EFFECTS OF SYMBOL RATE ON THE CLASSIFICATION OF DIGITAL MODULATION SIGNALS

Hussam Mustafa and Miloš Doroslovački

The George Washington University
Department of Electrical and Computer Engineering
Washington, DC 20052, USA

ABSTRACT

This paper considers features based on the multiplication of two consecutive signal values. Furthermore, three new classifiers using the features are proposed: fixed threshold tree classifier, dynamic threshold tree classifier and support vector machine (SVM) classifier. It is shown that the multiplication produces dependence of the features on the symbol rate. In order to quantify effects of this dependence the paper study the performance of the newly proposed classifiers as well as the maximum likelihood (ML) classifier [1,2], the qLLR classifier [3], and the cumulants based classifier [4]. Simulations show that the SVM classifier has promising results in the sense that it is closest to the theoretically optimal results obtained by the ML classifier.

1. INTRODUCTION

Recognition of modulation in received signals is important for many applications such as signal interception, interference identification, electronic warfare, enforcement of civilian spectrum compliance, radar and intelligent modems. The modulation recognition methods can be divided into two categories. The first is modulation recognition with prior information available. The information provides knowledge of signal parameters such as amplitude, carrier frequency, symbol rate, pulse shape, initial phase, channel characteristic and noise power. The second, and more challenging, is modulation recognition without any prior information about signal parameters.

In the past years there have been different approaches to solve the modulation recognition problem. These approaches can be classified in three groups. The first are approaches that use memoryless nonlinearities and detect the spectrum lines occurring for specific modulation types [5]. The second are the feature based approaches, where the recognition is divided into two stages. The first stage maps the signal into a smaller feature domain. Usually the feature domain is independent of the signal's parameters. The second stage does the classification of the signal by comparing the measured values of features to a priori collocation of the feature values for each modulation type [4]. The third are the decision theoretic approaches, where all the signal parameters are known with some exceptions (the classifier in [3] does not need to know the initial phase). These approaches use the likelihood function to do recognition. They are optimal in the sense of the minimum probability of misclassification.

Here we present features to classify the following modulations: amplitude shift keying with two levels (ASK2), amplitude shift keying with four levels (ASK4), binary phase shift keying (PSK2), quadrature phase shift keying (PSK4), binary frequency shift keying (FSK2) and four frequency shift keying (FSK4).

The paper is organized in the following way. Section 2 presents the signal model. Section 3 presents the features used in the classification. Section 4 presents the classification algorithms used in simulation. Section 5 presents the simulation results with discussion.

2. SIGNAL MODEL AND ASSUMPTIONS

We consider the following complex baseband signal

$$s(k) = x(k) + n(k) \tag{1}$$

where $x(k)$ is the transmitted signal. In the case of ASK and PSK modulation the transmitted signal is

$$x(k) = \sum_{n=0}^{N-1} a_n e^{j(\theta_n + \theta_c)} p(k - nT) \tag{2}$$

and in the case of FSK modulation

$$x(k) = \sum_{n=0}^{N-1} e^{j(\omega_n k + \theta_c)} p(k - nT) \tag{3}$$

where $(a_n, \theta_n, \omega_n)$ are the amplitude, phase and the frequency of the modulating signal. θ_c is the initial phase. $p(k - nT)$ is the rectangular pulse shape function defined as

$$p(k) = \begin{cases} 1, & 0 \le k \le T-1 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

and T is symbol period. $n(k)$ is assumed to be complex white Gaussian noise with power σ^2.

The signal model used here is different from the one used by Wei and Mandel [1, Equation 1.6], where the model for bandpass signals is assumed. Since in [1] the carrier frequency is assumed known, without loss of generality we can work with signals in the baseband.

