

The Capacity of Noncoherent Continuous-Phase Frequency Shift Keying

Shi Cheng, Rohit Iyer Seshadri, and Matthew C. Valenti

West Virginia University

Morgantown, WV

Email: {shic, iyerr, valenti}@csee.wvu.edu

Don Torrieri

U.S. Army Research Lab

Adelphi, MD

Email: dtorrieri@arl.army.mil

Abstract—Capacity analysis is used to determine the fundamental tradeoff between power and spectral efficiency for coded continuous-phase frequency shift keying (CPFSK). The modulation may use an arbitrary modulation index h and the number of tones M may be any power of two. Detection is noncoherent and the channel is either AWGN or fully-interleaved (ergodic) Rayleigh fading. Numerical results demonstrate the advantage of using nonorthogonal modulation and multiple tones when the bandwidth requirement is tight.

I. INTRODUCTION

Continuous-phase frequency-shift keying (CPFSK) is an attractive choice of modulation due to its well-behaved spectral characteristics, near-unity peak-to-average power ratio, and ability to be noncoherently detected. When spectral efficiency is of little concern, orthogonal CPFSK with a large number of tones (M) can be used. The capacity of noncoherent orthogonal modulation is discussed by Stark in [1], and numerical results for the binary case are presented ($M = 2$). In [2], we present further numerical results for multi-tone modulation ($M > 2$) and compare the so-called *coded modulation* (CM) capacity against the *bit-interleaved coded modulation* (BICM) capacity [3]. In the same paper, we propose using BICM with *iterative decoding* (BICM-ID) [4] to close the gap between the BICM and CM capacities, but make no attempt to optimize the code design. In [5], Guillén i Fàbregas and Grant also consider using BICM-ID with noncoherent orthogonal modulation, and furthermore optimize the code with respect to the EXIT chart [6].

A drawback of orthogonal modulation is its poor spectral efficiency. To improve the spectral efficiency of CPFSK, the tones can be placed closer together, which results in a modulation index h that is less than unity. As h decreases, the system becomes more spectrally efficient but the uncoded error rate will generally increase [7]. Some of the performance loss can be regained with the use of a rate r error correcting code, but the use of a code will decrease the spectral efficiency. For a system characterized by a particular choice of demodulator (e.g. coherent or noncoherent), channel (e.g. AWGN or Rayleigh fading), alphabet size M , and required spectral efficiency η , there will be an information-theoretic “optimal” value of the two-tuple (h, r) . By setting h and r to their optimal values, the value of \mathcal{E}_b/N_0 required to achieve an arbitrarily low error rate at spectral efficiency η is minimized.

A plot of η versus \mathcal{E}_b/N_0 is the constrained channel capacity of the system, where the constraints are (1) the use of CPFSK modulation, (2) the number of tones M , (3) the demodulator choice (i.e. a symbol-by-symbol noncoherent detector), and (4) the channel type (AWGN or fading).

This paper outlines a methodology for computing the capacity of CPFSK modulation with noncoherent detection. We begin in Section II by presenting a model for CPFSK modulation with symbol-by-symbol noncoherent detection, which is followed in Section III by a discussion of capacity under modulation constraints with an emphasis on CPFSK modulation. The capacity of noncoherent CPFSK under bandwidth constraints is presented in Section IV, and the paper concludes in Section V.

For the remainder of this paper, bold lowercase letters will be used to denote (column) vectors, e.g. \mathbf{x} , and bold uppercase will be used for matrices, e.g. \mathbf{X} . The scalar value $x_{i,j}$ is used to denote the $(i, j)^{th}$ entry of the matrix \mathbf{X} , while the scalar value x_i is used to denote the i^{th} element of the vector \mathbf{x} . All matrices and vectors are indexed starting at zero, $\mathbf{x} = [x_0, x_1, \dots, x_{M-1}]^T$. Matrices may be represented as a row of column vectors, e.g. $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1}]$.

II. NONCOHERENT CODED CPFSK

A set of K data bits is passed through a channel encoder defined over the alphabet $\mathcal{M} = \{0, 1, \dots, M-1\}$. The output of the encoder is a length N vector \mathbf{q} of M -ary symbols. The code rate is $r = K/N \leq \log_2 M$ information bits per symbol. Because the code alphabet is matched to the modulation alphabet, the results presented in this paper are for the *coded modulation* (CM) capacity. Alternatively, the data could first be passed through a binary encoder, bitwise interleaved, and then mapped to M -ary symbols, but in that case the results would be for the *bit-interleaved coded modulation* (BICM) capacity. Due to the data processing inequality, the BICM capacity is always lower than the CM capacity [3].

