MODULATION CLASSIFICATION IN FADING CHANNELS USING ANTENNA ARRAYS

Ali Abdi[1], Octavia A. Dobre[1], Rahul Choudhry[1], Yeheskel Bar-Ness[1], and Wei Su[2]

[1]CCSPR, Dept. of ECE, New Jersey Institute of Technology, Newark, NJ 07102, USA
[2]CECOM, Fort Monmouth, NJ 07703, USA

ABSTRACT

Blind modulation classification (MC) is an intermediate step between signal detection and demodulation, and plays a key role in various civilian and military applications In this paper, first we provide an overview of decision-theoretic MC approaches. Then we derive the average likelihood ratio (ALR) based classifier for linear and nonlinear modulations, in noisy channels with unknown carrier phase offset and also in Rayleigh fading channels. Since these ALR-based classifiers are complex to implement, we then develop a quasi hybrid likelihood ratio (QHLR) based classifier, where the unknown parameters are estimated using low-complexity techniques. This QHLR-based classifier is much simpler to implement and is also applicable to any fading distribution, including Rayleigh and Rice. Afterwards, we propose a generic multi-antenna classifier for linear and nonlinear modulations, using an antenna array at the receiver. This classifier has the potential to improve the performance of traditional single-antenna classifiers, including the proposed QHLR-based algorithm, via spatial diversity. Simulation results are provided to show the performance enhancement offered by the new QHLR-based multi-antenna classifier, in a variety of channel and fading conditions.

I. INTRODUCTION

Modulation identification of a received signal is of importance in a variety of military and commercial applications. Some examples include spectrum monitoring and management, surveillance and control of broadcasting activities, adaptive transmission schemes, and electronic warfare. Our comprehensive literature survey [3] shows that although this topic has been extensively studied, less attention has been paid to modulation classification in fading channels, and also the utilization of the spatial diversity, provided by antenna arrays, in such channels. In this paper, we study these two topics in a systematic way.

II. SIGNAL, NOISE, AND CHANNEL MODELS

Let $s(t;\mathbf{u}_{i,0})$ represent the noise-free baseband complex envelope of the received signal, coming from the i-th modulation format, $i=1,2,...,N_{\text{mod}}$, where N_{mod} is the number of candidates modulations that we are looking at. The vector $\mathbf{u}_{i,0}$, which corresponds to the i-th modulation format, denotes the vector of *unknown* quantities at the receiver. In this paper we consider a frequency-flat slowly-varying multipath fading channel with $\mathbf{u}_{i,0}=[\alpha_0\ \ \varphi_0\ \ \{s_k^{(i)}\}_{k=1}^N]^\dagger$, [1] where α_0 is the channel amplitude, φ_0 is the channel phase (which also includes the carrier phase offset), $\{s_k^{(i)}\}_{k=1}^N$ represents the N complex transmitted data symbols, taken from the i-th finite-alphabet modulation format, and † is the transpose operator. Depending on the classification approach, in Section IV we consider α_0 and φ_0 as random variables, denoted by α and φ, respectively, and average over them, whereas in Section V, α_0 and φ_0 are considered as unknown deterministic quantities, for which $\hat{\alpha}_0$ and $\hat{\varphi}_0$ are provided as their estimates. In any case, we treat the data symbols $\{s_k^{(i)}\}_{k=1}^N$ as independent and identically distributed (iid) randoms, and average over them. With symbol period T_0, rectangular unit-amplitude pulse shape $u_{T_0}(t)$ of length T_0, and $j^2=-1$, $s(t;\mathbf{u}_{i,0})$ is given by

$$s(t;\mathbf{u}_{i,0})=\alpha_0 e^{j\varphi_0}\sum_{k=1}^{N} s_k^{(i)}(t)u_{T_0}(t-(k-1)T_0),\ \ 0\le t\le NT_0\,. \quad (1)$$

Eq. (1) is applicable to M-ary ASK, PSK, QAM, and FSK modulations. In MASK we have $s^{(\text{MASK})}\in\{\pm1,\pm3,...,\pm(M-1)\}$, where M is even. The symbols of MPSK are given by $s^{(\text{MPSK})}=e^{j\theta_m}$, in which $\theta_m=2\pi m/M$, $m=0,1,...,M-1$, with M as a power of two. For square and cross QAM symbols, see [1]. In MFSK, the symbols are given by $s^{(\text{MFSK})}(t)=e^{j2\pi f_m t}$, such that $f_m=\pm f_{d0}, \pm 3f_{d0}, ..., \pm(M-1)f_{d0}$, with f_{d0} as the frequency spacing between any two adjacent constellation points and M as a power of two. As mentioned in the footnote of this page, for linear modulations, the data symbols do not depend on t. This property simplifies the general classifier derived in Section IV.

Let us define the variance of the zero-mean i-th constellation as $\vartheta^{(i)}=E[|s_k^{(i)}|^2]$, $i=1,2,...,N_{\text{mod}}$, where $E[.]$ denotes the mathematical expectation. Then the signal power, defined by

$$S_0^{(i)}=(NT_0)^{-1}\int_0^{NT_0}E[|s(t;\mathbf{u}_{i,0})|^2]dt\,, \quad (2)$$

can be shown to be $S_0^{(i)}=\alpha_0^2\vartheta^{(i)}$, after substituting (1) into (2). For equiprobable M_i constellation points of the i-th modulation, obviously one has $\vartheta^{(i)}=M_i^{-1}\sum_{m=1}^{M_i}|s_m^{(i)}|^2$. Note that for PSK and FSK $|s_m^{(i)}|^2=1$, $m=1,2,...,M_i$, whereas in ASK and QAM, $|s_m^{(i)}|^2$ takes different values.

In the presence of noise, for the baseband received complex envelope we have

$$r(t)=s(t;\mathbf{u}_{i,0})+n(t),\ \ 0\le t\le NT_0\,, \quad (3)$$

where $n(t)$ is the complex additive white Gaussian noise (AWGN) with two-sided power spectral density N_0 (W/Hz), and the correlation $E[n(t)n^*(t+\tau)]=N_0\delta(\tau)$, such that * is the complex conjugate and $\delta(.)$ is Dirac delta.

