A PHASE LIKELIHOOD-BASED ALGORITHM FOR BLIND IDENTIFICATION OF PSK SIGNALS

Daimei Zhu, V. John Mathews

University of Utah
Department of Electrical and Computer Engineering
Salt Lake City, Utah 84112
Email: daimei.zhu@utah.edu (D. Zhu)
mathews@ece.utah.edu (V. J. Mathews)

David H. Detienne

Raytheon Applied Signal Technology
Salt Lake City, Utah 84119
Email: david.detienne@raytheon.com

ABSTRACT

This paper presents a phase likelihood-based method for automatically identifying different phase-shift keying (PSK) modulations. This method identifies the PSK signals as the hypothesis for which the likelihood function of phase difference between nearby samples of the received signal is the maximum. This method does not need prior knowledge of carrier frequency or baud rate and can identify modulation types at relatively low signal-to-noise ratio (SNR) and using small number of input samples. Simulation results demonstrate that this algorithm can identify BPSK, QPSK and 8PSK signals with 100% accuracy with only 1000 symbols when the SNR of the input signal is better than 7 dB. Additional simulation results demonstrating the robustness of the algorithm to variations of the noise characteristics from the assumed Gaussian model are also included in the paper. Performance comparisons indicate that the approach of this paper can achieve 100% accuracy in modulation identification at 5-7 dB lower SNR than competing methods available in the literature.

***Index Terms*—** Modulation identification, phase likelihood function, PSK

1. INTRODUCTION

Automatic identification of the modulation type of a received signal is important in many military and civil applications. Blind modulation identification, *i. e.*, modulation identification without *a priori* knowledge of the carrier frequency, symbol rate and other parameters of signal transmission, is addressed in this paper. We assume that the received signal is phase-shift keying (PSK) modulated, but do not assume any prior knowledge of the order of PSK modulation.

In [1], the authors proposed a method for blind modulation recognition of PSK signals based on constellation reconstruction. The results showed in the paper indicated that this method can identify BPSK, QPSK and 8PSK with 100% accuracy at 15 dB signal-to-noise ratio (SNR) with 500 symbols. In [2], the authors implemented the Morlet wavelet to obtain the phase for different PSK types including BPSK, QPSK and 8PSK. The method performed the modulation identification based on a likelihood function of phase parameters extracted from Morlet wavelet transform of the input signal. Simulation results indicated that this approach can identify BPSK, QPSK and 8PSK with 100% recognition rate when the SNR was more than 12 dB. A clustering-based distribution fitting algorithm was used for modulation identification in [3]. The methods assumed known carrier frequency as well as other channel parameters. A survey of modulation identification systems was given in [4]. Unfortunately, most of the comparisons in the paper assumed knowledge of signal characteristics such as carrier frequency. Such knowledge is impractical in most applications of modulation identification.

In recent work, the authors of this paper presented a likelihood-based algorithm for identifying QAM modulations from received signals without knowing its carrier frequency or baud rate [5]. Performance evaluation of the algorithm indicated that the method was able to identify the modulation types accurately at lower SNR and using shorter durations of the received signal than previously possible. In this paper, we present a method for blind identification of PSK signals by deriving the likelihood function associated with the phase difference of nearby samples. To the best of the authors' knowledge, phase likelihood-based algorithms have not yet been developed for PSK identification. The likelihood function of [2] is not directly evaluated for the signal phase.

As will be shown in Section 4 and Section 5, the method of this paper performs substantially better than those competing algorithms available in the literatures. In addition, our algorithm is robust to variations of the signal characteristics from the signal model assumed in the derivation of the algorithm. The rest of the paper is organized as follows: the received signal model with pulse shaping is described in the next section. In Section 3, the phase likelihood-based identification method is presented. Section 4 contains simulation results demonstrating the probability of correct modulation

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

identification in different SNRs and under several noise environments. Finally, Section 5 contains the concluding remarks.

2. SIGNAL MODEL

We assume an additive white Gaussian channel under which the general model for the received signal [6] is

$$y(t) = Re\{\sum_k (s_k g_T(t - kT_b))e^{j2\pi f_c t} + N_0(t)\} \quad (1)$$

where s_k is a complex symbol sequence with $s_k = a_k + jb_k$, where a_k and b_k are the real and imaginary parts, T_b is the symbol period, $g_T(t)$ is the pulse shape filter, f_c is the carrier frequency, and $N_0(t)$ is additive white Gaussian noise.