The signal constellations are assumed to have unit average power. This assumption stays as it was in [1]. The pulse shape is rectangular and the noise power is known to the receiver in the case of maximum likelihood, qLLR, cumulants, dynamic threshold algorithm and SVM classifier. Although, the assumption that the carrier frequency is an integer multiple of the symbol rate is used in order to include the FSK modulated signals (in the case of maximum likelihood classification) , this assumption is not needed in the case of ASK and PSK signals. In all other aspects the assumed scenario is the same as the one described by [1, page 17].

3. PROPOSED FEATURES

The modulation recognition is based on two signal features. The first feature is

$$NP1(k) = x(k)x^*(k-1) \tag{5}$$

where $x\ (k)$ is the modulated signal.

Applying $NP1$ on ASK signal yields

$$NP1(k) = \begin{cases} a_n p(k-nT)e^{j\theta_c} a_n p(k-nT-1)e^{-j\theta_c}, & nT+1 \le k \le (n+1)T-1 \\ a_n p(k-nT)e^{j\theta_c} a_{n-1} p(k-1-(n-1)T)e^{-j\theta_c}, & k = nT \end{cases}$$
$$= \begin{cases} a_n{}^2, & nT+1 \le k \le (n+1)T-1 \\ a_n a_{n-1}, & k = nT. \end{cases} \tag{6}$$

Applying $NP1$ on PSK signal yields

$$NP1(k) = \begin{cases} e^{j\theta_n} p(k-nT)e^{j\theta_c} e^{-j\theta_n} p(k-nT-1)e^{-j\theta_c}, & nT+1 \le k \le (n+1)T-1 \\ e^{j\theta_n} p(k-nT)e^{j\theta_c} e^{-j\theta_{n-1}} p(k-1-(n-1)T)e^{-j\theta_c}, & k = nT \end{cases}$$
$$= \begin{cases} 1, & nT+1 \le k \le (n+1)T-1 \\ e^{j(\theta_n - \theta_{n-1})}, & k = nT. \end{cases} \tag{7}$$

Applying $NP1$ on FSK signal yields

$$NP1(k) = \begin{cases} e^{j\omega_n k} p(k-nT)e^{j\theta_c} e^{-j\omega_n (k-1)} p(k-nT-1)e^{-j\theta_c}, & nT+1 \le k \le (n+1)T-1 \\ e^{j\omega_n k} p(k-nT)e^{j\theta_c} e^{-j\omega_{n-1}(k-1)} p(k-1-(n-1)T)e^{-j\theta_c}, & k = nT \end{cases}$$
$$= \begin{cases} e^{j\omega_n}, & nT+1 \le k \le (n+1)T-1 \\ e^{j(\omega_n - \omega_{n-1})nT} e^{j\omega_{n-1}}, & k = nT. \end{cases} \tag{8}$$

From (6) and (7) it is clear that the mean of the imaginary part of $NP1$ is always 0 in the case of ASK and PSK2. In the case of PSK4 the imaginary part of $NP1$ has impulses at the transitions. The assumption of equally probable symbols makes the mean equal to 0. Moreover, from (8) we can see that the imaginary part of FSK signal is the sine function of the carrier frequencies. From this result we derive our first branch in the classification tree that distinguishes ASK and PSK signals form FSK signals.

From (6) and (7) also it is clear that the real part of $NP1$ applied to *x(k)* has multiple levels in the case of ASK2 and ASK4 and one level (if we exclude the transitions instants) in the case of PSK2 and PSK4. This phenomenon encouraged the idea of using the kurtosis of the real part of $NP1$ as a method to distinguish between ASK and PSK signals.