[1] As we discuss in the first paragraph after eq. (1), only for frequency shift keying (FSK), the data symbols depend on t. For others such as amplitude shift keying (ASK), phase shift keying (PSK), and quadrature amplitude modulation (QAM), $s_k^{(i)}$'s do not depend upon t. To simplify the notation, we most often drop this t-dependence, unless otherwise mentioned.

III. DECISION THEORY AND LIKELIHOOD

To design the modulation classifier, in order to determine what modulation has been received, out of N_{mod} equally likely candidates, we take the decision-theoretic approach, which is optimal under certain conditions [2]. This approach, which works based upon the maximum likelihood (ML) principle, requires the likelihood function of $r(t)$ over the interval $0 \le t \le NT_0$. Using the complex Gaussian distribution of $n(t)$ and for the i-th hypothesis H_i (the i-th modulation format), the conditional likelihood function of $r(t)$, conditioned on the unknown vector $\mathbf{u}_{i,0}$, can be shown to be [2] [5]

$$\Xi[r(t) \mid \mathbf{u}_{i,0}, H_i] =$$
$$\exp\left\{\frac{2}{N_0}\text{Re}\left[\int_0^{NT_0} r(t)s^*(t;\mathbf{u}_{i,0})dt\right] - \frac{1}{N_0}\int_0^{NT_0}\left|s(t;\mathbf{u}_{i,0})\right|^2 dt\right\}, \quad (4)$$

where Re[.] gives the real part. To derive the likelihood function of $r(t)$ from (4), $\Xi^{(i)}[r(t)]$, three techniques are proposed in the literature, that we discuss in the sequel. Once $\Xi^{(i)}[r(t)]$ is calculated for all the possible N_{mod} candidate modulations, based on the observed $r(t)$ over $0 \le t \le NT_0$, one can make the decision according to

Choose $\bar{i}$ as the received modulation

$$\text{if } \bar{i} = \arg\max_{1 \le i \le N_{\text{mod}}} \Xi^{(i)}[r(t)]. \quad (5)$$

Obviously our modulation classification is a multiple *composite* hypothesis testing problem, due to the unknown data symbols $\{s_k^{(i)}\}$, as well as the unknown parameters, which are α_0 and φ_0 in this paper. Based on our comprehensive literature survey [3] [4], three methods are proposed so far, to handle the unknown quantities: average likelihood ratio test (ALRT) [6]-[18], generalized likelihood ratio test (GLRT) [19], [20], and hybrid likelihood ratio test (HLRT) [19], [21]-[25].

The unknown quantities in ALRT are considered as random variables, with a certain joint probability density function (PDF), $p(\mathbf{u}_i \mid H_i)$, and the likelihood function is derived by averaging the conditional likelihood function with respect to it

$$\Xi_A^{(i)}[r(t)] = \int \Xi[r(t) \mid \mathbf{u}_i, H_i] p(\mathbf{u}_i \mid H_i) d\mathbf{u}_i \,. \quad (6)$$

If the chosen $p(\mathbf{u}_i \mid H_i)$ is the same as the true PDF, then ALRT is the optimal classifier, i.e., maximizes the probability of correct classification. Otherwise, the optimality is not guaranteed. Furthermore, the multivariate integration in (6) is mathematically tractable only for few cases, and normally one has to resort to approximations [8]-[12].

In GLRT, on the other hand, the unknown quantities are treated as unknown deterministics, and the likelihood function is obtained by replacing the unknown quantities in the conditional likelihood function, with their ML estimates

$$\Xi_G^{(i)}[r(t)] = \Xi[r(t) \mid \hat{\mathbf{u}}_{i,0}^{ML}, H_i], \quad (7)$$

where the ML estimate of $\mathbf{u}_{i,0}$ is obtained by

$$\hat{\mathbf{u}}_{i,0}^{ML} = \arg\max_{\mathbf{u}_i} \Xi[r(t) \mid \mathbf{u}_i, H_i]. \quad (8)$$

GLRT is a reasonable alternative to ALRT. However, the ML estimator is (6) usually does not have a closed from, and one needs to carry out a multidimensional exhaustive search, to find the ML estimates [18], for each hypothesis. Moreover, for nested constellations such as BPSK/QPSK, and 16-QAM/64-QAM, the likelihood function in GLRT can take the same numerical values, which in turn leads to incorrect classification [18] [26].

HLRT is a combination of ALRT and GLRT, in an attempt to avoid the disadvantages of both techniques, while employing their useful properties. In HLRT, the unknown data symbols are considered as random variables and are averaged out, whereas the unknown parameters are treated as deterministic unknowns, eventually replaced by their ML estimates. To write the likelihood function, we break $\mathbf{u}_{i,0}$ to two vectors $\mathbf{w}_0$ and $\mathbf{s}_i$ such that $\mathbf{u}_{i,0} = [\mathbf{w}_0{}^\dagger \mathbf{s}_i{}^\dagger]^\dagger$. The vector $\mathbf{w}_0$ contains the unknown parameters, whereas $\mathbf{s}_i = [s_1 \, s_2 \ldots s_N]^\dagger$ represents the unknown data symbols. With this notation, the likelihood function, conditioned on $\mathbf{w}_0$, is

$$\Xi^{(i)}[r(t) \mid \mathbf{w}_0] = \int \Xi[r(t) \mid \mathbf{w}_0, \mathbf{s}_i, H_i] p(\mathbf{s}_i \mid H_i) d\mathbf{s}_i \,. \quad (9)$$

With the ML estimate of $\mathbf{w}_0$ for each candidate modulation

$$\hat{\mathbf{w}}_{i,0}^{ML} = \arg\max_{\mathbf{w}} \Xi^{(i)}[r(t) \mid \mathbf{w}], \quad (10)$$

the HLRT likelihood function is eventually given by

$$\Xi_H^{(i)}[r(t)] = \Xi^{(i)}[r(t) \mid \hat{\mathbf{w}}_{i,0}^{ML}]. \quad (11)$$

Averaging over the data symbols in HLRT removes the nested constellations problem of GLRT. However, finding the ML estimates of the unknown parameters in HLRT still entails an exhaustive search, which makes its implementation complex.

