EFFECT OF CARRIER OFFSET ON THE CLASSIFICATION OF PHASE SHIFT KEYING MODULATION USING THE SUBTRACTION OF TWO CONSECUTIVE SIGNAL VALUES

H. Mustafa and M. Doroslovački
Department of Electrical and Computer Engineering
The George Washington University
Washington, DC 20052, USA

ABSTRACT

In this paper we study the effect of carrier offset on a feature distinguishing binary phase shift keying and quadrature phase shift keying. The feature is based on the subtraction of two consecutive signal values and the kurtosis of the absolute value of this subtraction. The paper compares performance of the proposed feature to a cumulant based feature when both are used in a support vector machine classifier. Also the paper compares the proposed-feature based classification to the maximum likelihood classification and to the quasi-log-likelihood ratio classifier. The simulation results showed that the proposed feature classifier has a robust performance with respect to carrier offset compared to the other above mentioned classifiers.

Index Terms— Phase shift keying, Quadrature phase shift keying, Pattern classification, Feature extraction, Maximum likelihood detection.

1. INTRODUCTION

Designing a classification algorithm to solve the digital modulation classification (DMC) problem is a challenging undertaking. When designing a classification algorithm, three main issues should be considered. First, the signal of interest propagates through a channel which affects the signal by distorting it with noise, fading and multi-path delay. Second, not all the signal parameters are known to the receiver; this makes the decision of the classifier more difficult since one or more of the main parameters shaping the signal of interest are unknown. Third, the complexity of the classification algorithm is a concerning issue. For this reason, noticeable effort has been directed towards the lowering of the complexity of the classification algorithm.

Today's communication systems more and more rely on intelligent receivers capable of distinguishing between many different signals. The receivers must have the ability to classify a signal in environments that are less than ideal with little prior information about the signal. This process constitutes the first block in the communication system which in the end targets retrieving the signal's information content. There are two main approaches to solve the DMC problem. The first is the decision theoretic approach. The second is the pattern recognition based approach.

In [1] Wei and Mendel present a maximum likelihood (ML) solution using the I-Q domain of the data as a sufficient statistic. In their approach they assume the ideal case scenario where all the signal parameters and noise power are known to the receiver. Polydoros in [2] uses the quasi-log-likelihood ratio (qLLR) (an approximation to ML) to classify between PSK2 and PSK4. Although in [2] it is not required to know the initial phase, all other signal parameters and noise power are known to the receiver. This is not always the case in real life problems, where sometimes all the signal parameters are unknown to the receiver. To solve these problems researchers turn to other alternative solutions which require less computation complexity and/or require less prior information about the signal parameters. These requirements lead to the pattern recognition approach.

In the pattern recognition approach the received signal is mapped to a feature space where the classification is being made. One source of the features used to solve the DMC problem using pattern recognition approach are cumulants [3]. The reason for using cumulants is that the classification algorithms benefit from the fact that higher order cumulants for Gaussian random variable equal zero [4].

This paper proposes a new feature and uses it to construct the support vector machine classification algorithm. The classification algorithm distinguishes binary phase shift keying modulation (PSK2) and quadrature phase shift keying modulation (PSK4).

In this paper we present the signal model we assume (Section 2), as well as the new proposed feature (Section 3) for modulation recognition. Further, we describe briefly the support vector machine (SVM) algorithm (Section 4). Finally, we present and comment on simulation results (Section 5).

2. SIGNAL MODEL

We consider the following complex baseband discrete-time signal

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

where $x(k)$ is the transmitted signal such that

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

θ_n is the phase of the modulating signal. θ_c is the initial phase and Δ is the carrier offset. As in the majority of relevant literature, $P(k-nT)$ is the rectangular pulse shape function defined as

$$P(k) = \begin{cases} 1, & 0 \le k \le T-1 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

and T is the number of samples per symbol period. $v(k)$ is assumed to be complex white Gaussian noise with the real and imaginary part having the same variance equal to $\sigma^2/2$.

