modulation order can be determined.

Accurate preprocessing is required for the effective implementation of most of the known AMC algorithms. Devising low complexity blind algorithms for joint parameter estimation is a topic of interest in AMC. In addition, development of classification methods which rely less on preprocessing is another topic for further investigation. New classification problems have raised as a result of emerging wireless technologies, such as, single carrier versus multicarrier modulation recognition, classification of signals received from single and multiple transmit antennas, identification of space-time modulation format, etc. These issues mean that AMC in real-world environments continues to be a dynamic research field.

APPENDIX A: DERIVATION OF (10), (13), (14) AND (15)

First we assume that the parameters are known, except for the symbols, i.e., $\mathbf{v}_i = \mathbf{u}_i = [\{s_k^{(i)}\}_{k=1}^K]$. Then, by inserting (5) into (4) and averaging w.r.t. the unknown symbols, the following LF can be written

$$\Lambda_A^{(i)}[r(t)] = E_{\{s_k^{(i)}\}_{k=1}^K} \left\{ \exp \left[2N_0^{-1} \operatorname{Re} \left\{ \int_0^{KT} r(t) s^*(t; \mathbf{u}_i) dt \right\} - N_0^{-1} \int_0^{KT} |s(t; \mathbf{u}_i)|^2 dt \right] \right\},$$

which upon using the signal model in (9)² changes to

$$E_{\{s_k^{(i)}\}_{k=1}^K} \left\{ \exp \left[2N_0^{-1} \operatorname{Re} \left\{ \int_0^{KT} r(t) \alpha \sqrt{S} \sum_{k=1}^K e^{-j\varphi} s_k^{(i)*}(t) u_T(t - (k-1)T) dt \right\} - N_0^{-1} \int_0^{KT} \left| \alpha \sqrt{S} \sum_{k=1}^K e^{j\varphi} s_k^{(i)}(t) u_T(t - (k-1)T) \right|^2 dt \right] \right\},$$

and after some manipulations, becomes

$$\Lambda_A^{(i)}[r(t)] = \prod_{k=1}^K E_{s_k^{(i)}} \exp \left[2\alpha \sqrt{S} N_0^{-1} \operatorname{Re} \left\{ e^{-j\varphi} R_k^{(i)} \right\} - \alpha^2 S T N_0^{-1} |s_k^{(i)}(t)|^2 \right]. \quad (37)$$

with $R_k^{(i)} = \int_{(k-1)T}^{kT} r(t) s_k^{(i)*}(t) u_T(t - (k-1)T) dt$.

In AWGN, with $\mathbf{v}_i = [\theta \{s_k^{(i)}\}]^\dagger$ and uniform distribution for the carrier phase θ over $[-\pi, \pi)$, the same for all K symbols, the LF can be obtained by averaging (37), simplified for the AWGN channel², over the phase

$$\begin{aligned}
\Lambda_A^{(i)}[r(t)] &= \mathbb{E}_{\{s_k^{(i)}\}_{k=1}^K} \left\{ e^{-STN_0^{-1} \sum_{k=1}^K |s_k^{(i)}|^2} \mathbb{E}_\theta \left\{ \exp \left[2\sqrt{S} N_0^{-1} \sum_{k=1}^K |R_k^{(i)}| \cos(\theta + \theta_{R_k^{(i)}}) \right] \right\} \right\}, \\
&= \mathbb{E}_{\{s_k^{(i)}\}_{k=1}^K} \left\{ e^{-STN_0^{-1} \sum_{k=1}^K |s_k^{(i)}|^2} \mathbb{E}_\theta \left\{ \exp \left[2\sqrt{S} N_0^{-1} \left| \sum_{k=1}^K R_k^{(i)} \right| \cos(\theta + \theta_{\sum_{k=1}^K R_k^{(i)}}) \right] \right\} \right\}, \\
&= \mathbb{E}_{\{s_k^{(i)}\}_{k=1}^K} \left\{ e^{-STN_0^{-1} \sum_{k=1}^K |s_k^{(i)}|^2} I_0 \left(2\sqrt{S} N_0^{-1} \left| \sum_{k=1}^K R_k^{(i)} \right| \right) \right\}, \tag{38}
\end{aligned}$$

where $\theta_{R_k^{(i)}}$ and $\theta_{\sum_{k=1}^K R_k^{(i)}}$ are the phases of $R_k^{(i)}$ and $\sum_{k=1}^K R_k^{(i)}$, respectively, and $I_0(u) = \mathbb{E}_\theta \{ \exp[u \cos(\theta + \psi)] \}$, where ψ is an arbitrary phase.

Similarly, in AWGN and under the assumption of per-symbol phase-incoherence due to phase jitter, i.e., $\mathbf{v}_i = [\{\phi_k\}_{k=1}^K \ \{s_k^{(i)}\}_{k=1}^K]^\dagger$, with $\{\phi_k\}_{k=1}^K$ as i.i.d. r.v.'s, one can easily obtain the expression of the LF,

$$\begin{aligned}
\Lambda_A^{(i)}[r(t)] &= \prod_{k=1}^K \mathbb{E}_{s_k^{(i)}} \left\{ e^{-STN_0^{-1} |s_k^{(i)}|^2} \mathbb{E}_{\phi_k} \left\{ \exp \left[2\sqrt{S} N_0^{-1} |R_k^{(i)}| \cos(\phi_k + \theta_{R_k^{(i)}}) \right] \right\} \right\}, \\
&= \prod_{k=1}^K \mathbb{E}_{s_k^{(i)}} \left\{ e^{-STN_0^{-1} |s_k^{(i)}|^2} I_0(2\sqrt{S} N_0^{-1} |R_k^{(i)}|) \right\}. \tag{39}
\end{aligned}$$