For every entry of \mathbf{q} , the continuous time modulated signal $x_i(t)$ is chosen as the q_i^{th} signal of the set $\mathcal{S} = \{s_k(t), k = 0, 1, \dots, M-1\}$, where

$$s_k(t) = \frac{1}{\sqrt{T_s}} e^{\frac{j h \pi (2k - (M-1)) t}{T_s}}, \quad t \in [0, T_s), \quad (1)$$

and h is the *modulation index*. In order to produce a more compact bandwidth, an additional phase ϕ_i is applied so that

the phase transition from symbol to symbol is continuous. This additional phase is accumulated as

$$\phi_{i+1} = \phi_i + (2q_i - (M - 1))\pi h \quad (2)$$

During the interval $iT_s \leq t \leq (i+1)T_s$, the received signal is

$$y_i(t) = a_i e^{j\theta_i} \sqrt{\mathcal{E}_s} e^{j\phi_i} x_i(t) + n_i(t), \quad (3)$$

where $n_i(t)$ is additive white Gaussian noise (AWGN) with noise spectral density N_0 , \mathcal{E}_s is the energy per symbol, and $a_i e^{j\theta_i}$ is a complex fading coefficient. In this paper, we consider two types of channel, AWGN and fully interleaved (ergodic) Rayleigh fading. In both channels, θ_i is uniformly distributed over the range $[0, 2\pi)$. For the AWGN channel $a_i = 1, \forall i$, while in the Rayleigh fading channel, the a_i 's are independent and identically distributed (i.i.d.) Rayleigh.

The front-end of the symbol-wise noncoherent detector is a bank of M pairs of matched filters, with one pair matched to the in-phase and quadrature components of each tone. Because the detector works on a symbol-by-symbol basis, we may drop the dependence on the symbol index i in the following discussion. Furthermore, because the symbol-wise noncoherent detector allows the received phase to be any arbitrary value, the CPFSK induced phase ϕ can be absorbed into the fading phase θ .

After it is matched filtered and sampled at the symbol epoch, the received signal $y(t)$ can be written in a vector form as

$$\mathbf{y} = a e^{j\theta} \sqrt{\mathcal{E}_s} \mathbf{x} + \mathbf{n}, \quad (4)$$

where each element of \mathbf{y} , \mathbf{x} and \mathbf{n} can be represented as

$$y_k = a e^{j\theta} \sqrt{\mathcal{E}_s} x_k + n_k \quad (5)$$

$$x_k = \int_0^{T_s} x(t) s_k^*(t) dt \quad (6)$$

$$n_k = \int_0^{T_s} n(t) s_k^*(t) dt. \quad (7)$$

The noise vector \mathbf{n} is now Gaussian with a covariance matrix $\mathbf{R} = \mathbf{E}(\mathbf{n}\mathbf{n}^H)$ with $(k, i)^{th}$ element

$$\begin{aligned} r_{k,i} &= N_0 \int_0^{T_s} s_k^*(t) s_i(t) dt \\ &= N_0 \rho_{k,i}, \end{aligned} \quad (8)$$

where

$$\rho_{k,i} = N_0 \frac{\sin(\pi(i-k)h)}{\pi(i-k)h} e^{j\pi(i-k)h}. \quad (9)$$

Since \mathbf{n} is Gaussian, the vector \mathbf{y} given \mathbf{x} , $a\sqrt{\mathcal{E}_s}$, and θ is Gaussian distributed with mean \mathbf{x} and covariance \mathbf{R} . Thus, the conditional joint probability density function (pdf) is

$$\begin{aligned} & p(\mathbf{y}|\mathbf{x}, a\sqrt{\mathcal{E}_s}, \theta) \\ &= \frac{1}{\pi^M \det(\mathbf{R})} \exp^{-(\mathbf{y} - a e^{j\theta} \sqrt{\mathcal{E}_s} \mathbf{x})^H \mathbf{R}^{-1} (\mathbf{y} - a e^{j\theta} \sqrt{\mathcal{E}_s} \mathbf{x})} \end{aligned} \quad (10)$$

