16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

# CLASSIFICATION OF LINEAR AND NONLINEAR MODULATIONS USING THE BAUM-WELCH ALGORITHM

*Anchalee Puengnim*(1), *Nathalie Thomas*(2), *Jean-Yves Tourneret*(2) *and Josep Vidal*(3)

(1) TéSA, 14-16 Port Saint Etienne, 31000 Toulouse, France
(2) University of Toulouse, IRIT/ENSEEIHT/TéSA, Toulouse, France
(3) Universitat Politécnica de Catalunya, Barcelona, Spain
anchalee.puengnim@tesa.prd.fr, {nathalie.thomas,jean-yves.tourneret}@n7.fr, pepe@gps.tsc.upc.es

## ABSTRACT

Nonlinear GMSK modulations are becoming new modulation standards for Telemetry/Telecommand satellite links. These new satellite systems will have to co-exist with other space systems using different modulation schemes. In this context, modulation recognition using the received communication signal is useful to identify a possibly perturbing system. This paper studies a Bayesian classifier to recognize BPSK, QPSK, 8PSK and two standardized GMSK modulations in the presence of additive white Gaussian noise. A modified version of the Baum-Welch (BW) algorithm is used to compute the posterior probabilities of the received signal, given each possible model, and to estimate the unknown model parameters. The posterior probabilities are then plugged into the optimal Bayes decision rule. The performance of the proposed classifier is assessed through several simulation results.

## 1. INTRODUCTION

Continuous Phase Modulations (CPMs) are non-linear constant amplitude modulations which are very interesting for satellite transmissions because of their high robustness to amplifier non-linearities. Moreover, with proper choice of pulse shape and parameters, CPMs allow one to obtain higher bandwidth efficiency than that obtained with the traditional MPSK schemes. These properties have motivated the use of CPMs in new modulation standards for satellite communications. For example, a particular CPM modulation, called GMSK (Gaussian Minimum Shift Keying), is the new modulation standard for telemetry/telecommand satellite links. Two different GMSK schemes characterized by two different bandwidths have been adopted by the consultative committee for space data system (CCSDS) for future space missions [1]. These new schemes will have to co-exist with other space systems using different modulation schemes. As a consequence, it is important to be able to identify the authorized and non authorized systems. Equivalently, the problem consists of recognizing the modulation associated to a received communication signal.

Various strategies have been proposed in the literature for the classification of linear modulations (the interested reader is invited to consult [2, 3] and references therein for details). However, classifying nonlinear modulations has received less attention in the literature. Several methods for classifying full response binary CPMs with rectangular pulse shape and different modulation indexes have been studied in [4, 5]. A new methodology for classifying the two non-linear GMSK modulations recommended by CCSDS was proposed in [6]. The classifier was based on a state trellis representation (exploiting the fact that the GMSK modulation is a modulation with memory) allowing the use of a modified version of the BW algorithm. The BW algorithm was used to estimate the posterior probabilities of the received modulated signal (conditionally to each class). These posterior probabilities were then plugged into the optimal Bayes decision rule.

However, the algorithm proposed in [6] assumed that nonlinear modulations were pre-identified from other linear modulation candidates, which is not always a simple task. This paper goes a step further and shows that linear modulations used in satellite systems (BPSK, QPSK, 8PSK) as well as the non-linear standardized GMSK modulation schemes can be identified using the same recognition process. The emitted linearly or nonlinearly modulated signals are assumed to be corrupted by an additive Gaussian noise whose variance is estimated by the BW algorithm. The performance of the proposed classifier is assessed through several simulation results.

This paper is organized as follows: Section 2 gives some useful information regarding the linear and nonlinear modulations considered in this study. Section 3 presents the model of received baseband communication signal and its associated first order hidden Markov model (HMM) (required for the BW algorithm). Section 4 recalls the main steps of the BW algorithm allowing one to estimate the posterior probability of the observation sequence given each possible modulation and the channel noise variance. Section 5 studies the performance of the MAP rule based on the posterior probabilities computed by the BW algorithm. Simulation results and conclusions are reported in Sections 6 and 7.