With $\mathbf{v}_i = [\alpha \ \varphi \ \{s_k^{(i)}\}_{k=1}^K]^\dagger$, α a Rayleigh distributed r.v. and φ the uniform phase of the fading channel, the LF can be obtained by averaging (38), written for the block fading channel², w.r.t. α

$$\Lambda_A^{(i)}[r(t)] = \mathbb{E}_{\{s_k^{(i)}\}_{k=1}^K} \mathbb{E}_\alpha \left\{ e^{-\alpha^2 TN_0^{-1} \sum_{k=1}^K |s_k^{(i)}|^2} I_0 \left(2\alpha N_0^{-1} \left| \sum_{k=1}^K R_k^{(i)} \right| \right) \right\}. \tag{40}$$

By using

$$\mathbb{E}_\alpha \left\{ e^{-K_1 \alpha^2} I_0(\alpha K_2) \right\} = \int_0^{+\infty} e^{-K_1 \alpha^2} I_0(\alpha K_2) 2\alpha \Omega^{-1} e^{-\alpha^2/\Omega} d\alpha = \Omega^{-1} \int_0^{+\infty} 2\alpha e^{-K_3 \alpha^2} I_0(\alpha K_2) d\alpha = \Omega^{-1} K_3^{-1} e^{4^{-1} K_3^{-1} K_2^2}, \tag{41}$$

the LF becomes

$$\Lambda_A^{(i)}[r(t)] = \mathbb{E}_{\{s_k^{(i)}\}_{k=1}^K} \left\{ \frac{1}{1 + \Omega TN_0^{-1} \sum_{k=1}^K |s_k^{(i)}|^2} \exp \left[\frac{\Omega N_0^{-2} \left| \sum_{k=1}^K R_k^{(i)} \right|^2}{1 + \Omega TN_0^{-1} \sum_{k=1}^K |s_k^{(i)}|^2} \right] \right\}. \tag{42}$$

APPENDIX B: DERIVATION OF (20) FOR DISCRIMINATING BETWEEN BPSK AND QPSK

The n th order moments of symmetric constellations are zero for n odd. Using the values of the moments $m_{s^{(i)},n,q}$ for $n=2,4,6,8$ and $q=0,\dots,\lfloor n/2 \rfloor$, for BPSK and QPSK, given in the first and second columns of Table II, (19) can be written, respectively as

$$\Lambda_A^{(BPSK)}[r(t)] \approx \exp \left[K \left((\sqrt{S} N_0^{-1})^2 |\hat{m}_{r,2,0}| + 2(\sqrt{S} N_0^{-1})^2 |\hat{m}_{r,2,1}| \right) \right]$$

$$\begin{aligned}
& +12^{-1}(\sqrt{S}N_0^{-1})^4 |\hat{m}_{r,4,0}| + 3^{-1}(\sqrt{S}N_0^{-1})^4 |\hat{m}_{r,4,1}| + 4^{-1}(\sqrt{S}N_0^{-1})^4 |\hat{m}_{r,4,2}| \\
& + 360^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,0}| + 60^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,1}| + 24^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,2}| + 36^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,3}| + \dots \Big], \quad (43)
\end{aligned}$$

and

$$\begin{aligned}
\Lambda_A^{(QPSK)}[r(t)] \approx \exp \Big[& K \left(2(\sqrt{S}N_0^{-1})^2 |\hat{m}_{r,2,1}| + 12^{-1}(\sqrt{S}N_0^{-1})^4 |\hat{m}_{r,4,0}| + 4^{-1}(\sqrt{S}N_0^{-1})^4 |\hat{m}_{r,4,2}| \right. \\
& \left. + 60^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,1}| + 24^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,2}| + 36^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,3}| + \dots \right) \Big]. \quad (44)
\end{aligned}$$

The difference between the log-LFs is therefore given by

$$\ln \Lambda_A^{(BPSK)}[r(t)] - \ln \Lambda_A^{(QPSK)}[r(t)] \approx K \left((\sqrt{S}N_0^{-1})^2 |\hat{m}_{r,2,0}| + 3^{-1}(\sqrt{S}N_0^{-1})^4 |\hat{m}_{r,4,1}| + 360^{-1}(\sqrt{S}N_0^{-1})^6 |\hat{m}_{r,6,0}| + \dots \right). \quad (45)$$

As one can easily notice from (45), the lowest order statistic which can be used to discriminate between BPSK and QPSK is $K |\hat{m}_{r,2,0}| = \left| \sum_{k=1}^K r_k^2 \right|$, i.e., $v_{2,BPSK}$.

REFERENCES

- [1] J. A. Sills, "Maximum-likelihood modulation classification for PSK/QAM," in *Proc. IEEE MILCOM*, 1999, pp. 57-61.
- [2] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, pp. 189-193, 2000.
- [3] K. Kim and A. Polydoros, "Digital modulation classification: the BPSK versus QPSK case," in *Proc. IEEE MILCOM*, 1988, pp. 431-436.
- [4] A. Polydoros and K. Kim, "On the detection and classification of quadrature digital modulations in broad-band noise," *IEEE Trans. Commun.*, vol. 38, pp. 1199-1211, 1990.
- [5] C. Y. Huang and A. Polydoros, "Likelihood methods for MPSK modulation classification," *IEEE Trans. Commun.*, vol. 43, pp. 1493-1504, 1995.
- [6] P. C. Sapiano and J. D. Martin, "Maximum likelihood PSK classifier," in *Proc. ICASSP*, 1996, pp. 1010-1014.
- [7] C. Long, K. Chugg, and A. Polydoros, "Further results in likelihood classification of QAM signals," in *Proc. IEEE MILCOM*, 1994, pp. 57-61.
- [8] L. Hong and K. C. Ho, "Classification of BPSK and QPSK signals with unknown signal level using the Bayes technique," in *Proc. IEEE ISCAS*, 2003, pp. IV.1-IV.4.
- [9] B. F. Beidas and C. L. Weber, "Higher-order correlation-based approach to modulation classification of digitally frequency-modulated signals," *IEEE Journal on Sel. Areas in Commun*, vol. 13, pp. 89-101, 1995.
- [10] B. F. Beidas and C. L. Weber, "Asynchronous classification of MFSK signals using the higher order correlation domain," *IEEE Trans. Commun.*, vol. 46, pp. 480-493, 1998.
- [11] B. F. Beidas and C. L. Weber, "General framework for the high-order correlation domain," in *Proc. IEEE MILCOM*, 1995, pp. 180-185.
- [12] B. F. Beidas and C. L. Weber, "Higher-order correlation-based classification of asynchronous MFSK signal," in *Proc. IEEE MILCOM*, 1996, pp. 1003-1009.