The exponent of the conditional pdf can be simplified as,

$$\begin{aligned} & -(\mathbf{y} - a e^{j\theta} \sqrt{\mathcal{E}_s} \mathbf{x})^H \mathbf{R}^{-1} (\mathbf{y} - a e^{j\theta} \sqrt{\mathcal{E}_s} \mathbf{x}) \\ &= -\mathbf{y}^H \mathbf{R}^{-1} \mathbf{y} - a^2 \mathcal{E}_s \mathbf{x}^H \mathbf{R}^{-1} \mathbf{x} \\ & \quad + 2 \operatorname{Re}(a e^{-j\theta} \sqrt{\mathcal{E}_s} \mathbf{x}^H \mathbf{R}^{-1} \mathbf{y}) \end{aligned} \quad (11)$$

$$\begin{aligned} &= -\mathbf{y}^H \mathbf{R}^{-1} \mathbf{y} - a^2 \mathcal{E}_s \mathbf{x}^H \mathbf{R}^{-1} \mathbf{x} \\ & \quad + 2a \sqrt{\mathcal{E}_s} |\mathbf{x}^H \mathbf{R}^{-1} \mathbf{y}| \cos(\psi - \theta), \end{aligned} \quad (12)$$

where ψ is the complex angle of $\mathbf{x}^H \mathbf{R}^{-1} \mathbf{y}$.

Define $\mathbf{K} \triangleq \frac{1}{N_0} \mathbf{R}$, which is the normalized version of \mathbf{R} , not dependant upon SNR. Note that when $x(t) = s_\nu(t)$ or equivalently the symbol index $q = \nu$, \mathbf{x} is the ν^{th} column of \mathbf{K} , i.e. $\mathbf{x} = \mathbf{k}_\nu$. Therefore, when $x(t) = s_\nu(t)$, the exponent becomes

$$-\frac{\mathbf{y}^H \mathbf{K}^{-1} \mathbf{y} + a^2 \mathcal{E}_s}{N_0} + 2 \frac{a \sqrt{\mathcal{E}_s}}{N_0} |y_\nu| \cos(\psi - \theta) \quad (13)$$

Integrating $p(\mathbf{y}|\mathbf{x}, a\sqrt{\mathcal{E}_s}, \theta)$ over the uniform random variable θ yields,

$$\begin{aligned} & p(\mathbf{y}|\mathbf{x} = \mathbf{k}_\nu, a\sqrt{\mathcal{E}_s}) \\ &= \frac{1}{2\pi} \int_0^{2\pi} p(\mathbf{y}|\mathbf{x} = \mathbf{k}_\nu, a\sqrt{\mathcal{E}_s}, \theta) d\theta \\ &= \frac{\exp^{-\frac{1}{N_0} (\mathbf{y}^H \mathbf{K}^{-1} \mathbf{y} + a^2 \mathcal{E}_s)}}{\pi^M N_0^M \det(\mathbf{K})} I_0 \left(2 \frac{a \sqrt{\mathcal{E}_s}}{N_0} |y_\nu| \right), \end{aligned} \quad (14)$$

where $I_0(\cdot)$ is the zeroth order modified Bessel function of the first kind. Notice that the factor that premultiplies the Bessel function is constant for all postulated symbols \mathbf{x} and therefore the conditional probability is proportional to just the Bessel function factor. Interestingly, this is the exact same metric as for the orthogonal case with known $a\sqrt{\mathcal{E}_s}/N_0$ [2], and therefore as will be made more clear in the next section, the normalized correlation matrix \mathbf{K} does not need to be directly used to form the decision statistic (though knowledge of the signal set is still needed for the receiver to compute the matched filter outputs).

III. CPFSK CONSTRAINED CAPACITY

The capacity of a channel with input symbol \mathbf{x} and output symbol \mathbf{y} is

$$C = \max_{p(\mathbf{x})} I(\mathbf{x}; \mathbf{y}), \quad (15)$$

where $p(\mathbf{x})$ is the joint pdf of \mathbf{x} , the average mutual information is

$$I(\mathbf{x}; \mathbf{y}) = E[i(\mathbf{x}; \mathbf{y})], \quad (16)$$

and the mutual information random variable is

$$\begin{aligned} i(\mathbf{x}; \mathbf{y}) &= \log \frac{p(\mathbf{x}, \mathbf{y})}{p(\mathbf{x})p(\mathbf{y})} \\ &= \log \frac{1}{p(\mathbf{x})} + \log p(\mathbf{x}|\mathbf{y}) \end{aligned} \quad (17)$$

Note that the expectation in (16) is with respect to the joint pdf $p(\mathbf{x}, \mathbf{y})$.