## 2. LINEAR AND NON-LINEAR MODULATIONS

The emitted signal $s(t)$ can be written as

$$s(t) = Re\left[\tilde{s}(t)e^{j\omega_c t}\right],$$

where $\tilde{s}(t) = I(t,\mathbf{a}) + jQ(t,\mathbf{a})$ is the complex envelope (or equivalent low-pass (LP) signal) and $\omega_c = 2\pi f_c$, where $f_c$ is the carrier frequency. The modulation is called *linear* when $\tilde{s}(t)$ linearly depends on the independent identically distributed (i.i.d.) complex symbol sequence $\mathbf{a} = \{a_k\}$ to be transmitted, and *nonlinear* in the other cases.

### 2.1 Linear M-PSK modulations

The baseband complex envelope of a linearly modulated signal can be written as $\tilde{s}(t) = \sum_k a_k h(t - kT)$, where $h(t)$

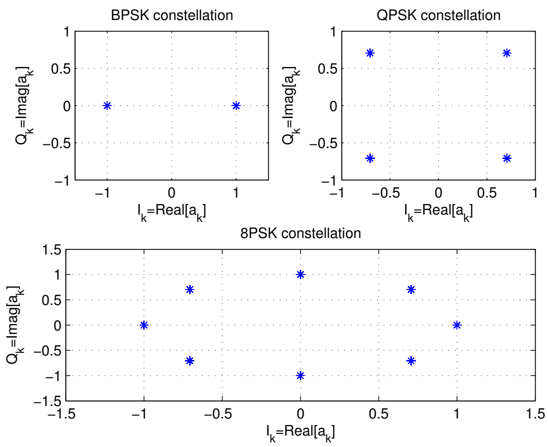


Figure 1: Classical linear modulation constellations.

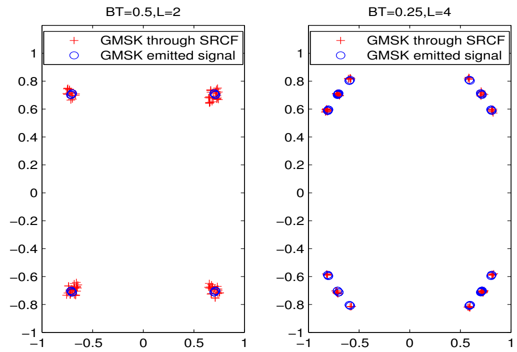


Figure 2: Emitted (circles) and received (crosses) GMSK *constellations* displayed for the two standardized cases $BT = 0.5, L = 2$ and $BT = 0.25, L = 4$.

is the pulse shape and $T$ represents the symbol duration. The i.i.d. complex symbol sequence $\mathbf{a} = \{a_k\}$ to be transmitted takes its values into a set of $M$ complex numbers $\{S_0, S_1, \ldots, S_{M-1}\}$ called *constellation* representing a particular modulation. MPSK modulations are defined by

$$S_m = \exp\left(j2\pi\frac{m-1}{M}\right), \quad m = 1, \ldots, M.$$

For instance, BPSK, QPSK and 8PSK constellations are displayed in Figure 1.

### 2.2 Non-linear GMSK modulations

The GMSK signal is a partial response CPM signal with modulation index $\frac{1}{2}$ and a smooth shape frequency pulse $g(t)$ of length $LT$, where $L \in \mathbb{N}$. The function $g(t)$ is the global impulse response of two consecutive filters. The first filter is rectangular of length $T$ whereas the second one is Gaussian with a normalized 3dB bandwidth $BT$:

$$g(t) = \frac{1}{2T}\left\{Q\left(2\pi B\frac{t-\frac{T}{2}}{\sqrt{\ln 2}}\right) - Q\left(2\pi B\frac{t+\frac{T}{2}}{\sqrt{\ln 2}}\right)\right\}.$$

where $Q(t) = \int_t^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{\tau^2}{2}) d\tau$. The complex envelope of the GMSK signal can be written as $\tilde{s}(t) = e^{j\Phi(t,\mathbf{a})}$ [7], where the transmitted i.i.d. symbol sequence $\mathbf{a} = \{a_k\}$ taken from $\{\pm 1, \pm 3, \ldots, \pm(M-1)\}$ is embedded in the time-varying phase:

$$\Phi(t,\mathbf{a}) = \pi \sum_k a_k q(t-kT),$$

where $q(t) = \int_{-\infty}^{t} g(\tau) d\tau$. For $t \in [kT, (k+1)T]$, the time-varying phase can be written $\Phi(t,\mathbf{a}) = \theta_k(t,\mathbf{a}) + \phi_k$ where