Taking kurtosis of the real part of $NP1$ when applied to *s(k)* gives the following expressions:

$$Kurtosis_{ASK2} = \frac{(1-p)[8+24\sigma^2+24\sigma^4+6\sigma^6+1.875\sigma^8+16c_{ASK2}+}{(1-p)^2[2+\sigma^2+.5\sigma^4+2c_{ASK2}+c^2{}_{ASK2}]^2+}$$
$$\frac{24\sigma^2 c_{ASK2}+12\sigma^4 c_{ASK2}+12c^2{}_{ASK2}+6\sigma^2 c^2{}_{ASK2}+3\sigma^4 c^2{}_{ASK2}+4c^3{}_{ASK2}+c^4{}_{ASK2}]+}{2p(1-p)[2+\sigma^2+.5\sigma^4+2c_{ASK2}+c^2{}_{ASK2}]}$$
$$\frac{p[4+12\sigma^2+13.5\sigma^4+6\sigma^6+1.5\sigma^8+8c_{ASK2}+6\sigma^4 c_{ASK2}+12\sigma^2 c_{ASK2}+6c^2{}_{ASK2}+}{[1+\sigma^2+.5\sigma^4+c_{ASK2}+c^2{}_{ASK2}]+}$$
$$\frac{6\sigma^2 c^2{}_{ASK2}+3\sigma^4 c^2{}_{ASK2}+2c^3{}_{ASK2}+c^4{}_{ASK2}]}{[1+\sigma^2+.5\sigma^4+c_{ASK2}+c^2{}_{ASK2}]^2} \tag{9}$$

where p is the ratio between the symbol rate and the sampling rate, and

$$c_{ASK2} = -(1-.5p); \tag{10}$$

$$Kurtosis_{ASK4} = \frac{(1-p)[11.3548+27.7716\sigma^2+24\sigma^4+6\sigma^6+1.875\sigma^8+}{(1-p)^2[2+\sigma^2+.5\sigma^4+2c_{ASK4}+c^2{}_{ASK4}]^2+}$$
$$\frac{4c_{ASK4}(4.6286+6\sigma^2+3\sigma^4)+6c^2{}_{ASK4}(2+\sigma^2+\sigma^4)+4c^3{}_{ASK4}+c^4{}_{ASK4}]+}{2p(1-p)[2+\sigma^2+.5\sigma^4+2c_{ASK4}+c^2{}_{ASK4}]}$$
$$\frac{p[4+12\sigma^2+13.5\sigma^4+6\sigma^6+1.5\sigma^8+4c_{ASK4}(1.8887+1.9284\sigma^4+3.3056\sigma^2)}{[1+\sigma^2+.5\sigma^4+1.2856c_{ASK4}+c^2{}_{ASK4}]+}$$
$$\frac{+6c^2{}_{ASK4}(1+\sigma^2+.5\sigma^4)+2.5712c^3{}_{ASK4}+c^4{}_{ASK4}]}{[1+\sigma^2+.5\sigma^4+1.2856c_{ASK4}+c^2{}_{ASK4}]^2} \tag{11}$$

where

$$c_{ASK4} = -[(1-p)+.6428p]; \tag{12}$$

$$Kurtosis_{PSK} = \frac{(1-p)[1+6\sigma^2+12\sigma^4+6\sigma^6+1.875\sigma^8+4c_{PSK}(1+3\sigma^2+3\sigma^4)+}{(1-p)^2[1+\sigma^2+.5\sigma^4+2c_{PSK}+c^2{}_{PSK}]^2+}$$
$$\frac{6c^2{}_{PSK}(1+\sigma^2+.5\sigma^4)+4c^3{}_{PSK}+c^4{}_{PSK}]+p[1+6\sigma^2+12\sigma^4+6\sigma^6+1.5\sigma^8+}{2p(1-p)[1+\sigma^2+.5\sigma^4+2c_{PSK}+c^2{}_{PSK}][1+\sigma^2+.5\sigma^4+c^2{}_{PSK}]+}$$
$$\frac{6c^2{}_{PSK}(1+\sigma^2+.5\sigma^4)+c^4{}_{PSK}]}{[1+\sigma^2+.5\sigma^4+c^2{}_{PSK}]^2} \tag{13}$$

where

$$c_{PSK} = p-1. \tag{14}$$

The second feature is

$$NP2(k) = x(k)x(k-1). \tag{15}$$

This feature compensates the effect of using the conjugate in the first feature, where the information content in the phase, which helps in distinguishing between PSK2 and PSK4, is lost. Applying $NP2$ on (1) directly (and assuming $x(k)$ is a PSK signal), we have

