

# MODULATION CLASSIFICATION IN UNKNOWN DISPERSIVE ENVIRONMENTS

Bernd-Peter Paris

Geoffrey C. Orsak

Hongda Chen

Nirmal Warke

Dept. of Electrical and Computer Engineering  
George Mason University  
Fairfax, VA 22030  
USA

## ABSTRACT

The problem of distinguishing reliably between signaling formats in the presence of noise, interference, unknown dispersive channel conditions, as well as timing and frequency mismatches is addressed. Methods based on a combination of blind equalization and universal classification are presented and their performance is assessed through simulations.

## 1. INTRODUCTION

Signal modulation classification in unknown dispersive environments continues to be an open problem. Significant progress has been made in the areas of modulation classification in non-dispersive environments and blind equalization with known modulation formats in Gaussian channels. Unfortunately, solutions for the problem in which both the signaling format and the characteristics of the communication channel are unknown remain elusive.

To address this, our research aims to develop and investigate novel approaches for modulation classifiers for distinguishing reliably between signaling formats in the presence of noise, interference, unknown dispersive channel conditions, as well as timing and frequency mismatches. We present classifiers that incorporate methods for blind channel identification into universal classifiers previously shown to be effective in non-dispersive environments [3].

## 2. THE UNIVERSAL CLASSIFIER

The signal classification problem can be formulated as an M-ary hypothesis testing problem based upon test data  $\mathbf{X}^n$  and ambient channel training data  $\mathbf{Y}^N$  with a rejection region. The decision regions are defined in terms of M discriminant functions corresponding to the M hypotheses as follows:

$$\Lambda_j = \{ \mathbf{X}^n, \mathbf{Y}^N : \mathcal{K}_q^j(\mathbf{X}^n, \mathbf{Y}^N, \lambda) < 0, \mathcal{K}_q^i(\mathbf{X}^n, \mathbf{Y}^N, \lambda) \geq 0, i=1, 2, \dots, M, i \neq j \} \quad (1)$$

and the rejection region of the Universal Classifier is  $\Lambda_R = (\bigcup_{i=1}^M \Lambda_i)^c$ . The discriminant function corresponding to the  $i$ th hypothesis,  $\mathcal{K}_q^i(\mathbf{X}^n, \mathbf{Y}^N, \lambda)$  can be written as:

$$\mathcal{K}_q^i(\mathbf{X}^n, \mathbf{Y}^N, \lambda) = \min_{i \in \{1, 2, \dots, M\}} \left\{ \frac{N+n}{n} H(P_{((\mathbf{X}^n - \mathbf{S}_i(\theta)), \mathbf{Y}^N)}) - H(P_{(\mathbf{X}^n - \mathbf{S}_i(\theta))}) \right\} - \frac{N}{n} H(P_{\mathbf{Y}^N}) - \lambda \quad (2)$$

Here  $\mathbf{S}_i(\theta)$  represents the random signal corresponding to hypothesis  $H_i$  with  $\theta$  as the unknown parameter for the hypothesis  $i = 1, 2, \dots, M$ .  $H(P(\cdot))$  denotes the entropy of the type of a signal vector. For the true hypothesis the discriminant function will be close to negative  $\lambda$  since all of the entropy terms in the discriminant function should be roughly equal. The parameter  $\lambda$  controls both the size of the rejection region and the rate at which the error probability tends to zero.

This so-called Image-map classifier have been shown in previous work to distinguish reliably between different QAM formats even in severe Gaussian interference [3]. Because of this past work, significant progress has been made in the area of modulation classification in non-dispersive environments, but unfortunately the classifier is highly sensitive to the unknown dispersive channel.

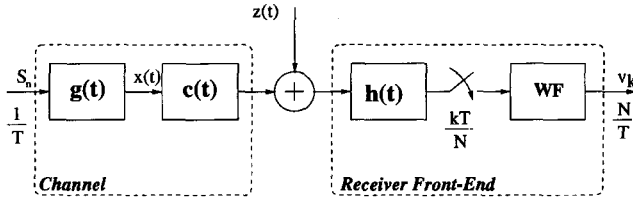
## 3. ALGORITHMIC APPROACHES

Towards integrating channel estimates into the universal modulation classifier we have two possible approaches. The first is to blindly identify the channel and "train" the classifier to recognize the filtered signal by translating the noise training sequence to all possible filtered signal locations. While this approach avoids the possibly problematic inverse filtering of the observations, it suffers from the fact that the number of possible signal locations is frequently large ( $\mathcal{A}^L$ , where  $\mathcal{A}$  equals the size of the constellation and  $L$  is the delay spread of the channel in sampling periods) and correspondingly long training sequences may be required. Alternatively, one can attempt to blindly identify the inverse of the channel to filter the test data. Since the output of such an inverse filter ideally contains the transmitted symbols (possibly multiplied by a complex exponential reflecting phase and frequency mismatch) embedded in filtered noise, the universal classifier would necessarily have to be trained with the inverse filtered training sequence. In this approach the classifier would only be "trained" to recognize the candidate constellation of size  $\mathcal{A}$ . In this paper, we focus on the second approach.