When there are no constraints on the modulation, (15) is maximized by letting \mathbf{x} be jointly Gaussian. However, when there are constraints on the modulation, we may not choose $p(\mathbf{x})$ in such a way. Rather, the distribution $p(\mathbf{x})$ is inextricably linked to the choice of modulation.

Because $\sum_{\mathbf{x}' \in \mathcal{S}} p(\mathbf{x}'|\mathbf{y}) = 1$, we may rewrite (17) as

$$\begin{aligned} i(\mathbf{x}; \mathbf{y}) &= \log \frac{1}{p(\mathbf{x})} + \log \frac{p(\mathbf{x}|\mathbf{y})}{\sum_{\mathbf{x}' \in \mathcal{S}} p(\mathbf{x}'|\mathbf{y})} \\ &= \log \frac{1}{p(\mathbf{x})} - \log \frac{\sum_{\mathbf{x}' \in \mathcal{S}} p(\mathbf{x}'|\mathbf{y})}{p(\mathbf{x}|\mathbf{y})}. \end{aligned} \quad (18)$$

From Bayes' law and assuming equiprobable \mathbf{x} , this becomes

$$i(\mathbf{x}; \mathbf{y}) = \log M - \log \frac{\sum_{\mathbf{x}' \in \mathcal{S}} p(\mathbf{y}|\mathbf{x}')}{p(\mathbf{y}|\mathbf{x})}. \quad (19)$$

For a given realization of the channel input \mathbf{x} and output \mathbf{y} , the mutual information random variable may be computed using (19) and the conditional pdf $p(\mathbf{y}|\mathbf{x})$. For noncoherent CPFSK, the conditional pdf is given by (14). Because of the ratio of conditional pdfs in (19), factors that are common to all \mathbf{x} will cancel, and thus

$$i(\mathbf{x}; \mathbf{y}) = \log M - \log \frac{\sum_{\mathbf{x}' \in \mathcal{S}} f(\mathbf{y}|\mathbf{x}')}{f(\mathbf{y}|\mathbf{x})}, \quad (20)$$

where $f(\mathbf{y}|\mathbf{x}) = \kappa p(\mathbf{y}|\mathbf{x})$ for any value of κ that is constant for all postulated symbols \mathbf{x} . For CPFSK, $f(\mathbf{y}|\mathbf{x})$ is found from (14) to be

$$f(\mathbf{y}|\mathbf{x} = \mathbf{k}_\nu) = I_0 \left(2 \frac{a\sqrt{\mathcal{E}_s}}{N_0} |y_\nu| \right), \quad (21)$$

where to maintain a compact notation, the conditioning on $a\sqrt{\mathcal{E}_s}$ is no longer indicated.

According to (16), the average mutual information $I(\mathbf{x}; \mathbf{y})$ is found by taking the expected value of the mutual information random variable $i(\mathbf{x}; \mathbf{y})$ with respect to the joint pdf $p(\mathbf{x}, \mathbf{y})$. For CPFSK, this expectation cannot be expressed in closed form and thus numerical techniques must be applied. While for low dimensionality (e.g. $M = 2$), the expectation can be evaluated using numerical integration, for higher dimensionality ($M \geq 4$), Monte Carlo integration is most effective. In the Monte Carlo integration, a very long sequence of symbols \mathbf{q} is drawn at random and then equation (5) is used to produce simulated matched filter outputs with randomly generated noise and fading. Given the values of the y_k 's and knowledge of the actual symbols, the mutual information random variable may be evaluated by substituting (21) into (20) and then the average is taken over the full set of input symbols. Because the channel is ergodic, the value of $I(\mathbf{x}; \mathbf{y})$ estimated through Monte Carlo integration converges to the true value as the length of the simulated sequence goes to infinity.