In this paper, first we derive, in Section IV, closed-form expressions for the ALRT-based likelihood function in AWGN and Rayleigh fading channels, for any kind of memoryless modulation, including ASK, PSK, QAM, FSK, etc. Due to the exponential complexity of these ALRT-based classifiers, as well as not being applicable to other types of fading such as Rice, Weibull, Nakagami [1], we then propose a quasi HLRT (QHLRT) approach in Section V. In QHLRT, we use non-ML parameter estimators, which are simple yet accurate enough to provide a good classification performance, with much less computational complexity, also applicable to any fading distribution. Then in Section VI we introduce a multi-antenna classifier, which significantly improves the performance of classifiers, including the new QHLRT of Section V, with a moderate increase in the hardware complexity.

IV. AVERAGE LIKELIHOOD APPROACH

By inserting (1) into (4), the conditional likelihood function can be written as

$$\Xi[r(t) \mid \mathbf{u}_{i,0}, H_i] = \exp\left\{\frac{2\alpha_0}{N_0}\text{Re}\left[e^{-j\varphi_0}\mathsf{R}_N^{(i)}\right] - \frac{\alpha_0^2 T_0}{N_0}\eta_N^{(i)}\right\}, \quad (12)$$

where

$$\mathsf{R}_N^{(i)} = \sum_{k=1}^{N} R_k^{(i)}, \quad (13)$$

such that

$$R_k^{(i)} = \int_{(k-1)T_0}^{kT_0} r(t) s_k^{(i)^*}(t) dt, \; k = 1, \ldots, N, \; i = 1, \ldots, N_{\text{mod}}, \quad (14)$$

$$\eta_N^{(i)} = \sum_{k=1}^{N} |\, s_k^{(i)} \,|^2 \,. \quad (15)$$

Notice that for linear modulations, since $s_k^{(i)}$ is independent of t, (14) simplifies to $R_k^{(i)} = s_k^{(i)^*} r_k$, such that

$$r_k = \int_{(k-1)T_0}^{kT_0} r(t)dt \,. \tag{16}$$

Both $R_k^{(i)}$ and r_k can be regarded as the output of some matched filters at the receiver, sampled at $t = kT_0,\ k = 1, ..., N$. Also note that $\eta_N^{(i)}$ in (15), when divided by N, represents an estimate of $\vartheta^{(i)}$, the constellation variance of the i-th modulation format.

Now we write the conditional likelihood function of (12) in a convenient product form

$$\Xi[r(t) \mid \mathbf{u}_{i,0}, H_i] = \prod_{k=1}^{N} \exp\left\{ \frac{2\alpha_0}{N_0} \mathrm{Re}[e^{-j\varphi_0} R_k^{(i)}] - \frac{\alpha_0^2 T_0}{N_0} \mid s_k^{(i)} \mid^2 \right\}. \tag{17}$$

In the sequel, we consider three cases and derive the associated likelihood function, in compact forms, by averaging (17) with respect to the random quantities. In all the situations, data symbols will be treated as iid random variables and will be averaged out.

A. AWGN with no Unknown Parameters

This is an optimistic case where there is no unknown quantity in (1), except for the data symbols, and AWGN in (3) is the only source of uncertainty in modulation classification. Hence, it can serve as a benchmark, i.e., the best performance that one can expect.

With both α_0 and φ_0 known in (1), obviously $\mathbf{u}_{i,0} = \mathbf{u}_i = [\{s_k^{(i)}\}_{k=1}^N]^\dagger$. Due to the independence of the data symbols, averaging of (17) with respect to $\{s_k^{(i)}\}_{k=1}^N$ results in the following unconditional likelihood function

$$\Xi_{A-AWGN}^{(i)}[r(t)] = \prod_{k=1}^{N} E_{s_k^{(i)}}\left[\exp\left\{ \frac{2\alpha_0}{N_0} \mathrm{Re}[e^{-j\varphi_0} R_k^{(i)}] - \frac{\alpha_0^2 T_0}{N_0} \mid s_k^{(i)} \mid^2 \right\} \right]. \tag{18}$$

Note that $E_{s_k^{(i)}}[.]$ in (18) is nothing but a finite summation over all the M_i possible alphabets of the i-th modulation, divided by M_i, for the k-th interval. So, the implementation complexity of (18) for the i-th modulation may be considered as $O(M_i N)$, where $O(.)$ denotes the order. When compared to the other two classifiers derived subsequently in this section, implementation of the classifier built upon (18) is less complex.

B. AWGN with Unknown Carrier Phase Offset

Now we consider a less optimistic case, where α_0 in (1) is known but φ_0, regarded as the carrier phase offset (CPO), is not known. We change the notation from φ_0 to φ, to emphasize its random nature in the ALRT approach. Assuming a uniform distribution for φ over $[-\pi, \pi)$, different approximations to the likelihood function are proposed in the literature, for a variety of modulation schemes. In what follows, we derive a new closed-form expression for the likelihood function, which applies to ASK, PSK, QAM, and FSK.

Clearly $\mathbf{u}_{i,0} = \mathbf{u}_i = [\varphi\ \{s_k^{(i)}\}_{k=1}^N]^\dagger$. Let $R_{k,I}^{(i)} / R_{k,Q}^{(i)}$ be the real/imaginary parts of $R_k^{(i)}$, respectively. Also let $\mathsf{R}_N^{(i)} = \mathsf{R}_{N,I}^{(i)} + j\mathsf{R}_{N,Q}^{(i)}$. Now we rewrite (12) as

$$\Xi[r(t) \mid \mathbf{u}_i, H_i] = \exp\left\{ \frac{2\alpha_0}{N_0}\left[\mathsf{R}_{N,I}^{(i)} \cos\varphi + \mathsf{R}_{N,Q}^{(i)} \sin\varphi \right] - \frac{\alpha_0^2 T_0}{N_0} \eta_N^{(i)} \right\}. \tag{19}$$

Using eq. 3.338-4 [27], averaging of (19) with respect to φ yields

$$\Xi[r(t) \mid \{s_k^{(i)}\}_{k=1}^N, H_i] = I_0\left(\frac{2\alpha_0}{N_0} \mid \mathsf{R}_N^{(i)} \mid \right) \exp\left\{ -\frac{\alpha_0^2 T_0}{N_0} \eta_N^{(i)} \right\}, \tag{20}$$

where $I_0(.)$ is the zero order modified Bessel function of the first kind. Further averaging of (20) with respect to $\{s_k^{(i)}\}_{k=1}^N$ yields the required unconditional likelihood function

$$\Xi_{A-CPO}^{(i)}[r(t)] = E_{\{s_k^{(i)}\}_{k=1}^N}\left[I_0\left(\frac{2\alpha_0}{N_0} \mid \mathsf{R}_N^{(i)} \mid \right) \exp\left\{ -\frac{\alpha_0^2 T_0}{N_0} \eta_N^{(i)} \right\} \right]. \tag{21}$$

For the i-th modulation, note that the number of terms required to calculate $E_{\{s_k^{(i)}\}_{k=1}^N}[.]$ in (21) is M_i^N. Therefore, one may consider the exponential complexity of $O(M_i^N)$ for (21). Obviously the classifier which implements (21) is much more complex than (18).