- [13] A. E. El-Mahdy and N. M. Namazi, "Classification of multiple M -ary frequency-shift keying signals over a Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 50, pp. 967-974, 2002.
- [14] P. Panagiotou, A. Anastasopoulos, and A. Polydoros, "Likelihood ratio tests for modulation classification," in *Proc. IEEE MILCOM*, 2000, pp. 670-674.
- [15] K. M. Chugg, C. S. Long, and A. Polydoros, "Combined likelihood power estimation and multiple hypothesis modulation classification," in *Proc. ASILOMAR*, 1995, pp. 1137-1141.
- [16] L. Hong and K. C. Ho, "An antenna array likelihood modulation classifier for BPSK and QPSK signals," in *Proc. IEEE MILCOM*, 2002, pp. 647-651.
- [17] N. Lay and A. Polydoros, "Per-survivor processing for channel acquisition, data detection and modulation classification," in *Proc. ASILOMAR*, 1995, pp. 170-174.
- [18] N. Lay and A. Polydoros, "Modulation classification of signals in unknown ISI environments," in *Proc. IEEE MILCOM*, 1995, pp. 170-174.
- [19] O. A. Dobre, J. Zarzoso, Y. Bar-Ness, and W. Su, "On the classification of linearly modulated signals in fading channel," in *Proc. CISS Conf.*, 2004, Princeton University, Princeton, NJ, US.
- [20] O. A. Dobre and F. Hameed, "Likelihood-based algorithms for linear digital modulation classification in fading channels," in *Proc. IEEE CCECE*, 2006, Ottawa, Canada.
- [21] A. Abdi, O. A. Dobre, R. Choudhry, Y. Bar-Ness, and W. Su, "Modulation classification in fading channels using antenna arrays," in *Proc. IEEE MILCOM*, 2004, pp. 211-217.
- [22] H. Li, O. A. Dobre, Y. Bar-Ness, and W. Su, "Quasi-hybrid likelihood modulation classification with nonlinear carrier frequency offsets estimation using antenna arrays," in *Proc. IEEE MILCOM*, 2005, pp. 1-6.
- [23] D. Boiteau and C. Le Martret, "A generalized maximum likelihood framework for modulation classification," in *Proc. ICASSP*, 1998, pp. 2165-2168.
- [24] S. Lesage, J. Tourneret, and P. M. Djuric, "Classification of digital modulations by MCMC sampling," in *Proc. ICASSP*, 2001, pp. 2553-2555.
- [25] Y.-C. Lin and C.-C. J. Kuo, "Classification of quadrature amplitude modulated (QAM) signals via sequential probability ratio test (SPRT)," *Signal Processing*, vol. 60, pp. 263-280, 1997.
- [26] Y.-C. Lin and C.-C. J. Kuo, "Sequential modulation classification of dependent samples," in *Proc. ICASSP*, 1996, pp. 2690-2693.
- [27] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "The classification of joint analog and digital modulations," in *Proc. IEEE MILCOM*, 2005, pp. 1-6.
- [28] E. E. Azzouz and A. K. Nandi, *Automatic Modulation Recognition of Communication Signals*. Kluwer Academic, 1996.
- [29] A. K. Nandi and E. E. Azzouz, "Modulation recognition using artificial neural networks," *Signal Processing*, pp. 165-175, 1997.
- [30] M. L. D. Wong and A. K. Nandi, "Automatic digital modulation recognition using spectral and statistical features with multi-layer perceptrons," in *Proc. Int. Symp. Signal Processing and Its Applications*, Kuala Lumpur, Malaysia, 2001, pp. 390-393.
- [31] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic recognition of communication signals," *IEEE Trans. Commun.*, vol. 46, pp. 431-436, 1998.