$$NP2(k) = \begin{cases} e^{j\theta_n}p(k-nT)e^{j\theta_c}e^{j\theta_n}p(k-nT-1)e^{j\theta_c}, & nT+1 \le k \le (n+1)T-1 \\ e^{j\theta_n}p(k-nT)e^{j\theta_c}e^{j\theta_{n-1}}p(k-1-(n-1)T)e^{j\theta_c}, & k = nT \end{cases}$$
$$= \begin{cases} e^{2j(\theta_n+\theta_c)}, & nT+1 \le k \le (n+1)T-1 \\ e^{j(\theta_n+\theta_{n-1}+2\theta_c)}, & k = nT. \end{cases} \tag{16}$$

Calculating the kurtosis of the real part of (16) gives us:

$$Kurtosis_{PSK2} = \frac{(1-p)[\cos^4(2\theta_c)+6\sigma^2\cos^2(2\theta_c)+9\sigma^4\cos^2(2\theta_c)+3\sigma^4+6\sigma^6+1.5\sigma^8}{(1-p)^2[\cos^2(2\theta_c)+\sigma^2+.5\sigma^4+2c_{PSK2}\cos(2\theta_c)+c^2{}_{PSK2}]^2+}$$
$$\frac{4c_{PSK2}(\cos^3(2\theta_c)+3\sigma^2\cos(2\theta_c)+3\sigma^4\cos(2\theta_c))+6c^2_{PSK2}(\cos^2(2\theta_c)+\sigma^2+.5\sigma^4)+}{2p(1-p)[\cos^2(2\theta_c)+\sigma^2+.5\sigma^4+2c_{PSK2}\cos(2\theta_c)+c^2{}_{PSK2}]}$$
$$\frac{4c^3_{PSK2}\cos(2\theta_c)+c^4_{PSK2}]+p[.75+.5\cos(8\theta_c)+\cos(4\theta_c)+6\sigma^2\cos^2(2\theta_c)+9\sigma^4\cos^2(2\theta_c)}{[\cos^2(2\theta_c)+\sigma^2+.5\sigma^4+c^2{}_{PSK2}]+}$$
$$\frac{3\sigma^4+6\sigma^6+1.5\sigma^8+6c^2_{PSK2}(\cos^2(2\theta_c)+\sigma^2+.5\sigma^4)+c^4_{PSK2}]}{(1-p)^2[\cos^2(2\theta_c)+\sigma^2+.5\sigma^4+c^2{}_{PSK2}]^2} \tag{17}$$

and

$$Kurtosis_{PSK4} = \frac{(1-p)[\sin^4(2\theta_c)+6\sigma^2\sin^2(2\theta_c)+9\sigma^4\sin^2(2\theta_c)+3\sigma^4+6\sigma^6+1.5\sigma^8}{(1-p)^2[\sin^2(2\theta_c)+\sigma^2+.5\sigma^4+c^2{}_{PSK4}]^2+}$$
$$\frac{+6c^2_{PSK4}(\sin^2(2\theta_c)+\sigma^2+.5\sigma^4)+c^4_{PSK4}]+p[0.375+.125\cos(8\theta_c)+3\sigma^2+7.5\sigma^4+}{2p(1-p)[\sin^2(2\theta_c)+\sigma^2+.5\sigma^4+c^2{}_{PSK4}][.5+\sigma^2+.5\sigma^4+c^2{}_{PSK4}]+}$$
$$\frac{6\sigma^6+1.5\sigma^8+6c^2_{PSK4}(.5+\sigma^2+.5\sigma^4)+c^4_{PSK4}]}{(1-p)^2[.5+\sigma^2+.5\sigma^4+c^2{}_{PSK4}]^2} \tag{18}$$