C. Rayleigh Fading

In this more realistic case, both α_0 and φ_0 in (1) are unknown. Since in the ALRT approach, unknown quantities are visualized as random variables, we replace both with α and φ, respectively, to reflect their randomness. In Rayleigh fading channels, $\alpha e^{j\varphi}$ is a complex Gaussian variable and α and φ are independent. The Rayleigh PDF of α is given by $p(\alpha) = (2\alpha / \Omega_0) e^{-\alpha^2/\Omega_0}$, $\alpha \geq 0$, with $\Omega_0 = E[\alpha^2]$ as the average fading power, whereas for φ we have $p(\varphi) = 1/(2\pi)$ over $[-\pi, \pi)$. Now we derive a new compact form for the likelihood function, applicable to both linear and nonlinear modulations.

For the vector of unknown parameters and data symbols, now we have $\mathbf{u}_{i,0} = \mathbf{u}_i = [\alpha\ \varphi\ \{s_k^{(i)}\}_{k=1}^N]^\dagger$. Notice that the average fading power Ω_0 is not included in the vector, as it is assumed to be known. Otherwise, we have to assign a pdf to it, for example, lognormal or gamma [28], and average over it. This makes it very difficult, if not impossible, to derive an expression for the likelihood function. Using the following integral identity

$$\int_0^\infty x \exp\left(-c_1 x^2\right) I_0\left(c_2 x\right) dx = \frac{1}{2c_1} \exp\left(\frac{c_2^2}{4c_1} \right), \tag{22}$$

where c_1 and c_2 are positive constants, averaging of (20) first with respect to α, the random representative of α_0, and then over the data symbols results in the unconditional likelihood function of interest

$$\Xi_{A-Rayleigh}^{(i)}[r(t)] = E_{\{s_k^{(i)}\}_{k=1}^N}\left[\left(1 + \frac{\Omega_0 T_0\, \eta_N^{(i)}}{N_0} \right)^{-1} \times \exp\left\{ \left(1 + \frac{\Omega_0 T_0\, \eta_N^{(i)}}{N_0} \right)^{-1} \frac{\Omega_0 \mid \mathsf{R}_N^{(i)} \mid^2}{N_0^2} \right\} \right]. \tag{23}$$

Note that (23) for Rayleigh fading is simpler than (21) for AWGN with the unknown CPO, as it does not have the Bessel function. However, still it suffers from the same exponential implementation complexity, due to the averaging over the data symbols $\{s_k^{(i)}\}_{k=1}^N$.

V. QUASI HYBRID LIKELIHOOD APPROACH

As we discussed in Section III, HLRT is a reasonable alternative to ALRT and GLRT, yet still has a high computational complexity, to calculate the ML estimates of unknown parameters. Here we propose the QHLRT approach, in which we average over the data symbols, whereas for the unknown parameters, we use simple non-

ML estimators, which are accurate enough. Method-of-moment (MOM) estimators are attractive candidates due to their simplicity, whereas there is a possibility of getting near-ML performance, depending on the estimation problem at hand [29]. In the rest of this section, we concentrate on linear modulations, and QAM in particular, and consider MOM estimators for the amplitude and phase.

To estimate α_0 in (1), for linear modulations we use the output of the matched filter r_k in (16). After substituting (1) into (3) and then the result into (16), we get

$$r_k = \alpha_0 e^{j\varphi_0} s_k^{(i)} T_0 + n_k,\ \ k = 1,2,\ldots,N\ , \tag{24}$$

where

$$n_k = \int_{(k-1)T_0}^{kT_0} n(t)dt\,. \tag{25}$$

Based on (24) and the independence of signal and noise, one can show that

$$\mathsf{M}_2 = \alpha_0^2 \vartheta^{(i)} T_0^2 + N_0 T_0\ , \tag{26}$$

$$\mathsf{M}_4 = \alpha_0^4 E[|\ s_k^{(i)}\ |^4] T_0^4 + 4\alpha_0^2 \vartheta^{(i)} N_0 T_0^3 + 2N_0^2 T_0^2\ , \tag{27}$$

where $\mathsf{M}_2 = E[|\ r_k\ |^2]$ and $\mathsf{M}_4 = E[|\ r_k\ |^4]$ are the second and fourth absolute moment of r_k. By calculating N_0T_0 from (26) and substituting it into (27), eventually we get a result similar to [30], which provides the basis for estimating α_0

$$\alpha_0^4 \vartheta^{(i)^2} T_0^4 = (2\mathsf{M}_2^2 - \mathsf{M}_4)[2 - (E[|\ s_k^{(i)}\ |^4] / \vartheta^{(i)^2})]^{-1}\,. \tag{28}$$

For 16QAM, 32QAM, 64QAM, and 128QAM, one can show that $E[|\ s_k^{(i)}\ |^4] / \vartheta^{(i)^2} = 1.32,\ 1.31,\ 1.38,\ \text{and}\ 1.35$, respectively. Since these are close enough, we use the approximation $E[|\ s_k^{(i)}\ |^4] / \vartheta^{(i)^2} \approx 1.35$ for all of them. This simplifies the estimator for the right-hand side of (28) to $(2\hat{\mathsf{M}}_2^2 - \hat{\mathsf{M}}_4)/0.65$, which is independent of the modulation type. Note that $\hat{\mathsf{M}}_2 = N^{-1}\sum_{k=1}^N |\ r_k\ |^2$ and $\hat{\mathsf{M}}_4 = N^{-1}\sum_{k=1}^N |\ r_k\ |^4$.

To estimate φ_0 in (1) for QAM, we employ this property of QAM at the output of the matched filter [5]

$$E[r_k^4] = Ce^{j4\varphi_0},\ \ k = 1,2,\ldots,N\ , \tag{29}$$

where C is a positive constant, independent of the number of constellation points. This suggests the following estimator

$$\hat{\varphi}_0 = \frac{1}{4}\,\text{angle}\left(\sum\nolimits_{k=1}^N r_k^4\right). \tag{30}$$

Simulation results for the proposed QHLRT QAM classifier are provided in Section VII.