- $\theta_k(t,\mathbf{a}) = \pi \sum_{i=k-L+1}^{k} a_i q(t-iT)$ represents the changing part of the time-varying phase in $[kT, (k+1)T]$,
- $\phi_k = \left[\frac{\pi}{2} \sum_{i=-\infty}^{k-L} a_i\right] \mathrm{mod}(2\pi)$ is the cumulant phase (where $[x]\,\mathrm{mod}(2\pi)$ denotes the angle of $x$ modulo $2\pi$). It represents the constant part of the time-varying phase in $[kT, (k+1)T]$ and can be recursively computed as $\phi_{k+1} = \phi_k + \pi a_{k-L+1}$.

A *state* of the GMSK signal is classically defined at $t = kT$ as the vector $s_k = \{\phi_k, a_{k-1}, a_{k-2}, \ldots, a_{k-L+1}\}$ and corresponds to a specific value of the time-varying phase $\Phi(kT, \mathbf{a})$. The number of states is limited, giving a set of possible values for the complex envelop of the GMSK modulated signal taken at $t = kT$. This set can be assimilated to a kind of *constellation* as shown in Figure 2.

## 3. MODEL OF RECEIVED SIGNAL

### 3.1 Complex envelope

The emitted signal $s(t)$ is assumed to be corrupted by a white Gaussian noise $w(t)$ with power spectral density $N_0/2$ (AWGN satellite channel). Note that the associated complex baseband Gaussian noise process will be represented by $\widetilde{w}(t)$. The received signal $r(t)$ is first down-converted by the receiver to recover its complex envelope $\widetilde{r}(t)$. Figure 3 recalls the structure of the standard down-convertor used in this study. After down-conversion, the received baseband signal $\widetilde{r}(t)$ can be written as:

$$\widetilde{r}(t) = \widehat{I}(t,\mathbf{a}) + j\widehat{Q}(t,\mathbf{a}) = \frac{1}{2}\widetilde{s}(t) \otimes f(t) + z(t), \quad t \in \mathbb{R},$$

where $f(t)$ is the impulse response of the two LP filters, $z(t) = \widetilde{w}(t) \otimes f(t)$ is a normalized complex-valued additive Gaussian noise with variance $\sigma_z^2$ and "$\otimes$" denotes convolution.

Assuming a perfect synchronization between the emitter and the receiver, the complex envelope of the received modulated signal, sampled at one sample per symbol ($t = kT$), can be written as:

$$\widetilde{r}(k) = \frac{1}{2}\widetilde{s}(k) \otimes f(k) + z(k), \quad k = 1, \ldots, N_s, \qquad (1)$$

where $N_s$ is the number of symbols in the observation interval. In absence of noise, the received constellations for linear modulations are exactly the same as the emitted ones when the Nyquist criterion is satisfied. The situation is different when the emitted signals are GMSK signals. Figure 2 shows emitted and received *constellations* associated to the two standardized GMSK modulations in the case of square root raised cosine LP filters. Note that these constellations

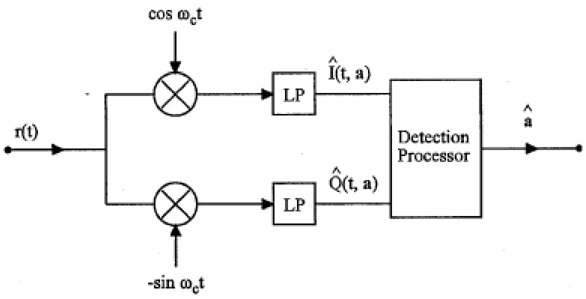


Figure 3: Basic quadrature receiver.

have been obtained in absence of noise, with a roll-off factor $\alpha = 0.35$ and a cutoff frequency adapted to symbol duration. Figure 2 indicates that the received signals corresponding to GMSK modulations and QPSK modulations are very similar. However, the classification rule proposed in this paper will allow us to distinguish these modulations.