where

$$c_{PSK2} = -(1-p)\cos(2\theta_c)$$
$$c_{PSK4} = 0. \tag{19}$$

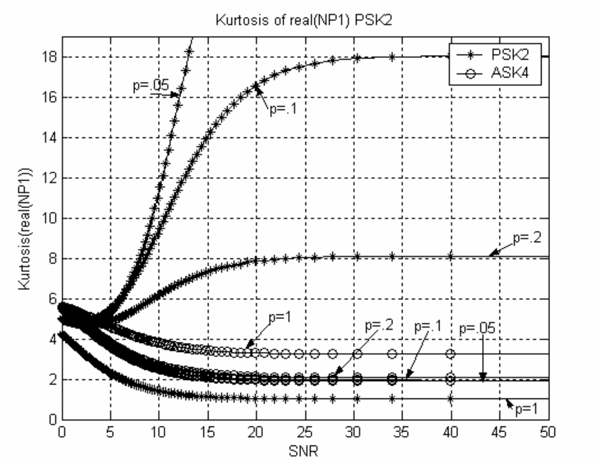

Fig. 1. Kurtosis curves of $NP1$ in the case of ASK4 and PSK4 for different values of p.

Calculating the kurtosis of the imaginary part of (16) gives us the same results as in (17-19), with only the following exception:

$$c_{PSK2} = -(1-p)\sin(2\theta_c). \tag{20}$$

The third feature distinguishing FSK2 from FSK4 is the ratio of the second maximum of the Fourier transform to the third maximum.

Figure 1 shows the plot of (11) and (13) as a function of SNR for different values of p. From the figure it is clear that as p decreases the distance between the kurtosis curves in the case of PSK and ASK4 increases at SNR higher than 5dB. It can be shown that the kurtosis curve in the case of ASK2 behaves in a similar manner as in the case of ASK4. Although, the distance between (9) and (11) is small compared to the distance of (9) or (11) to (13), the two curves are separated enough to distinguish them from each other. It should be noted at SNR<5 dB the curves of kurtosis for different modulations overlap for different values of p.

4. CLASSIFICATION

In the classification stage we compare the work done in [1], [3] and [4] to three newly proposed classification algorithms in the case of PSK2 and PSK4. However, we do compare [1] with the newly proposed classifiers for the other considered modulations. From the results of Section 3 it is clear that the kurtosis is SNR dependent. Because of that we constructed two classification trees. The first determines fixed thresholds and does the classification according. The second algorithm determines the threshold based on the knowledge of the SNR. The first is called the fixed threshold algorithm and the second is dynamic threshold algorithm. The last proposed algorithm is the

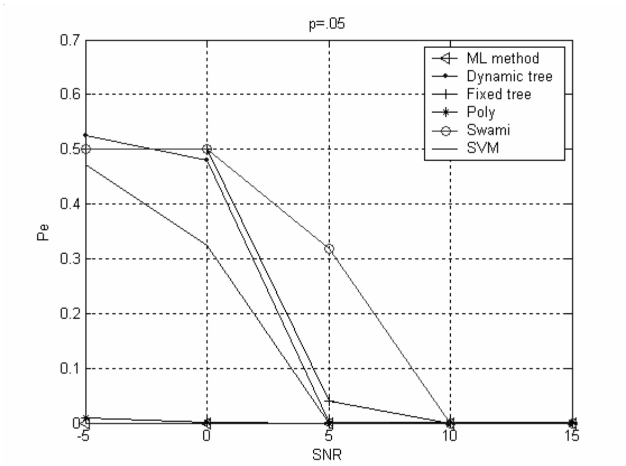

Fig. 2. Probability of classification error (P_e) for 1000 PSK2 and PSK4 for different SNRs. 'ML' represents the maximum likelihood classifier, 'Dynamic tree' is the proposed dynamic threshold classifier, 'Fixed tree' is the proposed fixed threshold classifier, 'Poly' is the qLLR classifier, 'Swami' is the cumulants classifier and 'SVM' is the proposed SVM classifier. p=.05.