3. PROPOSED FEATURE

The modulation recognition is based on the following signal feature

$$F(k) = \left| s(k) - s(k-1) \right| \; . \tag{4}$$

Applying F on PSK signals and assuming $v(k) = 0$ yields

$$F(k) = \begin{cases} 2\left|\sin\left(\frac{\Delta}{2}\right)\right|, & nT+1 \le k \le (n+1)T-1 \\ 2\left|\sin\left(\frac{\theta_n - \theta_{n-1} + \Delta}{2}\right)\right|, & k = nT. \end{cases} \tag{5}$$

Next, we take the sample kurtosis of (5)

$$\kappa(F(k)) = \frac{TN \sum_{k=0}^{NT-1} \left[F(k) - \bar{F}\right]^4}{\left[\sum_{k=0}^{NT-1} \left[F(k) - \bar{F}\right]^2\right]^2} \tag{6}$$

where

$$\bar{F} = \frac{1}{NT} \sum_{k=0}^{NT-1} F(k) \; . \tag{7}$$

Expanding (6), assuming $NT >> 1$ and using the law of large numbers we have

$$\kappa(F(k)) \approx \frac{E[F^4(k)] - 4E[F^3(k)]E[F(k)] + 6E[F^2(k)]E^2[F(k)] - 3E^4[F(k)]}{E^2[F^2(k)] + E[F^2(k)]E^2[F(k)] + E^4[F(k)]} \; . \tag{8}$$

The law of large number can be applied in this case based on results in [5], [6].

Calculating the second and fourth moments in (8) is straightforward. Assuming $\Delta = 0$ we determine the second and fourth moments in the PSK2 signal case as

$$E\left[F^2(k)\right] = 2p + 2\sigma^2 \; , \tag{9}$$

$$E\left[F^4(k)\right] = 8p + 16p\sigma^2 + 8\sigma^4 \tag{10}$$

where p is the ratio of symbol rate to sampling rate. In the case of PSK4 signals the second moment is given by (9) and the fourth moment is

$$E\left[F^4(k)\right] = 6p + 16p\sigma^2 + 8\sigma^4 \; . \tag{11}$$

On the other hand, it is difficult to determine closed-form expression for the first and third moments of (5). Therefore we calculated them using numerical integration. By letting $y(k) = x(k) - x(k-1)$ and $w(k) = v(k) - v(k-1)$ we modify (4) by using the cosine theorem such that

$$F(k) = \sqrt{|y(k)|^2 + |w(k)|^2 + 2|y(k)||w(k)|\cos(\theta)} \tag{12}$$

where $|y(k)|$ is a discrete random variable, $|w(k)|$ is a Rayleigh random variable and θ is uniformly distributed over $[0, 2\pi)$. $|y(k)|$, $|w(k)|$ and θ are independent. Next, we average over $|y(k)|$ assuming PSK2 is present and given $|w(k)|$ and θ:

$$E_{|y(k)|}\left[F(k) \middle| \text{ PSK2}, |w(k)|, \theta\right] = (1 - \frac{p}{2})|w(k)| + \frac{p}{2}\sqrt{4 + |w(k)|^2 + 4|w(k)|\cos(\theta)} \; . \tag{13}$$

Following that, we average over $|w(k)|$ and θ

$$E_{|w(k)|,\theta}\left[E_{|y(k)|}\left[F(k) \middle| \text{ PSK2}, |w(k)|, \theta\right]\right] = \frac{1}{2\pi}\int_0^\infty \int_0^{2\pi} \left[(1 - \frac{p}{2})|w(k)| + \frac{p}{2}\sqrt{4 + |w(k)|^2 + 4|w(k)|\cos(\theta)}\right] f(|w(k)|) d\theta dw \tag{14}$$

where $f(|w(k)|)$ is the probability density function of the random variable $|w(k)|$. Applying the change of variable rule of integration in (14)