VI. MULTI-ANTENNA CLASSIFIER AND SPATIAL DIVERSITY

The spatial diversity offered by an antenna arrays is an effective tool in wireless communications [1]. In this section we develop a generic decision-theoretic multi-antenna modulation classifier, by deriving the associated likelihood function, which is applicable to both linear and nonlinear modulations of any order, in multipath fading channels. In [24], the utility of an antenna array for classification of BPSK versus QPSK in an AWGN channel is studied.

Consider an L branch receive antenna array, such that the model given in (3) holds for each branch

$$r_\ell(t) = s_\ell(t;\mathbf{u}_{i,0}) + n_\ell(t),\ \ 0 \le t \le NT_0,\ \ \ell = 1,2,\ldots,L\ , \tag{31}$$

where similar to (1)

$$s_\ell(t;\mathbf{u}_{i,0}) = \alpha_{0,\ell} e^{j\varphi_{0,\ell}} \sum_{k=1}^N s_k^{(i)}(t) u_{T_0}(t-(k-1)T_0),\ \ 0 \le t \le NT_0\,. \tag{32}$$

Furthermore, for $\ell = 1,2,\ldots,L$, $n_\ell(t)$'s, are independent complex AWGNs, with the same two-sided power spectral density N_0. Let $\mathbf{r}(t) = [r_1(t)\ r_2(t)\ \ldots\ r_L(t)]^\dagger$. Then conditioned on the unknown vector $\mathbf{u}_{i,0} = [\{\alpha_{0,\ell}\}_{\ell=1}^L\ \{\varphi_{0,\ell}\}_{\ell=1}^L\ \{s_k^{(i)}\}_{k=1}^N]^\dagger$, the conditional likelihood function of $\mathbf{r}(t)$, based on (4), can be written as

$$\Xi[\mathbf{r}(t)\,|\,\mathbf{u}_{i,0},H_i] = \prod_{\ell=1}^L \exp\left\{\frac{2}{N_0}\text{Re}\left[\int_0^{NT_0} r_\ell(t) s_\ell^*(t;\mathbf{u}_{i,0})dt\right] - \frac{1}{N_0}\int_0^{NT_0} \left|s_\ell(t;\mathbf{u}_{i,0})\right|^2 dt\right\}. \tag{33}$$

Now let us define (13) and (14) for the ℓ-th branch, i.e.

$$\mathsf{R}_{N,\ell}^{(i)} = \sum\nolimits_{k=1}^N R_{k,\ell}^{(i)},\ \ i = 1,\ldots,N_{\text{mod}}\ , \tag{34}$$

$$R_{k,\ell}^{(i)} = \int_{(k-1)T_0}^{kT_0} r_\ell(t) s_k^{(i)^*}(t)dt,\ \ k = 1,\ldots,N,\ \ i = 1,\ldots,N_{\text{mod}}\,. \tag{35}$$

Therefore, (33) can be written as

$$\Xi[\mathbf{r}(t)\,|\,\mathbf{u}_{i,0},H_i] = \exp\left\{\frac{2}{N_0}\text{Re}\left[\sum\nolimits_{\ell=1}^L \alpha_{0,\ell} e^{-j\varphi_{0,\ell}} \mathsf{R}_{N,\ell}^{(i)}\right] - \frac{T_0}{N_0}\left(\sum\nolimits_{\ell=1}^L \alpha_{0,\ell}^2\right)\eta_N^{(i)}\right\}. \tag{36}$$

Clearly, (36) reduces to (12) for $L = 1$. We also rewrite the above conditional likelihood function in a convenient product form

$$\Xi[\mathbf{r}(t)\,|\,\mathbf{u}_{i,0},H_i] = \prod_{k=1}^N \exp\left\{\frac{2}{N_0}\text{Re}\left[\sum\nolimits_{\ell=1}^L \alpha_{0,\ell} e^{-j\varphi_{0,\ell}} R_{k,\ell}^{(i)}\right] - \frac{T_0}{N_0}\left(\sum\nolimits_{\ell=1}^L \alpha_{0,\ell}^2\right)|\ s_k^{(i)}\ |^2\right\}, \tag{37}$$

which is the natural generalization of (17) for antenna diversity.

When everything is known, expect for the data symbols, the optimal array classifier is the one that relies on ALRT, and its likelihood function can be derived by averaging (37) with respect to the data symbols

$$\Xi_{ArrayALRTnoUnknown}^{(i)}[\mathbf{r}(t)] = \prod_{k=1}^N E_{s_k^{(i)}}\left[\exp\left\{\frac{2}{N_0}\text{Re}\left[\sum\nolimits_{\ell=1}^L \alpha_{0,\ell} e^{-j\varphi_{0,\ell}} R_{k,\ell}^{(i)}\right] - \frac{T_0}{N_0}\left(\sum\nolimits_{\ell=1}^L \alpha_{0,\ell}^2\right)|\ s_k^{(i)}\ |^2\right\}\right]. \tag{38}$$

On the other hand, when in addition to the unknown data symbols, there are some unknown parameters, i.e., $\{\alpha_{0,\ell}\}_{\ell=1}^L$ and $\{\varphi_{0,\ell}\}_{\ell=1}^L$, integration over the unknown parameters becomes more difficult for the array classifier (notice the summations $\sum_{\ell=1}^L \alpha_{0,\ell} e^{-j\varphi_{0,\ell}} R_{k,\ell}^{(i)}$ and $\sum_{\ell=1}^L \alpha_{0,\ell}^2$ in (37), which used to be single terms in (17), the single-antenna classifier of Section IV). HLRT does not seem to be a good solution also, as finding the ML estimates of several parameters, depending on the number of antennas, can be very time consuming. Therefore, we propose the QHLRT approach for the array classifier, where $\{\alpha_{0,\ell}\}_{\ell=1}^L$ and $\{\varphi_{0,\ell}\}_{\ell=1}^L$ will be estimated using the low-complexity estimators that rely on (28) and (30), respectively. Based on (37), the associated likelihood function can be written as

$$\Xi^{(i)}_{ArrayQHLRT}[\mathbf{r}(t)] = \prod_{k=1}^{N} E_{s_k^{(i)}}\Bigg[$$

$$\exp\left\{\frac{2}{N_0}\mathrm{Re}\left[\sum\nolimits_{\ell=1}^{L}\hat{\alpha}_{0,\ell}e^{-j\hat{\varphi}_{0,\ell}}R^{(i)}_{k,\ell}\right]-\frac{T_0}{N_0}\left(\sum\nolimits_{\ell=1}^{L}\hat{\alpha}^2_{0,\ell}\right)|\,s_k^{(i)}\,|^2\right\}\Bigg]. \quad (39)$$

Interestingly, the simulation results of the HQLRT array classifier, given in the next section, show that by adding even one antenna to the standard single-antenna classifier, one gets a significant performance improvement.