- [32] S. Z. Hsue and S. S. Soliman, "Automatic modulation recognition of digitally modulated signals," in *Proc. IEEE MILCOM*, 1989, pp. 645-649.
- [33] S. Z. Hsue and S. S. Soliman, "Automatic modulation classification using zero crossing," *IEE Radar and Signal Processing*, vol. 137, pp. 459-464, 1990.
- [34] K. Assaleh, K. R. Farrell, and R. J. Mammone, "A new method of modulation classification for digitally modulated signals," in *Proc. IEEE MILCOM*, 1992, pp. 712-716.
- [35] K. R. Farrell and R. J. Mammone, "Modulation classification using a neural tree network," in *Proc. IEEE MILCOM*, 1993, pp. 1028-1032.
- [36] K. C. Ho, W. Prokopiw, and Y. T. Chan, "Modulation identification by the wavelet transform," in *Proc. IEEE MILCOM*, 1995, pp. 886-890.
- [37] K. C. Ho, W. Prokopiw, and Y. T. Chan, "Modulation identification of digital signals by the wavelet transform," *IEE Proc. Radar, Sonar and Navig.*, vol. 47, pp. 169-176, 2000.
- [38] L. Hong and K. C. Ho, "Identification of digital modulation types using the wavelet transform," in *Proc. IEEE MILCOM*, 1999, pp. 427-431.
- [39] F. F. Liedtke, "Computer simulation of an automatic classification procedure for digitally modulated communication signals with unknown parameters," *Signal Processing*, vol. 6, pp. 311-323, 1984.
- [40] H. Deng, M. Doroslovacki, H. Mustafa, J. Xu, and S. Koo, "Instantaneous feature based algorithm for HF digital modulation classification," in *Proc. CISS Conf.*, 2003, John Hopkins University, Baltimore, MD, US.
- [41] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, pp. 416-429, 2000.
- [42] W. Dai, Y. Wang, and J. Wang, "Joint power and modulation classification using second- and higher statistics," in *Proc. WCNC*, 2002, pp. 155-158.
- [43] G. Hatzichristos and M. P. Fargues, "A hierarchical approach to the classification of digital modulation types in multipath environments," in *Proc. ASILOMAR*, 2001, pp. 1494-1498.
- [44] Y. Yang and S. S. Soliman, "Optimum classifier for M-ary PSK signals," in *Proc. ICC*, 1991, pp. 1693-1697.
- [45] Y. Yang and C. H. Liu, "An asymptotic optimal algorithm for modulation classification," *IEEE Comm. Letters*, vol. 2, pp. 117-119, 1998.
- [46] Y. Yang and S. S. Soliman, "A suboptimal algorithm for modulation classification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, pp. 38-45, 1997.
- [47] Y. Yang and S. S. Soliman, "Statistical moments based classifier for MPSK signals," in *Proc. GLOBECOM*, 1991, pp. 72-76.
- [48] S. S. Soliman and S. Z. Hsue, "Signal classification using statistical moments," *IEEE Trans. Commun.*, vol. 40, pp. 908-916, 1992.
- [49] Y. Yang and S. S. Soliman, "An improved moment-based algorithm for signal classification," *Signal Processing*, pp. 231-244, 1995.
- [50] L. Lichun, "Comments on signal classification using statistical moments," *IEEE Trans. Commun.*, vol. 50, pp. 195, 2002.
- [51] P. C. Sapiano, J. Martin, and R. Holbeche, "Classification of PSK signals using the DFT of phase histogram," in *Proc. ICASSP*, 1995, pp. 1868-1871.

- [52] C. Schreyogg and J. Reichert, "Modulation classification of QAM schemes using the DFT of phase histogram combined with modulus information," in *Proc. IEEE MILCOM*, 1997, pp. 1372-1376.
- [53] A. Swami, S. Barbarossa, and B. Sadler, "Blind source separation and signal classification," in *Proc. ASILOMAR*, 2000, pp. 1187-1191.
- [54] C. Martret and D. M. Boiteau, "Modulation classification by means of different order statistical moments," in *Proc. IEEE MILCOM*, 1997, pp. 1387-1391.
- [55] M. P. DeSimio and G. E. Prescott, "Adaptive generation of decision functions for classification of digitally modulated signals," in *Proc. NAECON*, 1988, pp. 1010-1014.
- [56] J. Reichert, "Automatic classification of communication signals using higher order statistics," in *Proc. ICASSP*, 1992, pp. 221-224.
- [57] C. Schreyogg, C. Kittel, U. Kressel, and J. Reichert, "Robust classification of modulation types using spectral features applied to HMM," in *Proc. IEEE MILCOM*, 1997, pp. 1377-1381.
- [58] P. Marchand, J. L. Lacoume, and C. Le Martret, "Classification of linear modulations by a combination of different orders cyclic cumulants," in *Proc. Workshop on HOS*, 1997, pp. 47-51.
- [59] P. Marchand, J. L. Lacoume, and C. Le Martret, "Multiple hypothesis classification based on cyclic cumulants of different orders," in *Proc. ICASSP*, 1998, pp. 2157-2160.
- [60] C. M. Spooner, "Classification of cochannel communication signals using cyclic cumulants," in *Proc. ASILOMAR*, 1995, pp. 531-536.
- [61] C. M. Spooner, W. A. Brown, and G. K. Yeung, "Automatic radio-frequency environment analysis," in *Proc. ASILOMAR*, 2000, pp. 1181-1186.
- [62] C. M. Spooner, "On the utility of sixth-order cyclic cumulants for RF signal classification," in *Proc. ASILOMAR*, 2001, pp. 890-897.
- [63] O. A. Dobre, Y. Bar-Ness, and W. Su, "Higher-order cyclic cumulants for high order modulation classification," in *Proc. IEEE MILCOM*, 2003, pp. 112-117.
- [64] O. A. Dobre, Y. Bar-Ness, and W. Su, "Robust QAM modulation classification algorithm based on cyclic cumulants," in *Proc. WCNC*, 2004, pp. 745-748.
- [65] O. A. Dobre, A. Abdi, Y. Bar-Ness and W. Su, "Selection combining for modulation recognition in fading channels," in *Proc. IEEE MILCOM*, 2005, pp. 1-7.
- [66] J. Venalainen, L. Terho, and V. Koivunen, "Modulation classification in fading multipath channel," in *Proc. ASILOMAR*, 2002, pp. 1890-1894.
- [67] V. J. Stolman, S. Paranjpe, and G. C. Orsak, "A blind information theoretic approach to automatic signal classification," in *Proc. IEEE MILCOM*, 1999, pp. 447-451.
- [68] B.-P. Paris, G. C. Orsak, H. Chen, and N. Warke, "Modulation classification in unknown dispersive environments," in *Proc. ICASSP*, 1997, pp. 3853-3856.
- [69] W. Wei and J. M. Mendel, "A fuzzy logic method for modulation classification in nonideal environments," *IEEE Trans. Fuzzy Syst.*, vol. 7, pp. 333-344, 1999.
- [70] J. Lopatka and M. Pedzisz, "Automatic modulation classification using statistical moments and fuzzy classifier," in *Proc. IEEE ICSP*, 2000, pp. 1500-1505.