## 4. CHANNEL MODEL AND BLIND EQUALIZATION

The block diagram in Figure 1 summarizes the underlying communication system up to the receiver front-end [1].

There are many adaptive algorithms for blind equalization. Most of these algorithms are based on the use of stochastic gradient descent algorithms for self-adaptation, others are based on the use of higher-order statistics of the received signal to estimate the channel characteristics and to design the equalizer. Godard's blind equalization algorithm falls into the first category.



**Figure 1. Model of a communication system and receiver front-end.**

Let us consider a QAM data transmission system, where the data symbols are taken from a two-dimensional constellation. Throughout, we assume that the input sequence  $\{s_n^{q_i}\}$  consists of zero-mean i.i.d. random variables with discrete probability distribution  $q_i$  specified by hypothesis  $H_i$ . Denoting the baseband pulse shaping signal by  $g(t)$  and symbol interval by  $T$ , the transmitted signal under hypothesis  $H_i$  is of the form

$$x(t) = \text{Re}\left\{\sum_n s_n^{q_i} g(t - nT) e^{j2\pi f_0 t}\right\}$$

Assuming a dispersive transmission medium with additive noise  $w(t)$ , the receiver input signal can be expressed as

$$y(t) = \text{Re}\left\{\sum_n s_n^{q_i} a(t - nT) e^{j(2\pi f_0 t + \phi(t))}\right\} + w(t)$$

Where  $a(t)$  is a generally complex valued signal, incorporating both the pulse-shaping  $g(t)$  and the channel impulse response  $c(t)$ . We are currently focusing on the case where  $c(t)$  is not time varying.  $\phi(t)$  is a time-varying phase shift due to frequency offset and phase jitter.

We assume that demodulation by a local carrier with frequency  $f_0$  is carried out before equalization. Therefore, the equalizer has to process a complex signal of the general form

$$r(t) = \sum_n s_n^{q_i} a(t - nT) e^{j\phi(t)} + z(t)$$

here  $z(t)$  denotes the baseband equivalent of the noise process  $w(t)$ . The continuous time signal  $r(t)$  is then filtered and sampled at the symbol rate<sup>1</sup> to obtain the discrete time signal  $v_n$ . Hence, the signal  $v_n$  is given by

$$v_n = \sum_{l=0}^L s_{n-l}^{q_i} f_l e^{j\phi_n} + \eta_n$$

The following assumptions are made regarding the channel impulse response and the equalizer.

- The unknown channel  $\{f_l\}$  is a possibly non-minimum phase, linear, time-invariant filter in which the transfer function has no zeros on the unit circle.
- The equalizer  $\{c_l\}$  is assumed to be a FIR filter of sufficient length, so that truncation effects are insignificant.

<sup>1</sup>Clearly, an inherent assumption has been made regarding the availability of symbol timing. In the next section, we will drop this assumption and consider sampling at rates higher than the symbol rates.

With a linear equalizer, the equalizer output is given by  $z_n = \sum_{l=1}^{L_e} r_{n-l} c_l$ , where the equalizer coefficients  $\{c_l\}$ ,  $l = 1, 2, \dots, L_e$ , have to be adapted blindly. Godard's constant modulus criterion,  $D^{(p)} = E(|z_n|^p - R_p)^2$ , may be used for this purpose. Throughout this report, we focus on the case  $p = 2$ . Then, the stochastic gradient descent algorithm

$$c_{k+1} = c_k - \Delta_c r_k^* z_k (|z_k|^2 - 1)$$

may be used to estimate the channel coefficients. Here  $\Delta_c$  is step size parameter,  $c_k$  denotes the length  $L_e$  vector of coefficients and  $r_k$  denotes the vector of the  $L_e$  most recent observations. Notice that the blind criterion does not allow phase estimates. However, knowledge of the phase is not required for the Imagemap classifier.

## 5. FRACTIONALLY SPACED CHANNEL MODEL

In this section, we consider the use of general fractionally spaced, blind equalization (FSBE) in conjunction with image-map classifiers for modulation classification.

