

# The BICM Capacity of Coherent Continuous-Phase Frequency Shift Keying

Rohit Iyer Seshadri, Shi Cheng and Matthew C. Valenti  
Lane Dept. of Computer Sci. and Electrical Eng.  
West Virginia University  
Morgantown, WV 26506, USA  
{iyerr, shic, mvalenti}@csee.wvu.edu

**Abstract**—This paper presents a methodology for determining the capacity of coherently detected continuous-phase frequency shift keying (CPFSK) modulation under the constraints of binary coding and an ergodic channel. Building upon the capacity results, the coded CPFSK parameters of code rate, alphabet size, and modulation index are jointly optimized by using capacity as a cost function. From this optimization, it is possible to determine the minimum  $\mathcal{E}_b/N_0$  required for bit-interleaved coded CPFSK to achieve an arbitrarily low error rate as a function of spectral efficiency. Results showing this minimum  $\mathcal{E}_b/N_0$  are presented for a range of spectral efficiencies and several alphabet sizes in an AWGN channel.

## I. INTRODUCTION

*Continuous-phase frequency shift keying* (CPFSK) [1] is an attractive modulation choice due its small spectral side-lobes and constant envelope. In order to improve the energy efficiency, CPFSK can be combined with a channel code. There is however a fundamental tradeoff between the energy efficiency of the coded system and its spectral efficiency, since coding gain typically occurs at the expense of bandwidth.

In coded CPFSK, the power-bandwidth tradeoff is governed by the code rate  $r$  and the CPFSK parameters (the alphabet size  $M$  and the modulation index  $h$ ). For CPFSK (and more generally for continuous phase modulation (CPM)), signalling under bandwidth constraints involves a tradeoff between  $r$  and the modulation parameters. As an example, the spectral efficiency ( $\eta$ ) of CPFSK can be improved by placing the tones closer together which implies reducing the value of  $h$ . However, as  $h$  decreases, the error rate in general, increases [1]. The resulting performance loss may be overcome to some extent by using a code with lower rate  $r$ , which however decreases  $\eta$ .

In this paper, we attempt to better understand the above tradeoffs for coherently detected, bit-interleaved coded CPFSK. First, we outline a method to determine the capacity of the coded system using Monte-Carlo integration. This requires that the detector (demodulator) be capable of producing bit-wise log-likelihood ratios (LLRs) for the modulated symbols. From our capacity calculations, we identify the values of  $h$  and  $r$  that

minimize the  $\mathcal{E}_b/N_0$  required to achieve an arbitrarily low bit error rate, for a certain  $M$  and a required  $\eta$ . Thus we can determine what the minimum  $\mathcal{E}_b/N_0$  is at different values of  $\eta$  and the corresponding code rate  $r$  and CPFSK parameters. It is important to note that the above capacity is constrained not only by the choice of modulation and the bit to symbol mapping, but also by the demodulator formulation (coherent or noncoherent) and the channel (AWGN, Rayleigh etc.).

The remainder of this paper is organized as follows. Section II describes our system model. Since the capacity calculations require bit-LLRs from the detector, a soft-output coherent detector similar to [2] is employed. Section III deals with finding the constrained capacity and Section IV with the capacity under bandwidth constraints. Our results are presented and discussed in Section V. Finally, Section VI concludes the paper.

## II. SYSTEM MODEL

### A. Transmitter and Channel

A vector  $\mathbf{u} \in \{0, 1\}^K$  of information bits is passed through a linear binary encoder to produce the codeword  $\mathbf{b}' \in \{0, 1\}^N$ . The code rate is  $r = K/N$ . The row vector  $\mathbf{b}'$  is bit-interleaved to produce  $\mathbf{b}$ . The vector  $\mathbf{b}$  is arranged in a  $\log_2 M \times N_a$  matrix  $\mathbf{B}$  with  $(i, k)$  element  $B_{i,k} = b_{k \log_2 M + i}$ . The number of M-ary symbols to be transmitted is  $N_a = \lceil N / \log_2 M \rceil$ . Each column of  $\mathbf{B}$  is mapped to one of  $M$  symbols to produce the vector  $\mathbf{a} \in \{0, 1, \dots, M-1\}^{N_a}$  which represents the sequence of coded symbols to be transmitted. This process of concatenating a binary code with an M-ary modulator using a bit-interleaver is known as *bit-interleaved coded modulation* (BICM) [3].