### 3.2 Hidden Markov model

The received baseband signal $\widetilde{r}(k)$ can be modeled as a probabilistic function of an hidden state at time $k$ which is represented by a first order HMM model whose characteristics are summarized below:

- The state of the HMM at time instant $k$ is $s_k = a_k$ for MPSK modulated signals (memoryless linear modulation) whereas $s_k = (\phi_k, a_{k-1}, a_{k-2}, ..., a_{k-L+1})$ for GMSK modulated signals (non-linear modulation with memory). The state vector $s_k$ takes its values in a finite alphabet denoted as $\{s(1), s(2), ..., s(N)\}$ ($s(j)$ is the $j$th possible value of $s_k$). The size of this alphabet is $N = 4M^{L-1}$ for GMSK signals and $N = M$ for linear MPSK modulations.
- The state transition probability is defined by $d_{ij} = P[s_{k+1} = s(j)|s_k = s(i)]$ and equals $1/M$ when all symbols are equally likely.
- The initial state distribution vector $\pi = (\pi_1, ..., \pi_N)^T$ is defined by $\pi_i = P[s_1 = s(i)] = 1/N$ for $i = 1, \ldots, N$.
- Based on (1), the pdf of the observation $\widetilde{r}(k)$ conditioned on state $i$, denoted as $p_i(\widetilde{r}(k)) \triangleq p(\widetilde{r}(k)|s(i))$ can be written
$$p_i(\widetilde{r}(k)) = \frac{1}{\pi\sigma_z^2} \exp\left(-\frac{|\widetilde{r}(k) - m_i|^2}{\sigma_z^2}\right),$$
where $i = 1, ..., N$ and $m_i$ is the $i$th *constellation* point ($i$th possible value for $\frac{1}{2}\widetilde{s}(k) \otimes f(k)$). Note that $m_i = S_i$ for MPSKs when the transmitter and receiver filters are matched. For GMSK signals, $m_i$ is approximated by the $i$th value of $e^{j\Phi(kT,\mathbf{a})}$. We denote as $m = [m_1, ..., m_N]^T$ the vector containing all possible *constellation* points.

## 4. THE BW ALGORITHM

### 4.1 The standard BW algorithm

The BW algorithm proposed in [8] for speech recognition can be used to determine the posterior probability of the observation sequence $P(\tilde{r}|m, \sigma_z^2, \lambda)$, given a model $\lambda \in \{\lambda_1, \lambda_2, ..., \lambda_c\}$ (representing a modulation among the set of all $c$ possible modulations). The probability of the observation sequence for a given modulation is classically defined as a summation covering all possible state sequences. However, the direct computation of this summation requires high computational cost. The main idea of the BW algorithm is to use a forward-backward procedure which ensures a very efficient computation. The forward-backward procedure repeats the following three steps until convergence.

1. Computation of the normalized forward variable $\alpha_i(k)$.
Initialization:
$$\alpha_i(1) = \pi_i p_i(y(1)), \quad 1 \le i \le N,$$
$$c(1) = \left(\sum_{i=1}^{N} \alpha_i(1)\right)^{-1}.$$
Induction: for $k = 1, ..., N_s - 1$, $j = 1, ..., N$
$$\alpha_j(k+1) = c(k) p_j(y(k+1)) \sum_{i=1}^{N} \alpha_i(k) d_{ij},$$
$$c(k+1) = \left(\sum_{i=1}^{N} \alpha_i(k+1)\right)^{-1}.$$
2. Computation of the normalized backward variable $\beta_i(k)$.
Initialization: $\beta_i(N_s) = c(N_s)$, $1 \le i \le N$,
Induction: for $k = N_s - 1, ..., 1$, $i = 1, ..., N$,
$$\beta_i(k) = c(k) \sum_{j=1}^{N} d_{ij} p_j(y(k+1)) \beta_j(k+1),$$
3. Estimation of the model parameters
$$\widehat{m}_i = \frac{\sum_{k=1}^{N_s} \gamma_i(k) y(k)}{\sum_{k=1}^{N_s} \gamma_i(k)},$$
$$\widehat{\sigma}_z^2 = \frac{1}{N_s} \sum_{n=1}^{N_s} \sum_{i=1}^{N} \gamma_i(n) |m_i - y(n)|^2,$$
where $\gamma_i(k) = \alpha_i(k)\beta_i(k)$.