Applying Hilbert transformation to the received signals, an appropriately sampled version of this signal is given by

$$y(n) = \sum_k (s_k g_T(nT_s - kT_b))e^{j2\pi f_c nT_s} + N_0(nT_s) \quad (2)$$

where T_s is the sampling period. $N_0(nT_s)$ is the sampled version of noise, which is a band-limited white Gaussian noise. If we assume that the pulse shaping is such that the interference between the nearby symbols is negligible in a small interval around the midpoint of each baud [7], pulse shaping has little influence on the phase of the signal in these intervals. Let $y(m)$ represent the mid-point of the mth symbol.

Then,

$$y(m) = s_m g_T(0)e^{j2\pi f_c mT_b} + N_0(mT_b). \quad (3)$$

The contribution to the phase of $y(m)$ from the carrier frequency changes with m. Obtaining the phase difference between nearby samples of sub-sampled signal can avoid this variability. Let

$$y_d(m) = y(m)y^*(m - d) \quad (4)$$

where d is an appropriately selected lag value, then

$$\begin{aligned} y_d(m) &= \{s_m g_T(0)e^{j2\pi f_c mT_b} + N_0(mT_b)\} \\ &\quad \{s^*_{m-d} g^*_T(0)e^{-j2\pi f_c(m-d)T_b} + N^*_0((m - d)T_b)\} \\ &= s_m s^*_{m-d} g^2_T(0)e^{j2\pi f_c dT_b} + N_0(mT_b)N^*_0((m - d)T_b). \quad (5) \end{aligned}$$

The phase of $y_d(m)$ is

$$\theta_d(m) = \theta_s + \theta_c + \alpha \quad (6)$$

where $\theta_c = 2\pi f_c dT_b$ is the fixed value contributed by the carrier frequency f_c, θ_s is the phase difference between the symbols and α is the phase difference contributed by noise. In all the simulations in this paper, we selected d to be 1.

We can see from the above that, the phase difference between $y(m)$ and $y(m - d)$ will have a similar distribution (within a constant shift contributed by the carrier frequency) as the phase difference of the original symbol sequence. In the next section, we will derive the likelihood function of the phase difference between nearby samples of a PSK signal in a broadband Gaussian noise assuming one sample per symbol. In practice, we will compute the likelihood function of the phase sequence after appropriately sub-sampling the received signal so that the statistics of the signal approximately matches those in the derivations.

3. PHASE LIKELIHOOD FUNCTION AND PSK IDENTIFICATION

A schematic block diagram of the blind PSK identification algorithm is given in Figure 1. The algorithm first makes a coarse estimate of the baud rate, and then uniformly subsamples the received signal using the estimated baud rate. The likelihood functions of the observed phase differences of the sub-sampled signals are then calculated for each modulation type. The system identifies the signal modulation type as the corresponding hypothesis modulation type for which the log-likelihood function is the maximum.

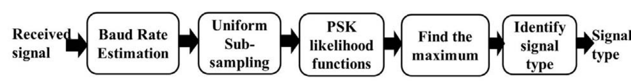


Fig. 1. Block diagram of modulation identification system.

3.1. Sub-sampling

In order to sub-sample the received signals, we need to estimate the baud rate. The algorithm for estimating the baud rate is the same as used in [5] and therefore not described here.

For the signal model in (2), the nearby samples most often correspond to the same symbol when the sampling rate is high. In such situation, the phase difference between nearby samples will be dominated by a fixed value contributed by the carrier frequency. This can be avoided by sub-sampling the received signals such that the nearby samples almost always come from different baud as shown in the signal model in (3). Thus we sub-sample the received signals in such a way that we pick one sample in each estimated symbol duration.