In contrast to the baud rate equalizer, a fractionally spaced equalizer is based on sampling the incoming signal more than once per symbol period. In general, a digitally implemented fractionally spaced equalizer has tap spacing of  $\frac{MT}{N}$  where  $M$  and  $N$  are integers and  $N > M$ . Throughout, we will assume  $M = 1$  to simplify our exposition.

The noise samples at the output of the fractionally spaced sampler,  $\int_{-\infty}^{\infty} z(t) h(t - k\frac{T}{N}) dt$ , are correlated if  $h(t)$  spans more than one sampling period of length  $\frac{T}{N}$ . In that case, a whitening filter (WF) is used to remove that correlation. We will assume here that the support of  $h(t)$  is confined to an interval of length  $\frac{T}{N}$  and, hence, no whitening filter is required.

Let us define the overall system impulse response as

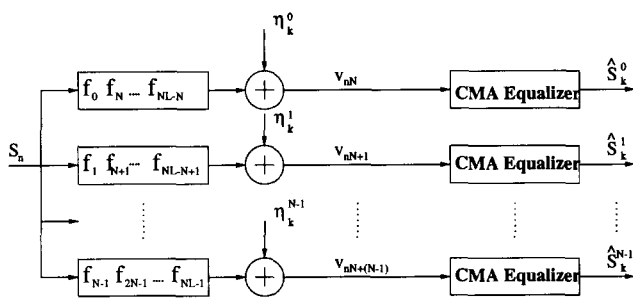
$$f(t) = g(t) \otimes c(t) \otimes h(t). \quad (3)$$

Then, the  $k$ -th sample out of the receiver front-end is given by

$$\begin{aligned} v_k &= s(t) \otimes f(t) + z(t) \otimes h(t) \big|_{t=k\frac{T}{N}} \\ &= \sum_n s_n f(k\frac{T}{N} - nT) + \eta_k, \end{aligned} \quad (4)$$

where  $s(t) = \sum_n s_n \delta(t - nT)$ . Equation (4) shows that the system in Figure 1 is equivalent to a discrete-time, multi-rate system [2]; the input rate is the symbol rate  $\frac{1}{T}$  while the output rate equals  $\frac{N}{T}$ . Two interpretations for (4) are possible. We can interpolate (up-sample) the low-rate input signal  $\{s_n\}$  and then model the channel as a FIR filter with approximately  $NL$  taps, where  $L$  is the delay spread of  $f(t)$  in symbol periods. Alternatively and consistent with the notion of polyphase filters [2], we can interpret (4) as describing a bank of  $N$  symbol-rate parallel filters with common input  $\{s_n\}$ . The coefficients of the  $l$ -th filter,  $l = 0, 1, \dots, N-1$  are  $\{f_{iN+l}\}$ , with  $i = 0, 1, \dots, L-1$ . This interpretation is illustrated in the left half of Figure 2.

Based on the two interpretations above, two avenues to proceed are possible. First, one could attempt to design an equalizer that takes the sequence  $\{v_k\}$  and seeks to extract the data symbols. Such an equalizer would necessarily be a multi-rate filter with input rate  $\frac{N}{T}$  and output rate  $\frac{1}{T}$ .



**Figure 2.** Polyphase filter interpretation of discrete-time equivalent channel with equalization.

It appears that symbol timing information is required for proper design/adaptation of this equalizer.

However, since we are not directly interested in the transmitted symbols we can instead use a symbol-rate blind equalizer for each of the polyphase filters as shown in Figure 2. The individual equalizer outputs do not need to be decimated (down-sampled) and can instead be concatenated and fed to the image-map classifier. In essence, we are using  $N$  channels with lower SNR but obtain  $N$  times as many observations as with a symbol rate sampler. Also, the precise symbol timing is not required as each branch in Figure 2 operates at the symbol rate.

## 6. SIMULATIONS

In this section, we evaluate empirically the performance of the Imagemap classifier with Godard's blind equalization algorithm in a practical QAM modulation classification scenario with an unknown dispersive environment with baud-rate and fractionally spaced sampling.