$$E_{V,\theta}\left[E_{|y(k)|}\left[F(k) \middle| \text{ PSK2}, V, \theta\right]\right] = \frac{1}{4\pi\sigma^2}\int_0^\infty \int_0^{2\pi} \left[(1 - \frac{p}{2})\sqrt{V} + \frac{p}{2}\sqrt{4 + V + 4\sqrt{V}\cos(\theta)}\right] e^{-\frac{V}{2\sigma^2}} d\theta dV \tag{15}$$

where $V = |w(k)|^2$. Similarly

$$E_{V,\theta}\left[E_{|y(k)|}\left[F^3(k) \middle| \text{ PSK2}, V, \theta\right]\right] = \frac{1}{4\pi\sigma^2}\int_0^\infty \int_0^{2\pi} \left[(1 - \frac{p}{2})\sqrt{V^3} + \frac{p}{2}\sqrt{4 + V^3 + 4\sqrt{V^3}\cos(\theta)}\right] e^{-\frac{V}{2\sigma^2}} d\theta dV \; . \tag{16}$$

In the case of PSK4 we obtain

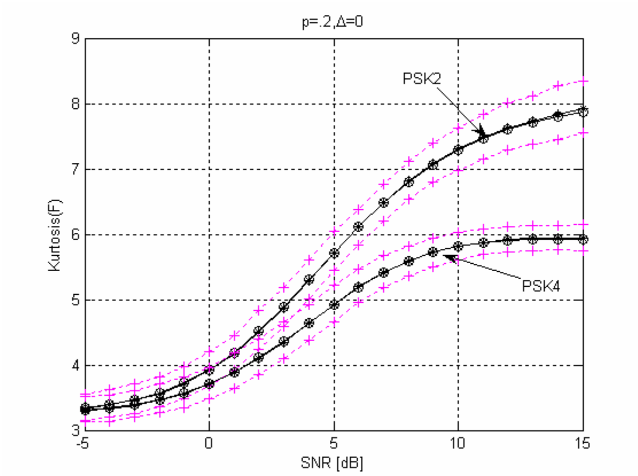

Figure 1: Kurtosis of $F(k)$ for 6000 PSK2 and PSK4 signals. 3000 realizations for each modulation. Solid lines with * are the mean and the dotted lines are the mean+standard deviation and the mean-standard deviation, all obtained by simulation. Solid lines with o are the calculated theoretical curves. $\Delta = 0$ and p=.2.

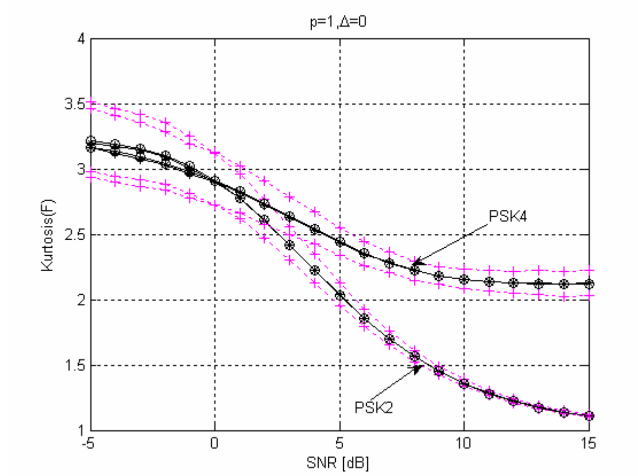

Figure 2: Kurtosis of $F(k)$ for 6000 PSK2 and PSK4 signals. 3000 realizations for each modulation. Solid lines with * are the mean and the dotted lines are the mean+standard deviation and the mean-standard deviation, all obtained by simulation. Solid lines with o are the calculated theoretical curves. $\Delta = 0$ and p=1.