VII. SIMULATIONS AND DISCUSSION

A. Simulation Setup and Parameters

In simulations, we generate normalized constellations, i.e., $\vartheta^{(i)} = E[|\, s_k^{(i)}|^2] = 1$ for 16QAM, 32QAM, and 64QAM, which makes the power of the signal $S_0^{(i)}$ in (2) independent of the modulation type. Furthermore, instead of working with α_0, the channel amplitude in (1), it is more convenient to work with the signal power for which we have $S_0 = \alpha_0^2$. Then, as discussed in Section V and also based on (28), we estimate the signal power using $\hat{S}_0 = T_0^{-2}[(2\hat{\mathsf{M}}_2^2 - \hat{\mathsf{M}}_4)/0.65]^{1/2}$. We also define the received signal-to-noise ratio (SNR) per symbol as $\gamma_0 = S_0T_0/N_0$. In all the simulations we set $T_0 = 1$, and by keeping $S_0 = 1$, SNR is changed by varying N_0. When simulating the fading, the average fading power is set to 1, i.e., $\Omega_0 = E[\alpha^2] = 1$.

To estimate φ_0, the channel phase in (1), we have used eq. (30), which has a four-fold phase ambiguity [5], since QAM has a $\pi/2$ symmetry. However, as expected, our simulation results have confirmed that the probability of correct classification, defined below, has a $\pi/2$ symmetry for QAM as well. So, the phase ambiguity of the estimator has no effect on the classifier performance. When phase is assumed to be known to the classifier, we set it to $\varphi_0 = 0$.

To evaluate the performance of classifiers, we have used the average probability of correct classification, defined by $P_{cc} = N_{\mathrm{mod}}^{-1}\sum_{i=1}^{N_{\mathrm{mod}}} P_c^{(i|i)}$ for equiprobable modulations, in which $P_c^{(i|i')}$ is the conditional classification probability. More specifically, $P_c^{(i|i')}$ is the probability to declare that the i-th modulation is received, where indeed the i'-th modulation has been originally transmitted. Here $N_{\mathrm{mod}} = 3$ and $P_c^{(i|i)}$ for each modulation has been estimated via 1000 Monte Carlo trials. The number of symbols is $N = 500$.

B. Single & Multi-Antenna Classifiers in AWGN & Fading

According to Fig. 1, for a single-antenna classifier when we have perfect (error free) estimates of all the parameters (so, there is no unknown parameter), ALRT with averaging over the unknown data symbols in AWGN shows the best performance. However, the same ALRT classifier in Rayleigh fading, again with perfect estimates of the unknown parameters and averaging over the unknown data symbols, exhibits a significant performance loss. This motivates the application of antenna array, to take advantage of spatial diversity. By adding only a second antenna, we obtain a huge performance improvement, when compared to the single antenna ALRT in Rayleigh fading. For example, at $P_{cc} = 0.9$ we get a 5 dB gain in SNR. By adding a second antenna, ALRT in Rayleigh fading also converges to ALRT in AWGN. Interestingly, ALRT in AWGN gets some improvement from antenna diversity, only at low SNRs. The four curves of Fig. 1 discussed here are obtained according to (38).

Clearly it is not possible in practice to find perfect estimates of the signal power and channel phase. The classification results of the proposed QHLRT, together with the previously described MOM estimators, is shown in Fig. 1 for Rayleigh fading. As expected, we get some performance degradation, when compared to the hypothetical perfect estimates. However, by adding only a second antenna, we again see the performance enhancement. For example, at the 10 dB SNR, P_{cc} in Rayleigh fading with imperfect estimates jumps from 0.82 to 0.94. Of course this will be further improved by using more antennas. Note that with QHLRT and $L = 2$ in Rayleigh fading, P_{cc} does not increase with SNR. This is mainly because of the simple phase estimator we have used, for which the performance does not improve at high SNR [31] and becomes less accurate for large QAMs [5]. Here we have used it just to demonstrate the utility of the QHLRT concept and of course one can consider more efficient but complex estimators, to get better performance. The two curves of Fig. 1 with MOM estimates and QHLRT are obtained via (39).

C. Effect of the Number of Antennas

In Fig. 2 we have plotted, versus the number of antennas L, the P_{cc} of ALRT in AWGN and Rayleigh fading, with no unknown parameters, as well as the new QHLRT with MOM estimates in Rayleigh fading. All the curves are generated according to (38) and (39). As expected, the performance enhances with L in all the situations. The performance improvement is more significant at low SNR. Furthermore, the largest jump in P_{cc} versus L comes by adding the second antenna. Interestingly, as we observed in Fig. 1, classification in AWGN does not get that much improvement from the antenna array. For $\mathrm{SNR} = 10\,\mathrm{dB}$, P_{cc} in Fig. 2 is almost 1, whereas with $\mathrm{SNR} = 5\,\mathrm{dB}$, P_{cc} in AWGN falls below the P_{cc} in Rayleigh fading for $L = 4$, and 5. To explain this, one can argue that in AWGN with no unknown parameters, the performance is affected only by SNR, which is increased by an antenna array. However, in fading channels, the antenna array not only increases SNR, but also combats the deep fades, which are major sources of performance loss.