Using Rimoldi's tilted phase representation [4], the CPFSK tilted phase during the interval  $nT \leq t < (n+1)T$  can be written as [4]

$$\varphi(t, \mathbf{a}) = [2\pi h V_n + \frac{2\pi h}{T} a_n(t - nT)] \mod 2\pi, \quad (1)$$

where  $T$  is the symbol period. Assuming the modulation index is rational and irreducible of the form  $h =$

$m_h/p_h$ ,  $V_n$  is

$$V_n = \left[ \sum_{i=0}^{n-1} a_i \text{ mod } p_h \right]. \quad (2)$$

According to [4], the modulator can be decomposed into a continuous phase encoder (CPE) followed by a memoryless modulator (MM). The CPE updates the MM input, such that

$$V_{n+1} = [V_n + a_n] \text{ mod } p_h. \quad (3)$$

From (1), the input to the MM which specifies the CPFSK phase during  $[nT, (n+1)T]$  is  $\mathbf{a}_n = [a_n, V_n]$ . Since  $\mathbf{a}_n$  can have one of  $p_h M$  possible values, the CPFSK signal will be one of  $p_h M$  possible signals at any symbol interval. The baseband transmitted signal can hence be written as

$$x(t, \mathbf{a}_n) = \sqrt{\frac{2\mathcal{E}_s}{T}} \exp(j\varphi(t, \mathbf{a}_n)), \quad (4)$$

where  $\mathcal{E}_s$  is the symbol energy.

The CPFSK signal is transmitted through an AWGN channel whose output is

$$r(t, \mathbf{a}_n) = x(t, \mathbf{a}_n) + n(t), \quad (5)$$

where  $n(t)$  is complex white Gaussian noise with zero mean and variance  $N_0$ .

### B. Soft-Output Coherent Detector

Soft output demodulation is performed by running the BCJR [5] algorithm on a trellis describing the underlying modulation. CPFSK can be represented as a finite state machine with  $p_h$  states [4], with the state at time instant  $t = nT$  given by  $S_n = V_n$ , and  $M$  branches emerging out of each state. Once the trellis is so defined, the LLR can be decomposed using the BCJR algorithm as

$$\mathcal{Z}_{i,n} = \log \frac{\sum_{\mathcal{S}^{(1)}} \alpha_n(s') \gamma_{n+1}(s', s) \beta_{n+1}(s)}{\sum_{\mathcal{S}^{(0)}} \alpha_n(s') \gamma_{n+1}(s', s) \beta_{n+1}(s)}, \quad (6)$$

where  $\mathcal{S}^{(1)}$  is the set of state transitions  $\{S_n = s'\} \rightarrow \{S_{n+1} = s\}$  corresponding to  $B_{i,n} = +1$ ,  $\mathcal{S}^{(0)}$  is defined similarly for  $B_{i,n} = 0$ , and  $\alpha$ ,  $\beta$  and  $\gamma$  are the metrics in the BCJR algorithm. Because [5] already fully describes how to recursively calculate  $\alpha$  and  $\beta$  from the branch metric  $\gamma$ , all that remains to completely describe the demodulator is a derivation of  $\gamma$ .

Because the interleaved code bits are equally likely, the metric  $\gamma_{n+1}(s', s)$  is defined as [5],

$$\begin{aligned} \gamma_{n+1}(s', s) &= P[r(t, \mathbf{a}_n) | (S_n \rightarrow S_{n+1}) = (s' \rightarrow s)] \\ &= P[r(t, \mathbf{a}_n) | x_{s \rightarrow s'}(t, \mathbf{c}_n)], \end{aligned} \quad (7)$$

where  $x_{s \rightarrow s'}(t, \mathbf{c}_n)$  is the CPFSK signal corresponding to the state transition  $\{S_n = s'\} \rightarrow \{S_{n+1} = s\}$  with the corresponding MM input being  $\mathbf{c}_n$ . It is pragmatic to use a sufficient statistic obtained from  $r(t, \mathbf{a}_n)$  in order

to evaluate the conditional pdf. In coherent receivers, the sufficient statistics may be provided by a bank of  $p_h M$  correlators, sampled at the end of every symbol duration. The output of the correlator during the  $n^{\text{th}}$  symbol interval is

$$\rho_n = \int_{nT}^{(n+1)T} r(t, \mathbf{a}_n) x_{s' \rightarrow s}^*(t, \mathbf{c}_n) dt. \quad (8)$$