$$E_{V,\theta}\left[E_{|y(k)|}\left[F(k)\middle|\ \text{PSK4},V,\theta\right]\right]=\frac{1}{4\pi\sigma^2}\int_0^\infty\int_0^{2\pi}\Bigg[(1-\frac{3p}{4})\sqrt{V}+\frac{p}{2}\sqrt{2+V+2\sqrt{2V}\cos(\theta)}+\frac{p}{4}\sqrt{4+V+4\sqrt{V}\cos(\theta)}\Bigg]e^{-\frac{V}{2\sigma^2}}d\theta dV. \qquad (17)$$

Similarly, in the case of PSK4

$$E_{V,\theta}\left[E_{|y(k)|}\left[F^3(k)\middle|\ \text{PSK4},V,\theta\right]\right]=\frac{1}{4\pi\sigma^2}\int_0^\infty\int_0^{2\pi}\Bigg[(1-\frac{3p}{4})\sqrt{V^3}+\frac{p}{2}\sqrt{2+V^3+2\sqrt{2V^3}\cos(\theta)}+\frac{p}{4}\sqrt{4+V^3+4\sqrt{V^3}\cos(\theta)}\Bigg]e^{-\frac{V}{2\sigma^2}}d\theta dV. \qquad (18)$$

Figure 1 and 2 show (8) as function of SNR. In Figure 1, $p = .2$ and in Figure 2, $p = 1$. From the figures it is clear that the theoretical curves match the simulation results. Also it is clear that for high SNR we can distinguish between PSK2 curves and PSK4 curves.

4. SUPPORT VECTOR MACHINE CLASSIFICATION

Support vector machine (SVM) is an empirical data 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 that all training data points belong to the right class. It is a two-class approach which can use multi-dimensions features. The two objectives of the SVC problem are then incorporated into an optimization problem. This is done by constructing the dual and primal problem of the classical Lagrangian problem with transferring the constraint of the second objective to become constraints on the Lagrange variables. The complete derivation of SVC is given in [7], [8].

5. SIMULATION AND DISCUSSION

We compare the performance of the proposed-feature classifier with different previously proposed classifiers. The classifiers include the maximum likelihood classifier proposed in [1], the qLLR classifier proposed in [2] and the modified cumulant based classifier proposed in [3]. In our cumulant based classifier a training set is used to construct the SVM classifier. This is different from what was used originally in [3] where the classification was based on theoretical asymptotic thresholds. Similarly, a training set was used to construct the proposed-feature classifier. The performances of these classifiers are shown in Figure 3 and 4.

To make a comprehensive comparison between the classifiers we need to determine the amount of information required for the classifiers to work from the receiver point of view. The receiver of the maximum likelihood classifier needs to know noise power and all signal parameters, i.e., carrier frequency, amplitude, symbol rate, initial phase, time offset and the values of the signal constellation points. The qLLR classifier works unaided by the knowledge of initial phase. Finally, in the case of cumulant based classifier and proposed-feature classifier, the values of the signal constellation points and the random initial phase (θ_c) are not needed.

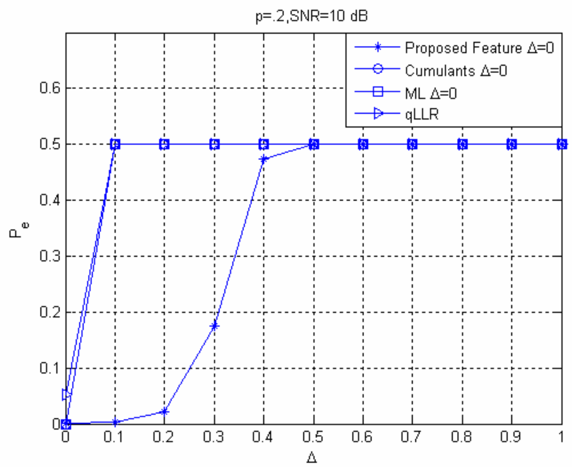

Figure 3: Probability of misclassification (P_e) for 2000 PSK2 and PSK4 signals. 1000 realizations for each modulation. "ML $\Delta = 0$" represents the maximum likelihood classifier designed for $\Delta = 0$, "Proposed Feature $\Delta = 0$" represents the SVM classifier whose input feature is the proposed feature and constructed based on $\Delta = 0$ and SNR= 10dB, and "Cumulants $\Delta = 0$" represents the SVM classifier whose input feature is the cumulants based feature proposed in [3] and constructed based on $\Delta = 0$ and SNR= 10dB . "qLLR" is the quasi log likelihood ratio classifier proposed in [2]. p=.2 and N=500.