D. Effect of the Correlation among the Antennas

In the derivation of the array classifier in Section VI and the simulations so far, we have assumed the antenna elements are far apart, and hence the branches are uncorrelated. To see the impact of correlation on the performance, in Fig. 3 we have considered Rayleigh fading and these two-branch classifiers: ALRT with perfect estimates in (38) and QHLRT in (39). The correlation between the two antennas is defined by $\rho_{12} = E[z_1z_2^*]$, where $z_1 = \alpha_1e^{j\varphi_1}$ and $z_2 = \alpha_2e^{j\varphi_2}$ are two zero-mean complex Gaussian variables. A non-complex correlation coefficient is simulated, such that $0 \le \rho_{12} \le 1$, which corresponds to fading channels with the common model of isotropic scattering [32]. As expected, P_{cc} decreases when ρ_{12} increases. The performance degradation seems to be less at high SNRs. Overall, the array classifiers which assume uncorrelated branches, appear to be reasonably robust to some possible correlations that may exist between the branches.

E. Effect of K Factor in Rice Fading with Antenna Array

As discussed in Section IV, one of the disadvantages of ALRT in fading channels with unknown parameters is that the classifier is going to be designed for a certain fading distribution, e.g., eq. (23) for Rayleigh fading, via averaging over the channel amplitude and phase. However, when the fading distribution is changed, then the classifier should be re-designed, which most likely does not have a closed form, due to the complicated distributions of the channel amplitude and phase. For example, look at [1] for the complicated amplitude and phase distributions in Rice fading. On the other hand, The QHLRT proposed in Section V does not depend on the fading model.

To observe the utility of the QHLRT approach, see Fig. 4 where P_{cc} is plotted versus the Rice K factor, for different number of antennas. The reasonable performance of QHLRT in (39) for all K's is noteworthy for $L=1$, which is significantly enhanced by adding only one extra antenna, i.e., $L=2$. The results of ALRT in Rice fading with error free estimates of the channel amplitudes and phases in (38) are also included as a reference, to which QHLRT becomes quite close when $L \geq 2$.

Note that for $K=0$ and ∞, Rice fading reduces to Rayleigh fading and no fading, respectively [1]. This is why all the P_{cc}'s in Fig. 4 increase with K.

VIII. CONCLUSION

In this paper, after an overview of decision-theoretic approaches to blind modulation recognition, we have focused on linear and nonlinear modulations in fading channels, via a unified notation. After formulating the problem in a systematic way, we have derived closed-form expressions for the average likelihood function, required to implement the classifier, under different channel conditions. Due to the complexity of these classifiers, a new low-complexity quasi-hybrid classifier is introduced. Then the classification problem using an antenna array is formulated and solved. From the simulation results one can conclude the usefulness of the proposed quasi hybrid technique, as well as the significant performance enhancement offered by multiple antennas. Interestingly, by adding even a second antenna to a single-antenna classifier, one can get a reasonable improvement in the classification accuracy.

REFERENCES

[1] G. L. Stuber, *Principles of Mobile Communication*, 2nd ed., Boston, MA: Kluwer, 2001.

[2] H. L. Van Trees, *Detection, Estimation and Modulation Theory: Part I*. New York: Wiley, 1968.

[3] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "A survey of automatic modulation classification techniques: Classical approaches and new developments," submitted to *IEEE J. Select. Areas Commun.*, 2004.

[4] A. Abdi, "A level crossing-based algorithm for the identification of DSB and SSB modulated spherically invariant processes," in *Proc. CISS*, 2002, pp. 652-656.

[5] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum, 1997.

[6] J. A. Sills, "Maximum-Likelihood Modulation Classification for PSK/QAM," in *Proc. IEEE MILCOM*, 1999, pp. 57-61.

[7] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, pp. 189-193, 2000.

[8] K. Kim and A. Polydoros, "Digital modulation classification: the BPSK versus QPSK case," in *Proc. IEEE MILCOM*, 1988, pp. 431-436.

[9] 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.

[10] C. Y. Huang and A. Polydoros, "Likelihood methods for MPSK modulation classification," *IEEE Trans. Commun.*, vol. 43, pp. 1493-1504, 1995.

[11] P. C. Sapiano and J. D. Martin, "Maximum likelihood PSK classifier," *in Proc. IEEE ICASSP*, 1996, pp. 1010-1014.

[12] C. Long, K. Chugg and A. Polydoros, "Further results in likelihood classification of QAM signals," *in Proc. IEEE MILCOM*, 1994, pp. 57-61.

[13] 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.

[14] B. F. Beidas and C. L. Weber, "Higher-order correlation-based approach to modulation classification of digitally frequency-modulated signals," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 89-101, 1995.

[15] B. F. Beidas and C. L. Weber, "Asynchronous classification of MFSK signals using the higher order correlation domain," *IEEE J. Select. Areas Commun.*, vol. 46, pp. 480-493, 1998.

[16] B. F. Beidas and C. L. Weber, "Higher-order correlation-based classification of asynchronous MFSK signal," *in Proc. IEEE MILCOM*, 1996, pp. 1003-1009.

[17] 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.

[18] C. Y. Hwang and A. Polydoros, "Advanced methods for digital quadrature and offset modulation classification," in *Proc. IEEE MILCOM*, 1991, pp. 841-845.

[19] P. Panagiotou, A. Anastasoupoulos, and A. Polydoros, "Likelihood ratio tests for modulation classification," in *Proc. IEEE MILCOM*, 2000, pp. 670- 674.

[20] N. Lay and A. Polydoros, "Modulation classification of signals in unknown ISI environments," in *Proc. IEEE MILCOM*, 1995, pp. 170-174.

[21] K. M. Chugg, C. S. Long, and A. Polydoros, "Combined likelihood power estimation and multiple hypothesis modulation classification," in *Proc. ASILOMAR*, 1996, pp. 1137-1141.

[22] L. Hong and K. C. Ho, "BPSK and QPSK modulation classification with unknown signal level," in *Proc. IEEE MILCOM* , 2000, pp. 976-980.

[23] L. Hong and K. C. Ho, "Modulation classification of BPSK and QPSK signals using a two element antenna array receiver," in *Proc. IEEE MILCOM*, 2001, pp. 118-122.

[24] 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.

[25] O. A. Dobre, J. Zarzoso, Y. Bar-Ness, and W. Su, "On the classification of linearly modulated signals in fading channel," in *Proc. CISS*, 2004.

[26] S. M. Kay, *Fundamental of Statistical Signal Processing: Detection Theory.* Englewood Cliffs, NJ: PTR Prentice-Hall, 1998.

[27] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed., A. Jeffrey, Ed., San Diego, CA: Academic, 1994.