From [2],  $\gamma_{n+1}(s', s)$  evaluates to

$$\gamma_{n+1}(s', s) \approx \exp(\text{Re}\{\rho_n\} / N_0). \quad (9)$$

Once the metrics are determined for every branch in the trellis, the demodulator executes the BCJR algorithm and using (6) produces the LLR  $\mathcal{Z}_{i,n}$  for each bit  $i$  of each symbol  $n$ . The LLRs are then placed into a row vector  $\mathbf{z}$  such that  $z_{n \log_2 M + i} = \mathcal{Z}_{i,n}$ . The vector is then deinterleaved and the resulting sequence  $\mathbf{z}'$  fed to the channel decoder for soft-decision decoding.

### III. BICM CAPACITY OF COHERENT CPFSK

Let  $X$  and  $Y$  be random variables denoting the input and output of a channel, respectively. The (Shannon) channel capacity is defined as the mutual information between  $X$  and  $Y$ , maximized over all possible input distributions  $p_X(x)$

$$C = \max_{p_X(x)} I(X; Y). \quad (10)$$

$I(X; Y)$  is the (average) mutual information between  $X$  and  $Y$  which can be written as [6]

$$I(X; Y) = \int \int p_{X,Y}(x, y) \log \frac{p_{X,Y}(x, y)}{p_X(x)p_Y(y)} dx dy, \quad (11)$$

where  $p_{X,Y}(x, y)$  is the joint density function between  $X$  and  $Y$  and  $p_Y(y)$  is the marginal density function of  $Y$ . As in [7], we define the *mutual information random variable*

$$i(X; Y) = \log \frac{1}{p_X(X)} + \log p_{X|Y}(X|Y), \quad (12)$$

where  $p_{X|Y}(X|Y)$  is the conditional pdf of  $X$  given  $Y$ . When  $X$  is discrete, the pdfs in the above expressions are replaced with the corresponding probability mass functions (pmfs).

The mutual information (11) can also be written as the expectation

$$I(X; Y) = E[i(X; Y)]. \quad (13)$$

An important consequence of writing the mutual information as an expectation is that it can now be accurately estimated using Monte Carlo integration with a large number of trials. This allows us to find  $I(X; Y)$  and hence  $C$  even in cases where they cannot be evaluated in closed form.

The mutual information in (10) is maximized when  $X$  is Gaussian. In a practical system however, the input

distribution is constrained by the choice of the modulation. Hence the capacity under practical modulation constraints is simply

$$C = I(X; Y), \quad (14)$$

where  $p_X(x)$  is now constrained by the choice of modulation.

As discussed in [3], BICM transforms the composite channel into  $\log_2 M$  parallel binary-input, continuous-output channels, which we call *BICM subchannels*. With a sufficiently large and random bit interleaver between the encoder and modulator, these subchannels will be independent. Because the capacity of parallel channels adds, the capacity of the overall BICM system is

$$C = \sum_{i=1}^{\log_2 M} C_i, \quad (15)$$

where  $C_i$  is the capacity of the  $i^{\text{th}}$  BICM subchannel. The capacity  $C_i$  is the mutual information between input  $b'_i \in \{0, 1\}$  and the corresponding LLR  $z'_i$  at the output of the demodulator,

$$C_i = E[i(b'_i; z'_i)]. \quad (16)$$

Using (12) and the fact that  $b'$  is discrete, we get

$$i(b'; z') = \log \frac{1}{P[b']} + \log P[b'|z'], \quad (17)$$

From [8], (17) evaluates to

$$i(b'; z') = \log 2 - \max^*(0, z'(-1)^{b'}), \quad (18)$$

where  $\max^*(x, y) = \log(e^x + e^y)$  as defined in [9]. Substituting (18) into (15) gives the expression for the BICM capacity

$$C = \sum_{i=1}^{\log_2 M} \left( \log 2 - E \left[ \max^*(0, z'_i(-1)^{b'_i}) \right] \right), \quad (19)$$

which is in units of nats per channel use. To convert to bits (19) must be divided by  $\log 2$ .