In a batch mode implementation, steps 1 to 3 are carried out iteratively with updated values of $p_j(y(k))$ until convergence. The posterior probability of the observation sequence given the model is then computed as follows

$$\widehat{P}(\widetilde{r}|m, \sigma_z^2, \lambda) = \frac{\sum_{i=1}^{N} \alpha_i(N_s)}{\sum_{i=1}^{N_s} c(i)}. \tag{2}$$

Different modifications have been applied to the standard BW algorithm to improve its performance or reduce computation complexity. One of these modifications is presented in Section 4.2.

### 4.2 The adaptive BW algorithm

An adaptive version of the BW algorithm was proposed in [9] to improve performance in terms of memory and computation speed. This LMS-type update algorithm is based on the following recursions

$$m_i(k) = m_i(k-1) + \mu_m \gamma_i(k) e_i(k),$$
$$\sigma_z^2(k) = (1-\mu_s)\sigma_z^2(k-1) + \mu_s \left( \sum_{i=1}^{N} \gamma_i(k) |e_i(k)|^2 \right),$$

where $e_i(k) = \widetilde{r}(k) - \widehat{m}_i(k-1)$ for $i = 1,...,N$. The initialization and time-induction steps for the forward variable can be computed as in the standard BW algorithm. The calculation of the backward variable can be obtained by using the fixed-lag or sawtooth-lag schemes [10]. In this paper, we have used the fixed-lag scheme as explained in [11].

## 5. CLASSIFICATION RULE

The classification rule used in this paper assigns the received signal $\widetilde{r}$ to the class $\lambda_i$ if

$$\widehat{P}(\widetilde{r}|\lambda_i)P(\lambda_i) \geq \widehat{P}(\widetilde{r}|\lambda_j)P(\lambda_j), \forall j = 1,...,c,$$

where $c$ is the number of possible modulations (or the number of classes) and $\widehat{P}(\widetilde{r}|\lambda_i) \triangleq \widehat{P}(\widetilde{r}|m, \sigma_z^2, \lambda_i)$ is obtained from (2). This strategy is sometimes referred to as plug-in maximum a posterior (MAP) rule [12]. It consists of replacing the class posterior probabilities in the optimal Bayesian classifier by their estimates. Note that the whole sequence of length $N_s$ is required to estimate $\widehat{P}(\widetilde{r}|\lambda_i)$ even if the online LMS-type update algorithm has been used for the computation of $\widehat{m}_i(k)$ and $\widehat{\sigma}_z^2(k)$. Note also that the observation length $N_s$ required to properly identify the different modulations should be greater than the maximum number of HMM states in the class dictionary to ensure that every possible state can be reached by the algorithm. This paper assumes that the different modulation formats are equally likely resulting in $P(\lambda_i) = 1/c$ for $i = 1,\ldots,c$.

## 6. SIMULATION RESULTS

Many simulations have been carried out to evaluate the performance of the proposed plug-in MAP classifier. This paper focuses on the classification of GMSK25, GMSK50, BPSK, QPSK and 8PSK modulations. All constellations have been normalized to unit energy. The signal to noise ratio per bit is defined as $\mathrm{E_b}/\mathrm{N_0}$ where $\mathrm{E_b}$ is the energy per bit at the input of the receiver. The classification performance is the average probability of correct classification defined as:

$$P_{\mathrm{cc}} = \frac{1}{c} \sum_{i=1}^{c} P\left[\text{assigning } \widetilde{r} \text{ to } \lambda_i | \widetilde{r} \in \lambda_i\right].$$

Tables 1-3 present the confusion matrices of the proposed classifier for different values of $\mathrm{E_b}/\mathrm{N_0}$ (the number of samples is $N_s = 500$ for these examples). It can be observed that the two GMSK signals as well as the MPSK signals can be distinguished even at very low values of $\mathrm{E_b}/\mathrm{N_0}$ (even if the *constellations* of GMSK and QPSK signals are very similar). However, to distinguish among linear modulations, the required operating $\mathrm{E_b}/\mathrm{N_0}$ is much higher especially when 8PSK modulations are present in the dictionary.