[28] A. Abdi, H. A. Barger, and M. Kaveh, "A simple alternative to the lognormal model of shadow fading in terrestrial and satellite channels," in *Proc. IEEE VTC*, 2001, pp. 2058-2062.

[29] A. Abdi, C. Tepedelenlioglu, M. Kaveh, and G. B. Giannakis, "On the estimation of the *K* parameter for the Rice fading distribution," *IEEE Commun. Lett.*, vol. 5, pp. 92-94, 2001.

[30] W. Dai, Y. Wang, and J. Wang, "Joint power estimation and modulation classification using second- and higher statistics," in *Proc. IEEE WCNC*, 2002, pp. 155-158.

[31] Y. Wang and E. Serpedin, "A class of blind phase recovery techniques for higher order QAM modulations: Estimators and bounds," *IEEE Signal Processing Lett.*, vol. 9, pp. 301-304, 2002.

[32] A. Abdi, J. A. Barger, and M. Kaveh, "A parametric model for the distribution of the angle of arrival and the associated correlation function and power spectrum at the mobile station," *IEEE Trans. Vehic. Technol.*, vol. 51, pp. 425-434, 2002.

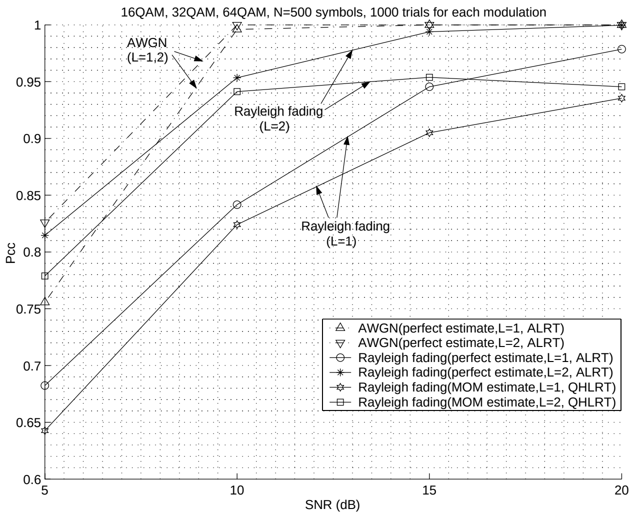

Fig. 1. Performance comparison versus SNR: ALRT in AWGN and Rayleigh fading, with no unknown parameters (perfect estimates), as well as the proposed QHLRT in Rayleigh fading (MOM estimates).

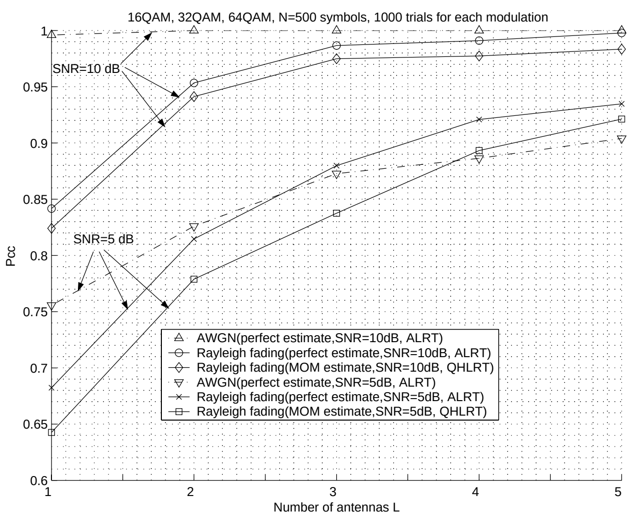

Fig. 2. Performance comparison of a multi-antenna classifier versus L: ALRT in AWGN and Rayleigh fading, with no unknown parameters (perfect estimates), as well as the new QHLRT in Rayleigh fading (MOM estimates).

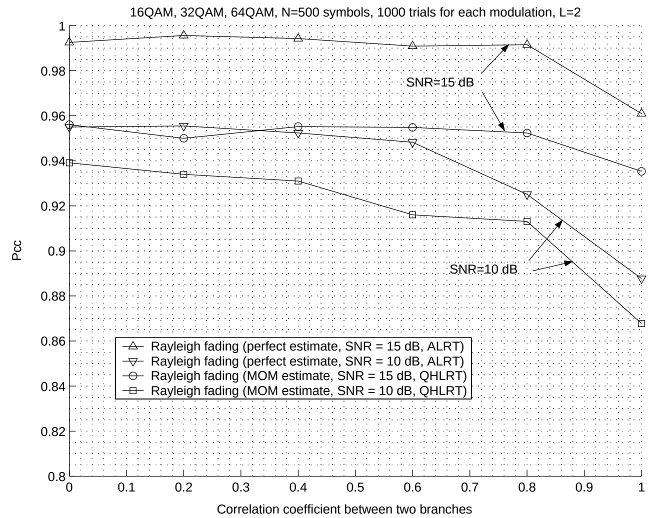

Fig. 3. Performance comparison of a two-branch array classifier in Rayleigh fading, versus the correlation among the branches: ALRT with no unknown parameters (perfect estimates), as well as the proposed QHLRT (MOM estimates).

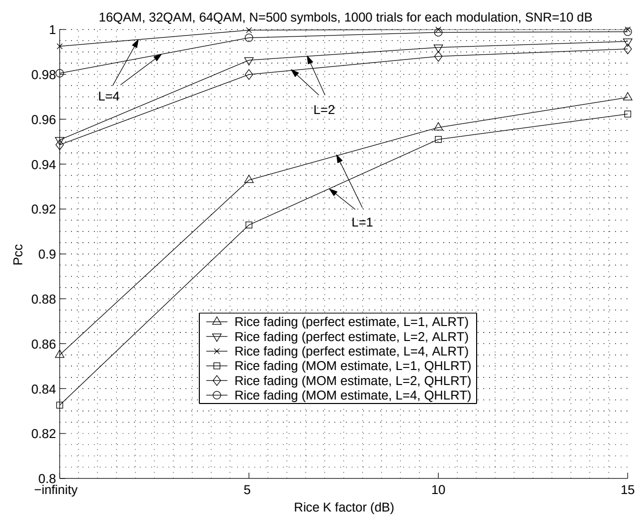

Fig. 4. Performance comparison of a multi-antenna classifier in Rice fading, versus the Rice K factor: ALRT with no unknown parameters (perfect estimates), as well as the proposed QHLRT (MOM estimates).