Fig. 1 shows  $C$  as a function of  $\mathcal{E}_s/N_0$  for 2-CPFSK at different  $h$ , in AWGN. Monte Carlo integration with 2 million symbols per  $\mathcal{E}_s/N_0$  were used to evaluate  $C$ . It is also possible to plot the information-theoretic minimum  $\mathcal{E}_b/N_0$  required for reliable signalling as a function of the code rate  $r = C$  since  $\mathcal{E}_b/N_0 = \mathcal{E}_s/N_0(r \log_2 M)$ . Fig. 2 shows the minimum required  $\mathcal{E}_b/N_0$  as a function of  $r$  for the same CPFSK parameters and channel as in Fig. 1. It is interesting to note from Fig. 2 that going to a lower  $r$  does not necessarily improve energy efficiency, this is a characteristic of CPFSK (and CPM in general) with BICM. One can see that for each choice of  $h$ , there is a particular value of  $r$  that minimizes the required  $\mathcal{E}_b/N_0$ . Furthermore, the results shown in Figs. 1 and 2 are for  $M = 2$  and would have to be repeated for all other  $M$ . Intuitively, the CPFSK design point that achieves capacity by

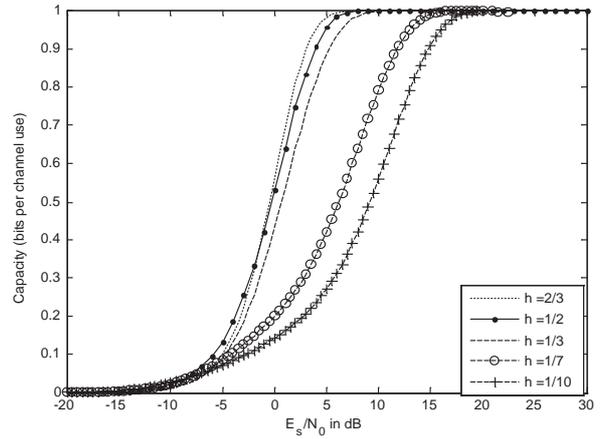


Fig. 1. Capacity versus  $\mathcal{E}_s/N_0$  for coherent CPFSK in AWGN with  $M = 2$  and modulation indices  $h = \{1/10, 1/7, 1/3, 1/2, 2/3\}$ .

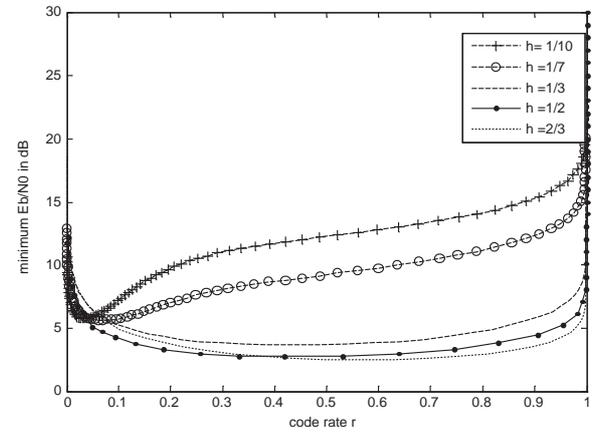


Fig. 2. Minimum  $\mathcal{E}_b/N_0$  required for coherent CPFSK to achieve an arbitrarily low error rate versus code rate  $r$  in AWGN with  $M = 2$  and modulation indices  $h = \{1/10, 1/7, 1/3, 1/2, 2/3\}$ .

maximizing the average mutual information is  $h = 1$  and  $M \rightarrow \infty$ . This is not a practical design because the infinite value of  $M$  will result in unbounded system complexity and the simultaneous use of orthogonal modulation will result in unbounded bandwidth.

Finally, it must be pointed out that the minimum  $\mathcal{E}_b/N_0$  obtained from the capacity is a very practical indicator of system performance due to the availability of “off-the-shelf” capacity-approaching, binary codes. Fig. 3 shows the bit error rate (BER) of bit-interleaved coded 2-CPFSK at modulation indices  $\{1/10, 1/7, 1/2\}$  using a rate  $1/2$  CDMA-2000 turbo code [10], after 10 decoder iterations with an interleaver length of 12282 bits, in AWGN. The vertical lines denote the information theoretic thresholds to achieve an arbitrarily low BER for the particular modulation index at  $r = 1/2$ .