Figure 4 displays the classification performance as a function of $\mathrm{E_b}/\mathrm{N_0}$, for different values of the number of observations $N_s$. A good classification performance can be observed especially for small values of $\mathrm{E_b}/\mathrm{N_0}$ which are typical for satellite space communications. The effect of roll-of mismatch on classification performance was also studied. Figure 5 displays the classification performance for several values of the roll-off factor $\alpha$ of the square root raised cosine filters used in the receiver. The proposed classifier seems to be robust to roll-off mismatch. The last simulations study the effect of a phase offset obtained by rotating the constellation with an angle $\phi$ (this phase offset is due to synchronization errors at the receiver). Figure 6 shows that the classification performance seems to be robust to moderate synchronization errors.

## 7. CONCLUSIONS

This paper addressed the problem of classifying linear and nonlinear modulations transmitted through AWGN channels. The classification was achieved according to a MAP rule. An hidden Markov model was associated to the received baseband signal, allowing the use of the famous Baum-Welch algorithm to estimate the posterior probabilities of all possible modulations. Several simulations showed the good performance of the proposed classifier.

An interesting perspective is to recognize modulations in new satellite communication standards such as digital video broadcasting satellite handheld (DVB-SH). This standard uses QPSK, 8PSK, 16APSK modulations and orthogonal frequency division multiplexing (OFDM) whose automatic recognition is a challenging problem.

### REFERENCES


[1] Consulative Committee for Space Data Systems (CCSDS), *Radio Frequency and Modulation Systems*. CCSDS, 2001, no. 401.

[2] A. Swami and B. Sadler, “Hierarchical digital modulation classification using cumulants,” *IEEE Trans. Comm.*, vol. 48, no. 3, pp. 416–429, March 2000.

[3] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, “Survey of automatic modulation classification techniques: classical approaches and new trends,” *IET Communications*, vol. 1, no. 2, pp. 137–156, April 2007.

[4] C. Y. Huang and A. Polydoros, “Two small snr classification rules for cpm,” *Proc. IEEE Milcom*, vol. 3, pp. 1236–1240, October 1992.

[5] C. D. Chung and A. Polydoros, “Envelope based classification schemes for continuous phase binary frequency shift keyed modulation,” in *Proc. IEEE Milcom*, Fort Monmouth, NJ, Oct. 1994, pp. 796–800.

[6] A. Puengnim, N. Thomas, J. Tourneret, and H. Guillon, “Classification of GMSK signals with different bandwidths,” in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, Las Vegas, USA, April 2008, pp. 2013–2016.

[7] J. G. Proakis, *Digital Communications*. Mc Graw Hill, 2001.

[8] L. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[9] J. A. R. Fonollosa and J. Vidal, “Application of hidden markov models to blind channel characterization and data detection,” in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, vol. 4, Adelaide, Australia, April 1994, pp. 185–188.

[10] V. Krishnamurthy and J. Moore, “On-line estimation of hidden markov model parameters based on the kullback-leibler information measure,” *IEEE Trans. Signal Process.*, vol. 41, no. 11, pp. 2557–2573, 1993.

[11] A. Puengnim, T. Robert, N. Thomas, and J. Vidal, “Hidden Markov models for digital modulation classification in unknown ISI channels,” in *Proc. European Sig. Proc. Conf. (EUSIPCO)*, Poznan, Poland, September 2007, pp. 1882–1885.