- [71] A. O. Hero and H. H. Mahram, "Digital modulation classification using power moment matrices," in *Proc. ICASSP*, 1998, pp. 3285-3288.
- [72] H. H. Mahram and A. O. Hero, "Robust QAM modulation classification via moment matrices," in *Proc. PIMRC*, 2000, pp. 133-137.
- [73] B. G. Mobasser, "Constellation shape as a robust signature for digital modulation recognition," in *Proc. IEEE MILCOM*, 1999, pp. 442-446.
- [74] X. Huo and D. L. Donoho, "A simple and robust modulation classification method via counting," in *Proc. ICASSP*, 1998, pp. 3289-3292.
- [75] D. L. Donoho and X. Huo, "Large-sample modulation classification using Hellinger representation," in *Proc. Signal Processing Advanced in Wireless Communications*, 1997, pp. 133-137.
- [76] A. Swami and B. Sadler, "Modulation classification via hierarchical agglomerative cluster analysis," in *Proc. IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, 1997, pp. 141-144.
- [77] C. Schreyogg, "Identification of voiceband data signal constellations using a divisive cluster algorithm," in *Proc. IEEE Digital Signal Processing Workshop*, 1996, pp. 474-477.
- [78] S. Taira and E. Murakami, "Automatic classification of analogue modulation signals by statistical parameters," in *Proc. IEEE MILCOM*, 1999, pp. 202-206.
- [79] C. Louis and P. Sehier, "Automatic modulation recognition with a hierarchical neural network," in *Proc. IEEE MILCOM*, 1994, pp. 713-717.
- [80] H. Yoshioka, Y. Shirato, I. Toyoda, and M. Umehira, "A fast modulation recognition technique using nearest neighbor rules with optimized threshold for modulation classification in Rayleigh fading channels," in *Proc. IEEE Wireless Personal Multimedia Communications Conf.*, 2002, pp. 1049-1052.
- [81] C.-D. Chung and A. Polydoros, "Envelope-based classification schemes for continuous-phase binary frequency-shift-keying modulations," in *Proc. IEEE MILCOM*, 1994, pp. 796-800.
- [82] Z. Zhijin and S. Junna, "A new method for modulation types recognition based on the time frequency representation," in *Proc. ICSP*, 2002, pp. 208-211.
- [83] L. V. Dominguez, J. M. P. Borrillo, and J. P. Garcia, "A general approach to the automatic classification of radio communication signals," *Signal Processing*, vol. 22, pp. 239-250, 1991.
- [84] K. Umebayashi, S. Ishii, and R. Kohno, "Blind adaptive estimation of modulation scheme for software defined radio," in *Proc. PIMRC*, 2000, pp. 43-47.
- [85] S. L. Wood, M. J. Ready, and J. R. Treichler, "Constellation identification using the Radon transform," in *Proc. ICASSP*, 1988, pp. 1878-1881.
- [86] S. L. Wood and J. R. Treichler, "Performance of the Radon transform method for constellation identification," in *Proc. ASILOMAR*, 1988, pp. 119-123.
- [87] S. L. Wood, M. J. Ready, and J. R. Treichler, "Modem constellation identification: A performance comparison of two methods," in *Proc. ICASSP*, 1990, pp. 1651-1654.
- [88] Z. Yu, Y. Q. Shi, and W. Su, "M-ary frequency shift keying signal classification based on discrete Fourier transform," in *Proc. IEEE MILCOM*, 2003, pp. 1167-1172.
- [89] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing,"

IEEE Trans. Commun., vol. 33, pp. 665-675, 1985.

- [90] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Artech House, 2000.
- [91] D. Grimaldi, S. Rapuano and G. Truglia, "An automatic digital modulation classifier for measurements on telecommunication networks," in *Proc. IEEE Instrumentation and Measurement Technology*, 2002, pp. 957-962.
- [92] W. Su and J. Kosinski, "A survey of digital modulation recognition methods," in *Proc. International Signal Processing Conf.*, 2003, Dallas, TX, US.
- [93] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Blind modulation classification: a concept whose time has come," in *Proc. IEEE Sarnoff Symposium on Advances in Wired and Wireless Communications*, 2005, Princeton University, pp. 223-228.
- [94] J. G. Proakis, *Digital Communications*. Fourth Ed., McGraw-Hill, 2001.
- [95] H. L. Van Trees, *Detection, Estimation and Modulation Theory- Part I*. New York: Wiley, 2001.
- [96] S. M. Kay, *Fundamental and Statistical Signal Processing – Detection Theory*. Englewood Cliffs, NJ: PTR Prentice-Hall, 1998.
- [97] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University, 2003.
- [98] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory and some applications," *Proc. of IEEE*, vol. 79, pp. 278-305, 1991.
- [99] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis. A Nonlinear Signal Processing Framework*. Englewood Cliffs, NJ: PTR Prentice-Hall, 1993.
- [100] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 4-37, 2000.
- [101] S. Barbarossa and A. Scaglione, "Blind equalization using cost-functions matched to the signal constellations," in *Proc. ASILOMAR*, 1997, pp. 550-54.
- [102] V. Dandawade and G. B. Giannakis, "Asymptotic theory of mixed time averages and kth- order cyclic-moment and cumulant statistics," *IEEE Trans. Informat. Theory*, vol. 41, pp. 216-232, 1995.
- [103] W. A. Gardner and C. M. Spooner, "The cumulant theory of cyclostationary time-series, Part I: Foundation," *IEEE Trans. Signal Processing*, vol. 42, pp. 3387-3408, 1994.

Table I A SUMMARY OF LIKELIHOOD-BASED CLASSIFIERS.