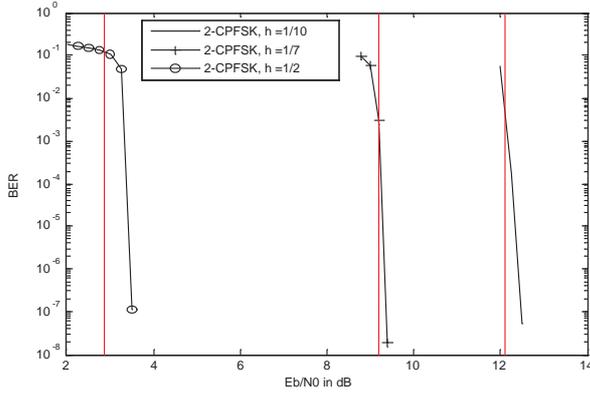


Fig. 3. Bit error rate in AWGN for bit-interleaved coded, 2-CPFSK with  $h = \{1/10, 1/7, 1/2\}$  using a rate  $1/2$  CDMA 2000 code after 10 turbo decoder iterations. The vertical lines denote the information theoretic  $\mathcal{E}_b/N_0$  in dB to achieve an arbitrarily low BER for the respective  $h$  and  $r = 1/2$ . The interleaver size is 12282 bits.

#### IV. CAPACITY UNDER BANDWIDTH CONSTRAINTS

The first step in quantifying the energy-bandwidth tradeoff is computing the bandwidth of the CPFSK signal. The power spectral density (PSD)  $\Phi_x(f)$  of the CPFSK signal  $x(t)$  is given in Section 4.4.2 of [11]. Using the PSD, the 99% power bandwidth  $B_{99}$  of  $x(t)$  is defined as

$$\int_{-B_{99}/2}^{B_{99}/2} \Phi_x(f) df = 0.99 \int_{-\infty}^{\infty} \Phi_x(f) df. \quad (20)$$

This bandwidth is a function of  $M$ ,  $h$ , and the symbol rate  $R_s = 1/T$ . The *normalized bandwidth* is defined to be  $B(M, h) = B_{99}T_b$  Hz/ baud, where  $T_b = T/\log_2 M$ . We can then define the *spectral efficiency*  $\eta = r/B(M, h)$ , which has units of bits-per-second-per-Hz (bps/Hz).

Next, one must determine the minimum value of  $\mathcal{E}_b/N_0$  for a desired spectral efficiency  $\eta$ . Due to the bandwidth constraint, the range of  $r$  that may be considered is restricted such that the threshold  $r'$  on code rate is given by

$$r' = \eta B(M, h), \quad (21)$$

where  $r \in [r', 1]$ . It is obvious that rates  $r < r'$  cannot be considered because for the particular  $h$  and  $M$ , as the spectral efficiency will be lower than  $\eta$ . Under tight bandwidth constraints, the optimal  $r$  is equal to the threshold value  $r'$ , but in looser bandwidth constraints the optimal  $r$  might be higher due to the non-monotonous nature of the  $\mathcal{E}_b/N_0$  versus  $r$  curves.

#### V. RESULTS AND DISCUSSION

$\mathcal{E}_b/N_0$  versus  $r$  curves were generated for  $M = \{2, 4, 8, 16\}$  in AWGN for modulation indices ranging from 0.1 to 0.9. For non-binary modulation, natural and gray labelling were considered. Next, for these CPFSK parameters, the minimum rate  $r'$  was determined from

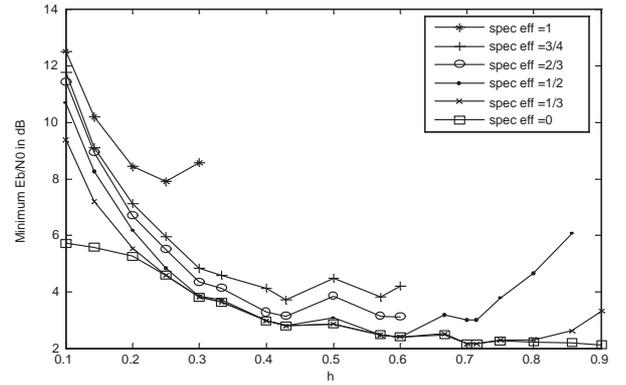


Fig. 4. Minimum  $\mathcal{E}_b/N_0$  required for coherent CPFSK to achieve an arbitrarily low error rate versus modulation index  $h$  in AWGN with  $M = 2$  for spectral efficiencies  $\eta = \{0, 1/3, 1/2, 2/3, 3/4, 1\}$ . For fixed  $h$ , the minimum  $\mathcal{E}_b/N_0$  increases with  $\eta$ .