SVM classifier. SVM is an empirical modeling algorithm that can be applied in classification problems. The first objective of the Support Vector Classification (SVC) is the maximization of the margin between the two nearest data points belonging to two separate classes. The second objective is to constrain all data points to belong to the right class. It is a two-class solution which can use multi-dimensional features. The two objectives of the SVC problem are then incorporated into an optimization problem. This is done by constructing the dual and primal problems of the classical Lagrangian problem by transferring the constraint of the second objective to become constraints on the Lagrange variables. The complete derivation of SVC is given in [6].

5. RESULTS AND DISCUSSION

Figure 2 and 3 represents a sample of the simulation results comparing the classifiers discussed in Section 4. It is clear the maximum likelihood classifier is the best classifier out of the six classifiers. However, in the maximum likelihood classifier all the signal parameters including the values of the signal constellation points are known to the receiver.

In the case of qLLR classifier the simulation results show that for small p the qLLR classifies at low SNR. However as we increase the value of p qLLR fails to classify between the PSK2 and PSK4 modulation signals. It should be noted that in the simulation we used the cumulants and qLLR classifier for PSK2 and PSK4 signals only. In the general case where all the six modulation are present these two classifiers are not used.

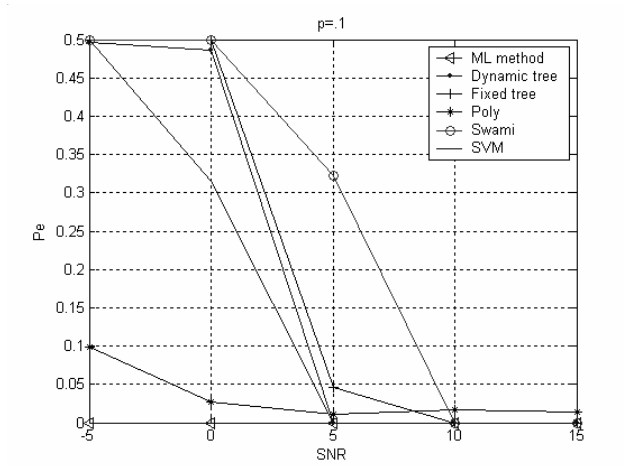

Fig. 3. Probability of classification error (P_e) for 1000 PSK2 and PSK4 for different SNRs. The acronyms are same as for Figure 2. p= .1.

As expected the dynamic threshold classifier outperforms the fixed threshold classifier. This is due to the curvature of the kurtosis curves at SNR<10 dB. Also, the simulation showed that the performance of the cumulants based classifier "Swami" is independent of p. The classifier achieves 0 probability of classification error at SNR =10dB. Finally, the simulation results demonstrate similar consistent performance for both the SVM classifier and dynamic threshold classifier for different values of p. However in the 0dB area the SVM outperforms the dynamic tree classifier. This is due to the fact that the curves of kurtosis overlap in that region and the SVM classifier is modified such that it can be used on nonseparable data.

6. REFERENCES

[1] W. Wei "Classification of Digital Modulation Using Constellation Analyzes." Ph.D. dissertation. Univ. of Southern California. 1998.

[2] W. Wei, JM. Mendel. "Maximum-likelihood classification for digital amplitude-phase modulations", *IEEE Transactions on Communications*, vol.48, no.2, pp.189-93, Feb. 2000.

[3] A. Polydoros, K. Kim. "On the detection and classification of quadrature digital modulations in broad-band noise", *IEEE Transactions on Communications*, vol.38, no.8, pp.1199-211 Aug. 1990.

[4] A. Swami, BM. Sadler. "Hierarchical digital modulation classification using cumulants", *IEEE Transactions on Communications,* vol.48, no.3, pp.416-29, March 2000.

[5] J. Reichert. "Automatic classification of communication signals using higher order statistics", *ICASSP*, vol.5, New York, NY, USA. 1992 pp.221-4 .

[6] C. Burges" A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, Vol. 2, 1998, pp 121-167.