Author(s)	Classifier(s)	Modulations ¹⁴	Unknown parameter(s)	Channel
Sills [1]	ALRT	BPSK, QPSK, 16QAM, V29, 32QAM, 64QAM	Carrier phase θ	AWGN
Wei and Mendel [2]	ALRT	16QAM, V29	-	AWGN
Kim and Polydoros [3], [4]	Quasi-ALRT	BPSK, QPSK	Carrier phase θ	AWGN
Huang and Polydoros [5]	Quasi-ALRT	UW, BPSK, QPSK, 8PSK, 16PSK	Carrier phase θ and timing offset ε	AWGN
Sapiano and Martin [6]	ALRT	UW, BPSK, QPSK, 8PSK	-	AWGN
Long et al. [7]	Quasi-ALRT	16PSK, 16QAM, V29	Carrier phase θ	AWGN
Hong and Ho [8]	ALRT	BPSK, QPSK	Signal level α	AWGN
Beidas and Weber [9]	ALRT Quasi-ALRT	32FSK, 64FSK	Phase jitter $\{\phi_k\}_{k=1}^K$	AWGN
Beidas and Weber [10], [12]	ALRT Quasi-ALRT	32FSK, 64FSK	Phase jitter $\{\phi_k\}_{k=1}^K$ and timing offset ε	AWGN
Panagiotu et al. [14]	GLRT HLRT	16PSK, 16QAM, V29	Carrier phase θ	AWGN
Chugg et al. [15]	HLRT	BPSK, QPSK, OQPSK	Carrier phase θ , signal power S and PSD N_0	AWGN
Hong and Ho [16]	HLRT	BPSK, QPSK	Angle of arrival ϑ	AWGN
Dobre et al. [19]	HLRT	BPSK, QPSK, 8PSK, 16PSK, 16QAM, 64QAM	Channel amplitude α and phase φ	Flat fading
Dobre and Hameed [20]	HLRT Quasi-HLRT	BPSK, QPSK, 8PSK, 16PSK	Channel amplitude α and phase φ , and noise PSD N_0	Flat fading
Abdi et al. [21]	ALRT Quasi-HLRT	16QAM, 32QAM, 64QAM	Channel amplitude α and phase φ	Flat fading
Li et al. [22]	Quasi-HLRT	4QAM, 16QAM, 64QAM	Frequency offset Δf	AWGN

¹⁴ These modulations have been used in the simulations in the original papers.

Table II THEORETICAL MOMENTS FOR SEVERAL UNIT-VARIANCE SIGNAL CONSTELLATIONS WITH EQUIPROBABLE SYMBOLS.

	BPSK	QPSK	8-PSK	16-PSK	V.29	16-QAM	64-QAM
$m_{s^{(i)},2,0}$	1	0	0	0	0	0	0
$m_{s^{(i)},2,1}$	1	1	1	1	1	1	1
$m_{s^{(i)},4,0}$	1	1	0	0	0.5158	-0.68	-0.619
$m_{s^{(i)},4,1}$	1	0	0	0	0	0	0
$m_{s^{(i)},4,2}$	1	1	1	1	-0.5816	1.32	1.38
$m_{s^{(i)},6,0}$	1	0	0	0	0	0	0
$m_{s^{(i)},6,1}$	1	1	0	0	-1.5243	-1.32	-1.298
$m_{s^{(i)},6,2}$	1	1	0	0	0	0	0
$m_{s^{(i)},6,3}$	1	1	1	1	1.4897	1.96	2.22
$m_{s^{(i)},8,0}$	1	1	1	0	-5.6304	2.2	1.91
$m_{s^{(i)},8,1}$	1	0	0	0	0	0	0
$m_{s^{(i)},8,2}$	1	1	0	0	9.4585	-2.48	-2.75
$m_{s^{(i)},8,3}$	1	0	0	0	0	0	0
$m_{s^{(i)},8,4}$	1	1	1	1	-8.5370	3.12	3.96

Table III A SUMMARY OF FEATURE BASED CLASSIFIERS.

Author(s)	Features	Modulations ¹⁴	Unknown parameters	Channel(s)
Azzouz and Nandi [28]	Maximum power spectral density of normalized centered amplitude, standard deviations of normalized centered amplitude, phase and frequency	2ASK, 4ASK, BPSK, QPSK, 2FSK, 4FSK	-	AWGN
Soliman and Hsue [32], [33]	Variance of the zero-crossing interval sequence, phase difference and zero-crossing interval histograms	UW, BPSK, QPSK, 8PSK, BFSK, 4FSK, 8FSK	-	AWGN
Soliman and Hsue [44]-[46]	PDF of phase	UW, BPSK, QPSK, 8PSK	-	AWGN
Soliman and Hsue [47]-[49]	Statistical moments of phase	UW, BPSK, QPSK, 8PSK	-	AWGN
Sapiano et al. [51]	DFT of phase PDF	UW, BPSK, QPSK, 8PSK	-	AWGN
Ho et al. [36], [37]	Variance of HWT magnitude, HWT magnitude and peak magnitude histograms	BPSK, QPSK, 8PSK, 2FSK, 4FSK, 8FSK, CP2FSK, CP4FSK, CP8FSK, MSK	-	AWGN
Hong and Ho [38]	Variance of HWT magnitude and normalized HWT magnitude	QPSK, 4FSK, 16QAM	-	AWGN
Swami and Sadler [41]	Normalized fourth-order cumulants of the received signal	BPSK, 4ASK, 16QAM, 8PSK, V32, V29, V29c	Carrier phase θ , frequency Δf and timing offsets ε	AWGN, impulsive noise, cochannel interference
Swami et al. [53]	Normalized fourth-order cumulants of the received signal and AMA cost function	BPSK, 4ASK, QPSK, 16QAM, V29, V32, 64QAM	-	Frequency selective channel
Martret and Boiteau [54]	Fourth- and second-order moments of the received signal	QPSK, 16QAM	-	AWGN
Marchand et al. [58], [59]	Fourth- and second-order cyclic cumulants of the received signal	QPSK, 16QAM, 64QAM	-	AWGN