### 6.1. Baud-Rate Sampling

We consider the binary modulation classification scenario of differentiating between 4-QAM and 8-QAM. The channel is an FIR filter with tap weights given by

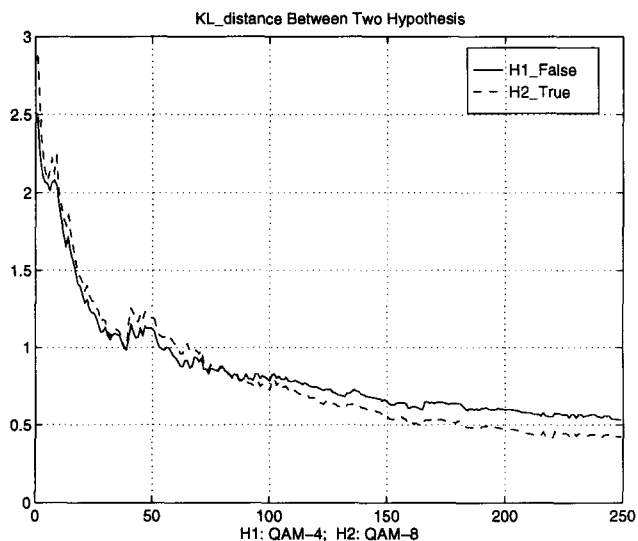
$$\{f_i\} = \{0.7474 \quad -0.5440 \quad 0.2989 \quad -0.0747 \quad 0.2247\}.$$

The noise is 10% Laplacian embedded in Gaussian noise. The average symbol energy for the two QAM schemes is taken to be unity.

The Kullback-Leibler distances between test and training data for each of the hypotheses is shown in Figure 3. Note that after about 80 test symbols, the distances separate to provide correct detection. In Figure 3, we have plotted the empirical detection probability as a function of the length of the test sequence. Even with relatively short test sequences highly reliable decisions are made. Figure 5 shows the empirical detection probabilities versus the signal-to-noise ratio (SNR). The classifier achieves more than 93% detection probability when the SNR is greater than or equal to 0 dB in this non-Gaussian environment and for the channel under consideration.

### 6.2. Fractionally Spaced Sampling

With a randomly chosen channel coefficients for  $\{f_i\}$ , we simulate the above classification scenario with  $\frac{1}{2}T$ -spaced equalizer. We assume that the delay spread of  $f(t)$  spans 3 symbol period and hence the fractionally spaced channel has 6 taps. We used two channels in our simulations, the second of which exhibited zeroes near the unit-circle. The



**Figure 3.** H1: 4-QAM, H2: 8-QAM. Kullback-Leibler Distance versus Length of Test,  $n$  Symbols. SNR = 3 dB, 10% Laplacian,  $C=2$ .

noise is 10% Laplacian embedded in Gaussian noise. The average symbol energy for the two QAM schemes is taken to be unity.

In Figure 6, we have plotted the empirical detection probability versus the signal-to-noise ratio (SNR). The figure shows the empirical detection probability if each of the branch outputs is considered alone and the detection probability if the two branch outputs are concatenated. Even with the relatively short equalizer, reliable decisions are made, and the figure shows that significant better performance is achieved when both branch outputs are considered.

The CMA blind equalizer with 12 taps is not quite capable of inverting the channel frequency response which has deep notches, particularly in the first branch. Figure 7 shows the empirical detection probabilities versus the SNR. The performance of the classifier in this channel is substantially worse than with Channel I. Note in particular that the first branch has a lower detection probability than the second which strongly supports the hypothesis that inadequate equalization is to fault for the loss in performance. Furthermore, the combined detection probability is dominated by the better individual detection probability (branch 2). Note however that short sequences of only 100 observations were used in these experiments.

## REFERENCES

- [1] K. M. Chugg and A. Polydoros, "MLSE for an Unknown Channel - Part I: Optimality Considerations," *IEEE Transactions on Communications*, vol. COM-44, July 1996.
- [2] P. P. Vaidyanathan, "Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial," *Proceedings of the IEEE*, vol. 78, January 1990.
- [3] N. Warke and G. C. Orsak, "Applications of Universal Methods to Modulation Classification in non-Gaussian Channels," *submitted to IEEE Trans. Communications*, 1996.

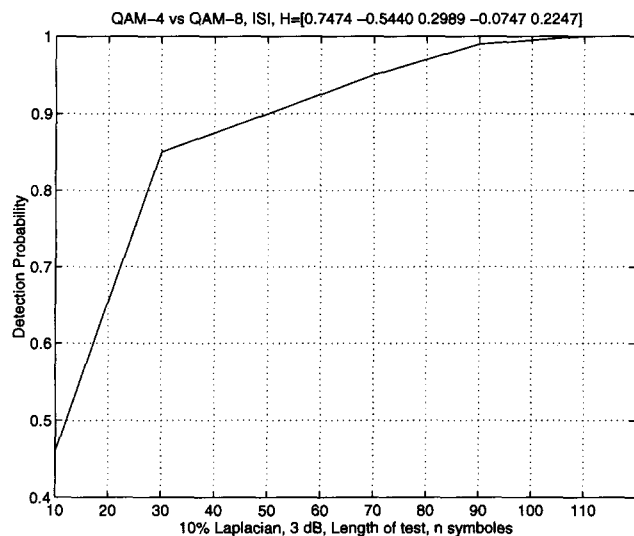


Figure 4. H1: 4-QAM, H2: 8-QAM. Detection Probability for Random Channel versus Length of Test. SNR = 3 dB, 10% Laplacian, Blind Equalizer With 12 Taps,  $C=4$ .