(21). For example, when  $M = 2$  and  $\eta = 1/2$  bps/Hz, the minimum values of  $r$  are 0.4, 0.59, 0.85 and 0.93 for  $h = 1/5, 1/2, 2/3$  and  $3/4$ , respectively. Since  $B(M = 2, h = 1) = 2.1309 > 1/\eta$ , no code of rate  $r \leq 1$  can be used at this  $\eta$  when  $h = 1$  and thus orthogonal modulation cannot be considered. Next, the minimum  $\mathcal{E}_b/N_0$  was found by inspecting the curve over the range of possible rates  $r \in [r', 1]$ . For a given  $\eta$  and  $M$ , this procedure was repeated for each value of  $h$  over a range  $(0, h')$ , where  $h' = \max_{h \leq 1} : B(M, h) \leq 1/\eta$  is a maximum modulation index. At low spectral efficiency,  $h' = 1$  but at high spectral efficiency, values of  $h > h'$  cannot be used because the bandwidth requirement cannot be met for any code rate  $r \leq 1$ . The minimum  $\mathcal{E}_b/N_0$  over possible values of  $h$  was determined. As an example, Fig. 4 shows the minimum  $\mathcal{E}_b/N_0$  versus the corresponding  $h$  for  $M = 2$  in AWGN and several values of  $\eta$  (the  $\eta = 0$  case corresponds to having no bandwidth constraint).

Fig. 4 reveals that for each value of  $\eta$  there is an optimal choice of  $h$  that minimizes  $\mathcal{E}_b/N_0$ . For the unlimited bandwidth case ( $\eta = 0$ ), the optimal  $h$  tends to 1, but as  $\eta$  increases, the optimal value of  $h$  decreases. The combination of  $\eta$  and the  $\mathcal{E}_b/N_0$  minimized over  $h$  is the constrained channel capacity for that value of  $M$ , channel (AWGN), and coherent detection. It is interesting to note that the popular MSK (2-CPFSK,  $h = 1/2$ ) is not the optimum choice at any spectral efficiency. A plot of minimum  $\mathcal{E}_b/N_0$  versus  $h$  for all  $M \leq 16$  and  $\eta = 1/2$  is shown in Fig. 5 for the AWGN channel.

By finding the minimum value of  $\mathcal{E}_b/N_0$  with respect to  $h$  for each  $M$  over a wide range of  $\eta$ , one can finally determine the capacity of CPFSK. Capacity can now be plotted in terms of spectral efficiency  $\eta$  versus the corresponding minimum  $\mathcal{E}_b/N_0$ , as shown for different  $M$  and  $\eta \leq 1$  in Fig. 6 in AWGN. Note that as expected, the minimum  $\mathcal{E}_b/N_0$  in dB increases with  $\eta$ . While there

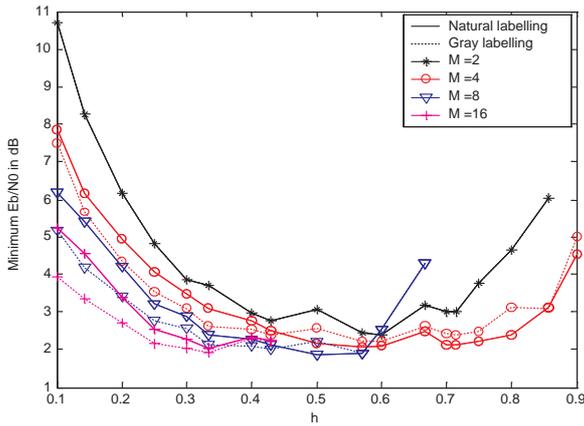


Fig. 5. Minimum  $\mathcal{E}_b/N_0$  required for coherent CPFSK to achieve an arbitrarily low error rate versus modulation index  $h$  in AWGN for modulation orders  $M = \{2, 4, 8, 16\}$  and spectral efficiency  $\eta = 1/2$ .