Spooner et al. [60]-[62]	Sixth-, fourth-, and second-order cyclic cumulants of the received signal	MSK, QPSK, BPSK, 8PSK, 8QAM, QPSK, 16QAM, 64QAM, V29	Frequency offset Δf , excess bandwidth EBW, symbol period T , signal amplitude a	AWGN, cochannel interference
Dobre et al. [63]	Eighth-order cyclic cumulants of the received signal	BPSK, QPSK, 8PSK, 4ASK, 8ASK, 16QAM, 64QAM, 256QAM	-	AWGN
Dobre et al. [64]	Eighth-, sixth-, and fourth-order cyclic cumulants of the received signal	4QAM, 16QAM	Carrier phase θ , phase jitter $\{\phi_k\}_{k=1}^K$, frequency offset Δf	AWGN, impulsive noise
Dobre et al. [65]	Eighth-order cyclic cumulants of the signal at the output of a selection combiner	4ASK, 8ASK, BPSK, QPSK, 16QAM, 32QAM, 64QAM	-	Rayleigh and Ricean fading channels
Yu et al. [88]	DFT of the received signal	2FSK, 4FSK, 8FSK, 16FSK, 32FSK	-	AWGN

Table IV THEORETICAL CUMULANTS FOR SEVERAL UNIT-VARIANCE SIGNAL CONSTELLATIONS WITH EQUIPROBABLE SYMBOLS.

	BPSK	QPSK	8-PSK	16-PSK	V.29	16-QAM	64-QAM
$c_{s^{(i)},2,0}$	1	0	0	0	0	0	0
$c_{s^{(i)},2,1}$	1	1	1	1	1	1	1
$c_{s^{(i)},4,0}$	-2	1	0	0	0.5158	-0.68	-0.619
$c_{s^{(i)},4,1}$	-2	0	0	0	0	0	0
$c_{s^{(i)},4,2}$	-2	-1	-1	-1	1.4148	-0.68	-0.619
$c_{s^{(i)},6,0}$	16	0	0	0	0	0	0
$c_{s^{(i)},6,1}$	16	-4	0	0	1.0683	2.08	1.7972
$c_{s^{(i)},6,2}$	16	0	0	0	0	0	0
$c_{s^{(i)},6,3}$	16	4	4	4	2.2551	2.08	1.7972
$c_{s^{(i)},8,0}$	-272	-34	1	0	3.7797	-13.9808	-11.5022
$c_{s^{(i)},8,1}$	-272	0	0	0	0	0	0
$c_{s^{(i)},8,2}$	-272	34	0	0	2.1993	-13.9808	-11.5022
$c_{s^{(i)},8,3}$	-272	0	0	0	0	0	0
$c_{s^{(i)},8,4}$	-272	-34	-33	-33	3.7797	-13.9808	-11.5022

Table V PERFORMANCE AND DETAILS OF SEVERAL CLASSIFIERS FOR BPSK VERSUS QPSK.

	Classifier	Unknown parameter	Model mismatch ¹⁵	SNR (dB)	P_{cc}	K
1	ALRT, $L = 1$ (10) and (8), with $\eta_A = 1$	-	-	-3	0.975	100
2 ¹⁶	ALRT, $L = 2$ (12) ³ and (8), with $\eta_A = 1$	-	-	-6	0.975	100
3 ¹⁶	ALRT, $L = 1$ (10) and (8), with $\eta_A = 1$	-	Carrier phase θ , $U[-\pi, \pi)$	-3	0.63	100
4	Quasi-ALRT, (20) and (22), with $M = 2$	Carrier phase θ	-	-2	0.96	100
5 ¹⁶	HLRT, (27) ¹⁷ and (8), with $\eta_H = 1$	Carrier phase θ	-	-2	0.968	100
6	Cumulant-based, (32) with $\hat{\omega} = \left \hat{c}_{r,4,0}(\mathbf{0}_3) / \hat{c}_{r,2,1}^2(\mathbf{0}) \right $, $N_{\text{mod}} = 2$, $\mu_{H_1} = 1$, $\mu_{H_2} = 2$	Carrier phase θ	-	4	0.96	100
7	Quasi-ALRT, (20) and (22), with $M = 2$	Carrier phase θ	Timing offset $\varepsilon = 0.15$	11	0.96	100
8	Cumulant-based (see the sixth row in this table)	Carrier phase θ	Timing offset $\varepsilon = 0.15$	6	0.96	100

¹⁵ These are not included when designing the classifier, and their impacts on the performance were studied to evaluate the classifier robustness.

¹⁶ These numerical results are not given in the papers where the classifiers are proposed, and we have simulated them to perform a comparison with other methods.

¹⁷ The maximization procedure w.r.t. the carrier phase θ was carried out by discrete search with 1 degree resolution. Due to the symmetry of the investigated constellations, the maximization procedure can be performed over $[0, \pi/2]$, instead of $[0, 2\pi]$ [14], [19], [21].

Table VI PERFORMANCE AND DETAILS OF SEVERAL CLASSIFIERS FOR 16QAM VERSUS V.29.

	Classifier	Unknown parameter	Model mismatch ¹⁵	SNR (dB)	P_{cc}	K
1 ¹⁶	ALRT (10) and (8), with $\eta_A = 1$	-	-	7	0.99	100
2 ¹⁶	ALRT (10) and (8), with $\eta_A = 1$	-	Carrier phase θ , $U[-\pi, \pi)$	7	0.50	100
3	Cumulant-based, (32) with $\hat{\omega} = \hat{c}_{r,4,0}(\mathbf{0}_3) / \hat{c}_{r,2,1}^2(0)$, $N_{\text{mod}} = 2$, $\mu_{H_0} = -0.68$, $\mu_{H_1} = 0.5185$	-	-	9	0.99	100
4 ¹⁶	Cumulant-based ¹⁶ , (see the third row of this table)	-	Carrier phase θ , $U[-\pi, \pi)$	9	0.50	100
5	Quasi-ALRT, (23), with $A = 0.0135$ and $B = -0.0246$, and $\eta_{QA,E}$ ¹²	Carrier phase θ	-	30	0.88	100
6	HLRT (27) ¹⁷ and (8), with $\eta_{H,E}$ ¹²	Carrier phase θ	-	9	0.99	100
7 ¹⁶	Quasi-HLRT (30) ¹³ , (29) and threshold set to one	Carrier phase θ	-	19	0.99	100
				11	0.91	6000
8	CC-based (35) and (36)	Carrier phase θ and timing error ε	-	9	0.99	6000

Table VII PERFORMANCE AND DETAILS OF SEVERAL CLASSIFIERS FOR 32FSK VERSUS 64FSK.