[12] A. K. Jain, R. P. W. Duin, and J. Mao, “Statistical pattern recognition: a review,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, 2000.

| In/Out | GMSK25 | GMSK50 | BPSK | QPSK | 8PSK |
|---|---|---|---|---|---|
| GMSK25 | 449 | 51 | 0 | 0 | 0 |
| GMSK50 | 13 | 487 | 0 | 0 | 0 |
| BPSK | 0 | 0 | 500 | 0 | 0 |
| QPSK | 0 | 0 | 0 | 498 | 2 |
| 8PSK | 0 | 0 | 0 | 0 | 500 |

Table 1: Confusion matrix for $E_b/N_0$=0dB.

| In/Out | GMSK25 | GMSK50 | BPSK | QPSK | 8PSK |
|---|---|---|---|---|---|
| GMSK25 | 406 | 94 | 0 | 0 | 0 |
| GMSK50 | 46 | 454 | 0 | 0 | 0 |
| QPSK | 0 | 0 | 500 | 0 | 0 |
| 4QAM | 0 | 0 | 0 | 457 | 43 |
| 8PSK | 0 | 0 | 0 | 5 | 495 |

Table 2: Confusion matrix for $E_b/N_0$=−2dB.

| In/Out | GMSK25 | GMSK50 | BPSK | QPSK | 8PSK |
|---|---|---|---|---|---|
| GMSK25 | 334 | 164 | 1 | 0 | 1 |
| GMSK50 | 123 | 375 | 0 | 1 | 1 |
| BPSK | 0 | 0 | 488 | 4 | 8 |
| QPSK | 0 | 0 | 0 | 313 | 187 |
| 8PSK | 0 | 0 | 0 | 81 | 419 |

Table 3: Confusion matrix for $E_b/N_0$=−6dB.

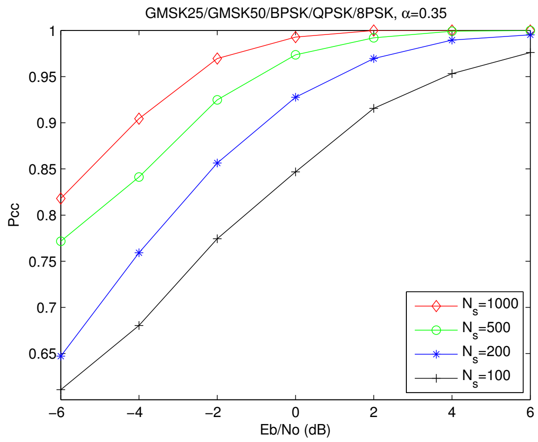


Figure 4: Classification performance versus $E_b/N_0$.

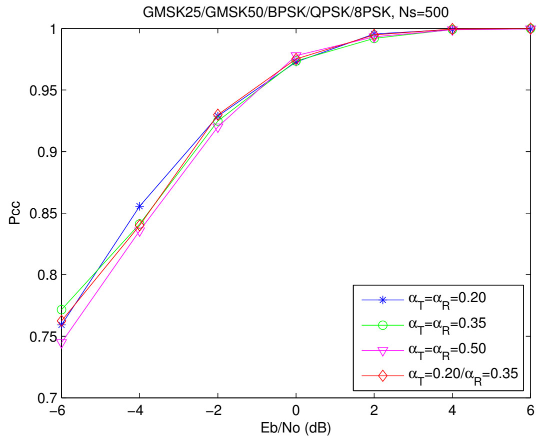


Figure 5: Classification performance versus $E_b/N_0$ for different values of the roll-off factor $\alpha$.

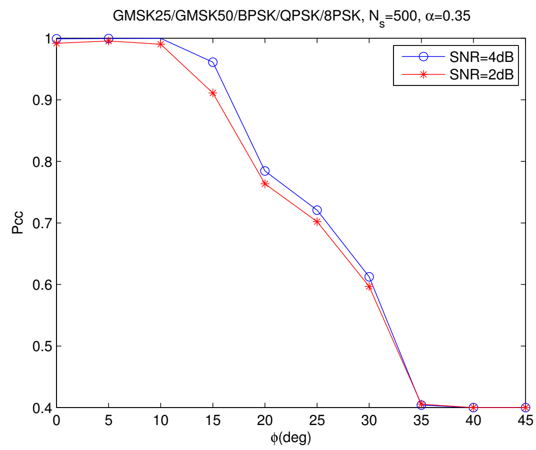


Figure 6: Classification performance versus phase offset.