Figure 3 shows the performance of the classifiers for different Δ. Out of the four classifiers the proposed feature classifier has the best performance. In the case of the cumulant based classifier, the $\sum_{k=0}^{TN-1} s^2(k)/TN$ term in the fourth order cumulant used to distinguish PSK2 from PSK4 causes the deterioration of performance when $\Delta \neq 0$. In the case of $\Delta = 0$ for noiseless PSK2 signal

$$\frac{1}{NT}\sum_{k=0}^{NT-1} s^2(k) = 1 \quad . \tag{19}$$

From the figures it is clear that for $\Delta = 0$ the maximum likelihood classifier has 0 probability of error. However as Δ increases the performance of the classifier deteriorates until it reaches .5 probability of misclassification.

Increasing Δ causes the qLLR classifier to deteriorate in performance. It should be noted that in the qLLR case even for $\Delta = 0$ the probability of error does not reach 0. The performance worsens as p increases. This happens since when we increase the value of p, we decrease the number of samples in the averaging process used in the qLLR classifier. This affects the approximation used in the algorithm.

The proposed feature classifier shows robust performance for small Δ. For example, in Figure 3 the proposed classifier has a probability of error less than .2 for $\Delta = 0.3$.

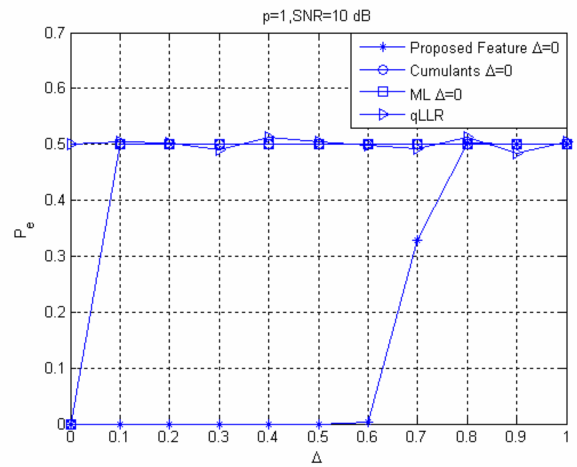

Figure 4: Probability of misclassification (P_e) for 2000 PSK2 and PSK4 signals. 1000 realizations for each modulation. Acronyms are the same as for Figure 3. p=1 and N=500.

6. CONCLUSION

In this paper we proposed a simple feature to distinguish between PSK2 and PSK4 modulations. The feature is based on the kurtosis estimate for the subtractions of two consecutive signal values. Following that we studied the performance of the SVM classifier whose input is the proposed feature for different Δ. Moreover, we compared the SVM classifier to other, previously proposed in literature, classifiers. The simulation results showed that the SVM classifier whose input is the proposed feature, is robust with respect to Δ when compared to the other classifiers.

7. REFERENCE

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

[2] A. Polydoros and K. Kim, "On the detection and classification of quadrature digital modulations in broad-band noise", *IEEE Transactions on Communications*, vol. 38, no. 8, Aug. 1990, pp.1199 – 1211.

[3] A. Swami and BM. Sadler, "Hierarchical digital modulation classification using cumulants", *IEEE Transactions on Communications,* vol. 48, no. 3, pp. 416-429, March 2000

[4] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework,* Prentice Hall, 1993.

[5] P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, New York, 1987.

[6] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[7] C. Burges," A Tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, 1998, pp. 121-167.

[8] S Gunn," Support Vector Machines for Classification and Regression," *ISIS Technical Report*, Image Speech and Intelligent Systems Group University of Southampton, 1998.