	Algorithm	Unknown parameter	Model mismatch ¹⁵	SNR (dB)	P_{cc}	K (32FSK)
1	ALRT, (14) with $ s_k^{(i)} =1$, and (8) with $\eta_d=1$	Phase jitter $\{\phi_k\}_{k=1}^K$	-	6	0.975	10
2	Quasi-ALRT (first-order correlation), (25) with the first three terms, calculated based on $K (K/2)$ symbols for 32FSK (64FSK) and $\eta_{HOC,E}^{12}$ as the threshold	Phase jitter $\{\phi_k\}_{k=1}^K$	-	6	0.74	10
3	Quasi-ALRT (second-order correlation), (25) with the first five terms, expressed for $K (K/2)$ symbols for 32FSK (64FSK) and $\eta_{HOC,E}^{12}$ as the threshold	Phase jitter $\{\phi_k\}_{k=1}^K$	-	6	0.90	10
4	ALRT (see the first row of this table)	Phase jitter $\{\phi_k\}_{k=1}^K$	Frequency offset ($\Delta f = f_d/2$)	6	0.57	10

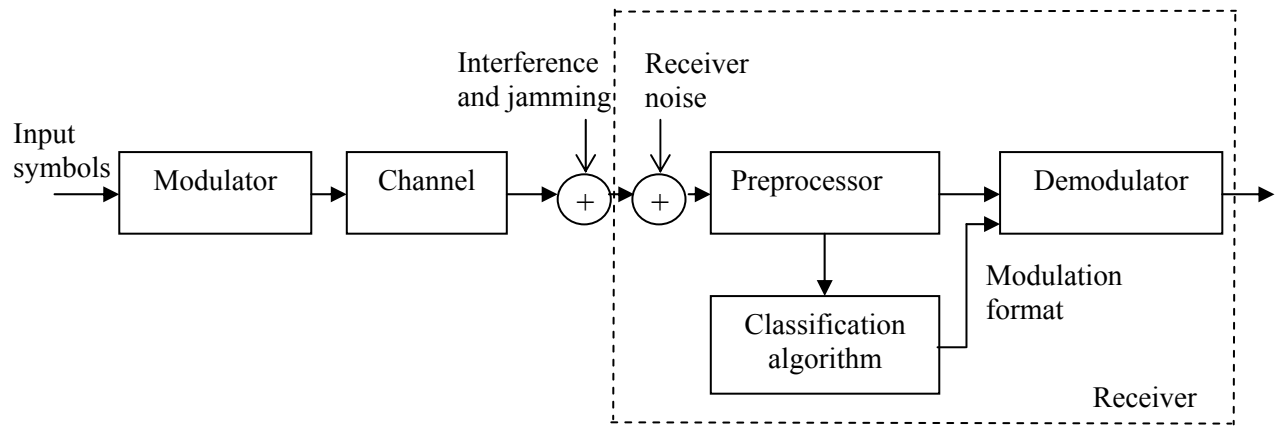


Fig. 1. System block diagram.

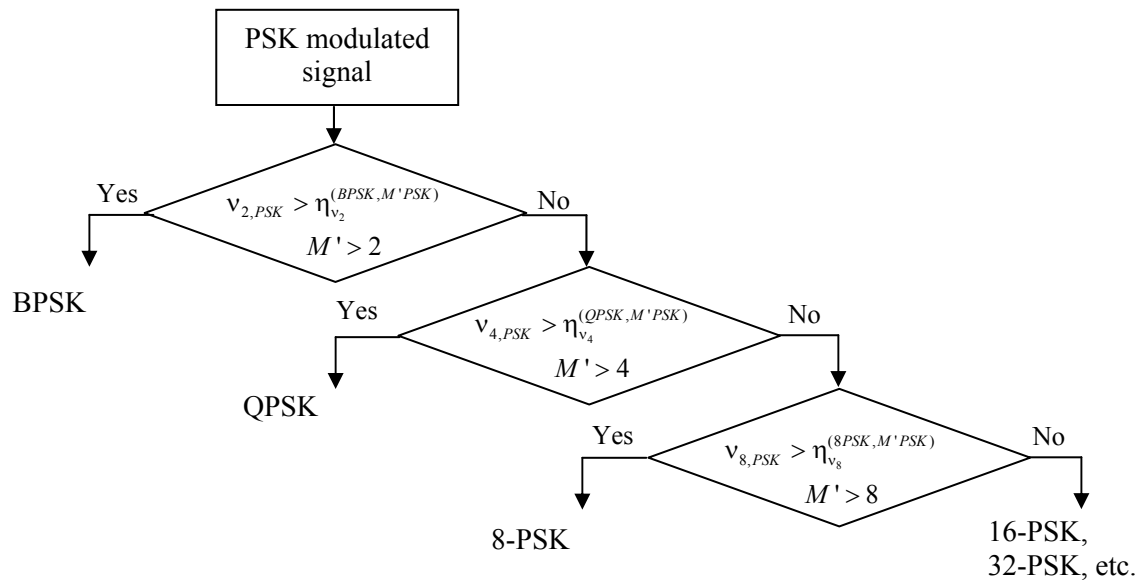


Fig. 2. Binary decision tree for PSK signals using the quasi-ALRT classifier.