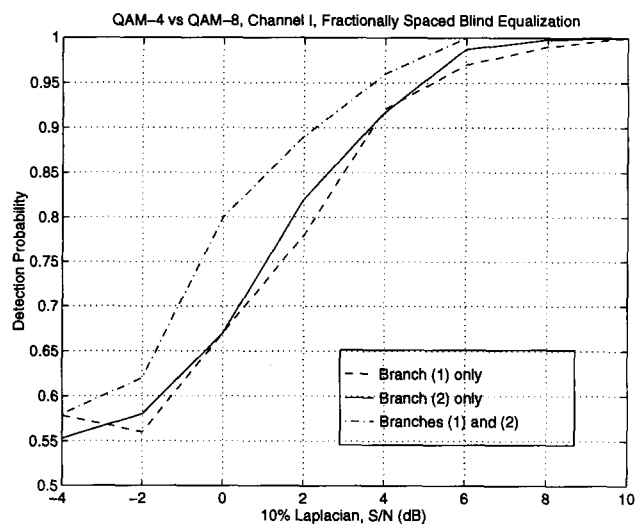


Figure 6. H1: 4-QAM, H2: 8-QAM. Detection probability for channel (I), 10% Laplacian,  $\frac{T}{2}$ -fractionally spaced blind equalizer with 4 taps, length of test 100 symbols,  $C=4$ .

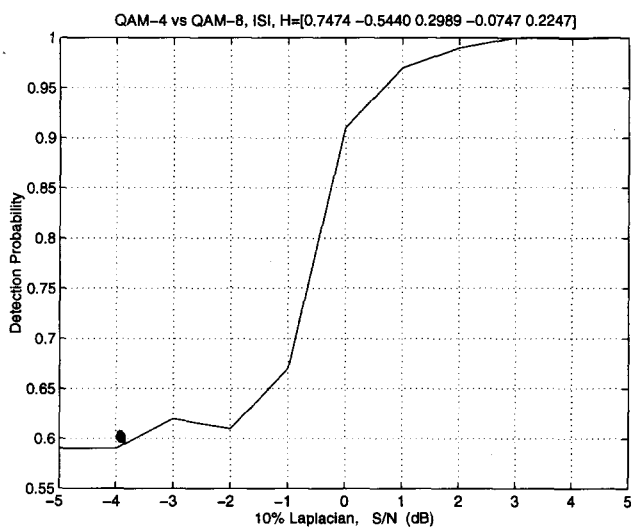


Figure 5. H1: 4-QAM, H2: 8-QAM. Detection Probability for Random Channel versus SNR. 10% Laplacian, Blind Equalizer With 12 Taps, Length of Test 250 Symbols,  $C=4$ .

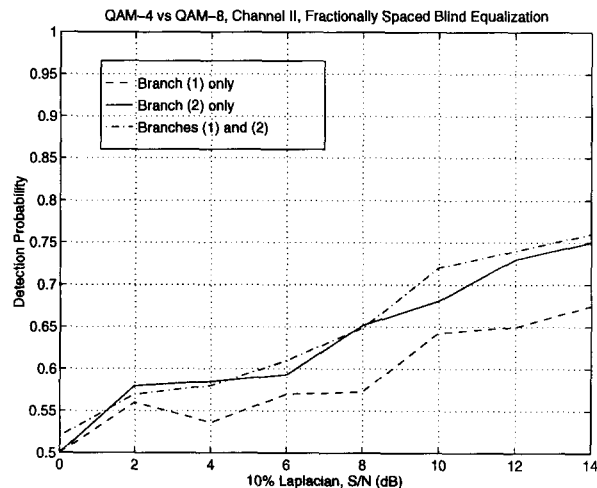


Figure 7. H1: 4-QAM, H2: 8-QAM. Detection probability for channel (II), 10% Laplacian,  $\frac{T}{2}$ -fractionally spaced blind equalizer with 12 taps, length of test 100 symbols,  $C=10$ .