Starting with the maximum amplitude in the first several estimated symbol durations (4 in the simulation results presented), the system sub-samples the received signal at the rate of the estimated symbol duration (rounded to the nearest integer.) The phase differences between the adjacent samples of the sub-sampled signal are computed and input to the likelihood function calculation block. Additional phase difference samples may be computed by changing the starting point for the sub-sampling process. Since the phase values of the symbols are almost the same throughout each baud, using these additional calculations will mitigate the effect of noise on modulation identification performance. The number of the

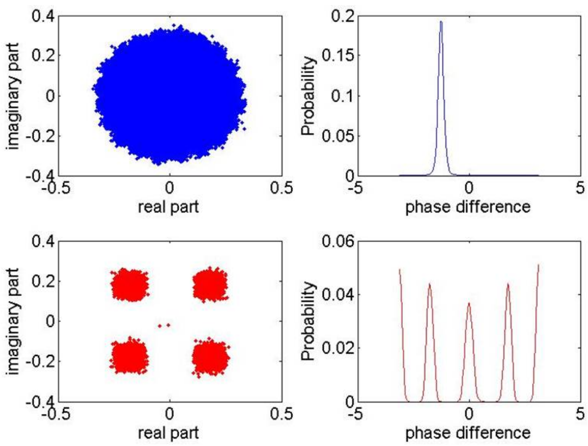


Fig. 2. "Scatter plot" and phase difference distribution of a QPSK signals before and after sub-sampling. Top left: "scatter plot" for received signal with pulse shaping: SNR=20 dB, N=10000; top right: histogram of phase difference for received signal with pulse shaping; bottom left: "scatter plot" after sub-sampling; bottom right: histogram of phase difference after sub-sampling

starting samples can be as many as the estimated symbol duration. The likelihood functions calculated in our simulations in Section 4 were based on the phase difference between adjacent samples of all such possible sub-sampled sequences.

Figure 2 shows the effects of the sub-sampling process on the phase difference sequences. The symbol sequence was independent and identically distributed in this example. The top left panel displays the scatter plot of the samples of the received signal. The top right panel shows the phase difference between the adjacent samples. The impact of the carrier frequency is clear in both figures. The effect of the pulse shaping is also seen in the top left. The corresponding results after sub-sampling by the estimated symbol duration are shown in the bottom panels. The four groups of phase difference between QPSK symbols can be observed after sub-sampling. Consequently, we can apply the likelihood functions derived for modulation identification of signals without pulse shaping to identify pulse shaped signals after the sub-sampling.

3.2. Phase Likelihood Function for PSK Signals

For two continuous sinusoids with the same frequency and initial phase that are independently perturbed by white Gaussian noise with variance σ^2 and zero mean value, the probability density function (PDF) of the phase difference α between them [8] is given by

$$P(\alpha) = \frac{1}{2\pi}\int_0^{\frac{\pi}{2}} \{\sin 2\beta[1 + \frac{S}{2}(1 + \cos\alpha \sin 2\beta)] \\ e^{-\frac{1}{2}S(1-\cos\alpha\sin 2\beta)}\} d\beta \tag{7}$$

where S is SNR.

The distribution of phase difference θ_d in (6) is

$$P(\theta) = \frac{1}{2\pi}\int_0^{\frac{\pi}{2}} \{\sin 2\beta[1 + \frac{S}{2}(1 + \cos(\theta - \theta_s - \theta_c)\sin 2\beta)] \\ e^{-\frac{1}{2}S(1-\cos(\theta-\theta_s-\theta_c)\sin 2\beta)}\} d\beta. \tag{8}$$

For the signal model in (2), the phase difference between symbols can take one of a finite number of values depending on the modulation order. Furthermore, assuming that the symbols sequence is independent and identically-distributed, the probability of these phase difference values can be predetermined for each PSK type.

Let there be N distinct phase difference values θ_s between symbols for the Mth modulation type. Let the set $\{\theta_s(M,i); i = 1, 2, \cdots, N\}$ represent these values and let $w_M[i]$ be the probability of the ith phase difference value for the Mth modulation type. The PDF for signal phase difference θ is

$$\begin{aligned} P(\theta) &= \sum_{i=1}^{N} P(\theta|\theta_s(M,i) w_M[i], -\pi \le \theta \le \pi, \\ &= \sum_{i=1}^{N}\int_0^{\frac{\pi}{2}} \{[1 + \frac{S}{2}(1 + \cos(\theta - \theta_s(M,i) - \theta_c)\sin 2\beta)] \\ &\quad \sin 2\beta e^{-\frac{1}{2}S(1-\cos(\theta-\theta_s(M,i)-\theta_c)\sin 2\beta)}\}\frac{w_M[i]}{2\pi} d\beta \end{aligned} \tag{9}$$

where $P(\theta|\theta_s(M,i)$ is the conditional PDF of the phase difference given that the modulation type is M.