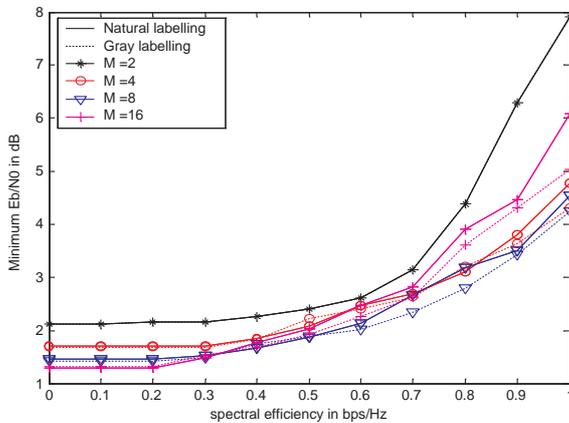


Fig. 6. Minimum  $\mathcal{E}_b/N_0$  required for coherent CPFSK to achieve an arbitrarily low error rate versus spectral efficiency  $\eta$  in AWGN for modulation orders  $M = \{2, 4, 8, 16\}$ .

is a benefit to increasing  $M$  at very low  $\eta$ , these benefits diminish as  $\eta$  is increased. There appears to be little benefit to using  $M > 4$  at higher spectral efficiencies. Also, Gray labelling is preferable to natural labelling under tight bandwidth constraints.

## VI. CONCLUSION

We have investigated the problem of designing bit-interleaved coded CPFSK systems under bandwidth constraints. A methodology has been outlined for solving the above problem using the constrained capacity as the cost function. Monte Carlo integration with a large number of trials is used to reliably determine the capacity over several modulation indices and alphabet sizes. The constrained capacity also takes into account the detector design.

Because many “off-the-shelf” codes are binary, it is convenient to use bit-interleaved coded modulation. The

key issue when extending BICM to nonorthogonal FSK is that the performance will be sensitive to how bits are mapped to symbols. Hence the optimization would have to be performed over all bit-to-symbol labelling. In this paper, we have limited ourselves to natural and gray labelling. Our results indicate at high spectral efficiencies, it is advantageous to use  $M = 4$  and gray labelling, whereas at low spectral efficiencies, there is a slight benefit of using a larger alphabet size and natural labelling.

The soft-output coherent detector has excellent energy efficiency. However, its main drawback is the high complexity involved in accurately estimating the signal phases at the receiver. The detector complexity also increases as  $h$  is reduced. Furthermore, trellis representation constrains the modulation index be a rational number [4], which restricts our design space. Noncoherent [12] and differential phase detectors [13] provide low complexity alternatives, albeit at the expense of energy efficiency.

An alternative approach to coded CPFSK design is to match the code alphabet to the modulation alphabet. The resulting scheme is known as *coded modulation* (CM), which due to the data processing inequality, has a higher capacity than BICM. Determining the CM capacity for coherent CPFSK is an open problem and a topic for future research. We have previously found the optimal parameters for noncoherently detected CPFSK using coded modulation in [12].

## REFERENCES

- [1] J. B. Anderson, T. Aulin, and C. E. Sundberg, *Digital Phase Modulation*. New York: Plenum Press, 1986.
- [2] P. Moqvist and T. Aulin, “Serially concatenated continuous phase modulation with iterative decoding,” *IEEE Trans. Commun.*, vol. 49, pp. 1901–1915, Nov. 2001.
- [3] G. Caire, G. Taricco, and E. Biglieri, “Bit-interleaved coded modulation,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.
- [4] B. E. Rimoldi, “A decomposition approach to CPM,” *IEEE Trans. Inform. Theory*, vol. 34, pp. 260–270, Mar. 1988.
- [5] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.
- [6] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [7] R. G. Gallager, *Information Theory and Reliable Communications*. Wiley, 1968.
- [8] R. I. Seshadri and M. C. Valenti, “Capacity-based optimization of bit-interleaved coded CPM with differential detection,” *IEEE Trans. Veh. Technol.*, 2008.
- [9] A. Viterbi, “An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes,” *IEEE J. Select. Areas Commun.*, vol. 16, pp. 260–264, Feb. 1998.
- [10] Third Generation partnership project 2 (3GPP2), “Physical layer standard for cdma2000 spread spectrum systems, release c,” *3GPP2 C.0002-C Version 1.0*, pp. 115–122, May 2002.
- [11] J. Proakis, *Digital Communications*. New York: McGraw-Hill, Inc., fourth ed., 2001.
- [12] S. Cheng, R. I. Seshadri, M. C. Valenti, and D. Torrieri, “The capacity of noncoherent continuous-phase frequency shift keying,” in *Proc., Conf. on Info. Sci and Sys.*, (Baltimore, MD), Mar. 2007.
- [13] I. Korn, “GMSK with differential phase detection in the satellite mobile channel,” *IEEE Trans. Commun.*, vol. 38, pp. 1980–1986, Nov. 1990.