Let H_M represent the hypothesis that the Mth modulation type is the modulation type of the received signal. Given n phase difference values $\theta_1, \theta_2, \ldots, \theta_n$ of the received signals, the likelihood function for H_M [5] is

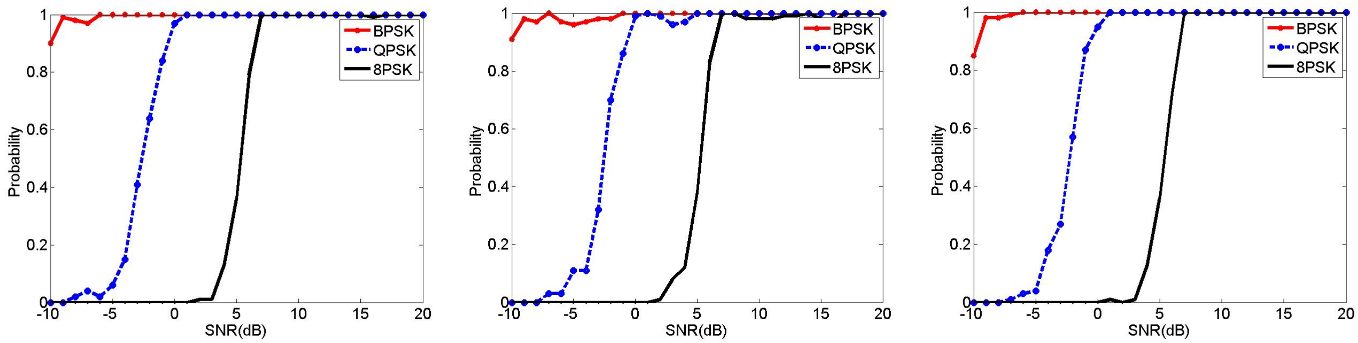
$$p(\theta_1, \theta_2, \ldots, \theta_n | H_M) = \prod_{i=1}^{n} P(\theta_i|H_M). \tag{10}$$

Combining (10) with (9), we obtain the likelihood function for the observed phase difference values given the hypothesis of the Mth modulation type as

$$p(\theta_1, \theta_2, \ldots, \theta_n | H_M) = \\ \prod_{j=1}^{n}\sum_{i=1}^{N}\int_0^{\frac{\pi}{2}} \{[1 + \frac{S}{2}(1 + \cos(\theta_j - \theta_s(M,i) - \theta_c)\sin 2\beta)] \\ \sin 2\beta e^{-\frac{1}{2}S(1-\cos(\theta_j-\theta_s(M,i)-\theta_c)\sin 2\beta)}\}\frac{w_M[i]}{2\pi} d\beta. \tag{11}$$

In order to simplify the calculations and since natural logarithm is monotonically increasing, we use the natural logarithm of (11) as the decision function for the identification problem. That is, we choose the modulation type that maximizes the log likelihood function given by

$$l_M = \sum_{j=1}^{n} \ln(P(\theta_j|H_M)). \tag{12}$$



(a) Additive Gaussian noise. (b) Additive uniform noise. (c) Additive Laplacian noise.

Fig. 3. Probability of correct modulation identification with different SNRs for pulse shaped signals corrupted by different noise; N=1000.

3.3. Estimation of Phase Contributed By The Carrier Frequency and Noise Variance

In order to apply (12), the SNR and θ_c must be estimated. θ_c has substantial influence on the PDF of the phase difference between samples. To ensure correct modulation identification, we consider θ_c as an unknown parameter that is estimated jointly with the modulation type by maximizing the log-likelihood function over the modulation type and θ_c.

With estimated SNR as described in [5], we can compute the log likelihood function for phase difference in (12). As will be shown in the next section, even for moderately small signal lengths (for example, 1000 symbols), the modulation identification algorithm performs with small or no errors above SNR values of 10 dB. Since the variable in the exponential part of the likelihood function as shown in (11) becomes very large for large SNR values, and can potentially drive the calculations to outside allowable ranges, we employed a regularization procedure in the calculations where SNR estimates above 15 dB were reset to 15 dB. This regularization provided good results while ensuring that the algorithm was operated in a stable manner.

4. PERFORMANCE EVALUATION

In this section, the performance of the algorithm is demonstrated by the probability of correctly identifying each modulation type under several SNR conditions. We also evaluate the performance of the algorithm in noise environments different from the assumed Gaussian model. In all the simulations, a root raised cosine filter with parameter $\beta = 0.5$ was applied to the transmitted symbol sequence, $N = 1000$ symbols were used, and 100 independent runs were used to calculate the probability of correct identification.

Figure 3 (a) shows the results when the noise is zero-mean white and Gaussian noise. The system can distinguish between BPSK and QPSK modulation with 100% accuracy at 1 dB SNR. For 8PSK, 100% correct identification occurs when SNR is at or above 7 dB.

Figures 3 (b) and (c) show the identification results in noise environments different from the assumed Gaussian model. The noise in Figure 3 (b) was zero-mean and uniformly distributed noise and the results in Figure 3 (c) were obtained with the Laplacian noise with zero mean value. We observe that the performances with uniform noise and Laplacian noise are comparable to that with Gaussian noise.

5. CONCLUSION REMARKS

The likelihood-based modulation identification algorithm presented in this paper performs substantially better than alternate methods available in the literature. The system identified BPSK, QPSK and 8PSK modulation with 100% accuracy at 7 dB SNR with 1000 symbols. Simulation results presented in [2] indicated that the method in that paper needed more than 12 dB SNR to identify BPSK, QPSK and 8PSK modulation with 100% accuracy with 1050 symbols. Our algorithm also is able to combat the effects of the pulse shaping in the signal model, which was not shown in [2]. Simulation results shown in [1] indicate that its method needed 15 dB SNR with 500 symbols for identifying BPSK, QPSK and 8PSK. In simulation results not included here, our method accurately identified all three PSK signals at or above 8 dB SNR using 500 symbols. Furthermore, the simulation results presented indicated that the performance was robust under different noise environments. Additional work on performance evaluation under a variety of impairments as well as algorithm refinements to reduce computational complexity and to improve performance is underway at this time.

Acknowledgment

The work done at the University of Utah was supported in part by a research contract with Raytheon Applied Signal Technology.

6. REFERENCES

[1] X. Tan, H. Zhang, Y. Shen and W. Lu, "Blind Modulation Recognition of PSK Signals Based on Constellation Reconstruction," *Wireless Communications and Signal Processing (WCSP), 2010 International Conference on*, pp. 1-6, Suzhou, China, Oct. 2010.

[2] D. Zhang and X. Wang, "MPSK Signal Modulation Recognition Based on Wavelet Transformation," *2009 International Conference on Networking and Digital Society*, pp. 203-205, Guiyang, Guizhou, China, May 2009.

[3] K.-T. Woo and C.-W. Kok, "Clustering Based Distribution Fitting Algorithm for Automatic Modulation Recognition," *Proc. IEEE Symposium on Computers and Communications*, pp. 13-18, Averio, Portugal, July 2007.

[4] O.A. Dobre, A. Abdi, Y. Bar-Ness and W. Su, "Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends", *Communications, IET*, pp. 137-156, vol. 1, No.2, Apr. 2007.

[5] D. Zhu and V. J. Mathews, "Likelihood-based Algorithms for Blind Identification of QAM Signals," *2013 Digital Signal Processing and Signal Processing Education Workshop*, Napa, CA, Aug. 2013.

[6] B. Farhang-Boroujeny, *Signal Processing Techniques for Software Radios*, LuLu Publishing House, 2010.

[7] J. Xi and Z. Wang, "MQAM Modulation Scheme Recognition Using Hilbert Transform," *Journal on Communications*, vol. 28, No. 6, June 2007.

[8] J. T. Fleck and E. A. Trabka, "Error Probabilities of Multiple-state Differentially Coherent Phase-shift Keyed Systems in The Presence of White, Gaussian noise," *Investigation of Digital Data Communication Systems*, Rep. UA-1420-S-1, J. G. Lawton, Ed., Cornell Aeronaut. Lab., Inc., Buffalo, NY, Jan. 1961, Detect Memo 2A; available as NTIS Doc. AD256584.

[9] R. F. Pawula, S. O. Rice, and J. H. Roberts, "Distribution of the Phase Angle between Two Vectors Perturbed by Gaussian Noise," *IEEE Transactions on Communications*, vol. COM-30, No. 8, pp. 1828-1841, Aug. 1982.

[10] Y. Yang, C.-H. Liu, and T.-W. Soong, "A Log-likelihood Function-based Algorithm for QAM Signal Classification," *Signal Processing*, vol. 70, No. 1, pp. 61-71, Oct. 1998