For a fixed modulation format, $p(\mathbf{x})$ is fixed and therefore the maximization in (15) goes away and the capacity is merely the average mutual information. However, if the modulation has parameters, the maximization in (15) is over the set

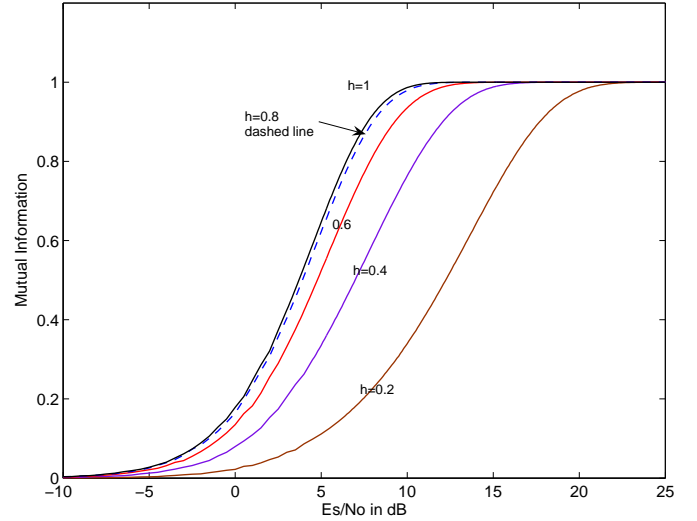


Fig. 1. Average mutual information versus \mathcal{E}_s/N_0 for noncoherent CPFSK in AWGN with $M = 2$ and several modulation indices $h = \{0.2, 0.4, 0.6, 0.8, 1\}$. For fixed \mathcal{E}_s/N_0 , the mutual information increases with h .

of parameters. In the case of CPFSK modulation, there are two parameters, M and h , and the capacity is found by optimizing the mutual information with respect to both of these parameters. An example is shown in Fig. 1 which plots the average mutual information as a function of \mathcal{E}_s/N_0 in AWGN with $M = 2$ and several values of h .

According to the Shannon coding theorem, it is possible to signal at an arbitrarily low error probability provided that the code rate $r \leq I(\mathbf{x}; \mathbf{y})$. When using a rate r code, the energy per information bit is $\mathcal{E}_b = \mathcal{E}_s/r$. By replacing $I(\mathbf{x}; \mathbf{y})$ with r and then plotting $\mathcal{E}_b/N_0 = (\mathcal{E}_s/N_0)/r$ as a function of r , the information-theoretic minimum \mathcal{E}_b/N_0 required for reliable signaling can be determined. An example is shown in Fig. 2, for AWGN, $M = 2$, and the same values of h that were shown in Fig. 1.

From Fig. 2 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 . This behavior is called the *noncoherent combining penalty* [1]. Unlike coherent systems, going to a lower r does not necessarily improve energy efficiency. Another trend to notice from Fig. 1 is that the minimum \mathcal{E}_b/N_0 decreases with increasing h , reaching its minimum value for orthogonal modulation ($h = 1$). Furthermore, the results shown in Figs. 1 and 2 are for the $M = 2$ case and would have to be repeated for all other M . As M increases, the minimum required \mathcal{E}_b/N_0 decreases [2]. With these behaviors in mind, it is quite apparent that the CPFSK design point that achieves capacity by 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. In the next section, a methodology is presented for determining capacity under bandwidth constraints. To manage complexity,

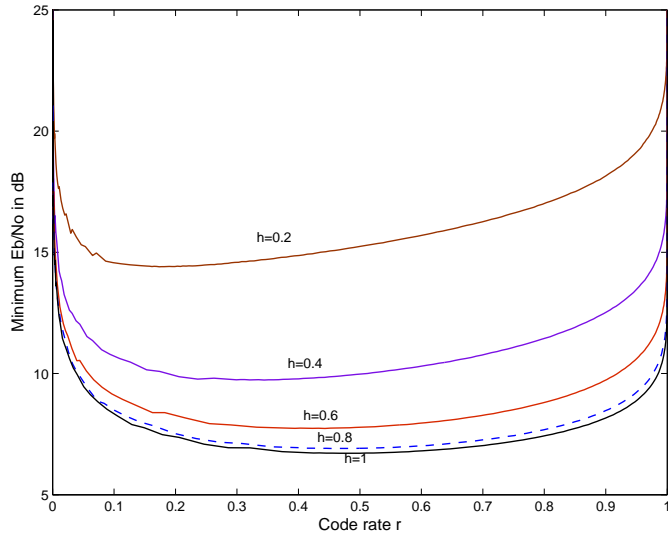


Fig. 2. Minimum \mathcal{E}_b/N_0 required for noncoherent CPFSK to achieve an arbitrarily low error rate versus code rate r in AWGN with $M = 2$ and several modulation indices $h = \{0.2, 0.4, 0.6, 0.8, 1\}$. For fixed r , the minimum \mathcal{E}_b/N_0 decreases with increasing h .

the capacity should be further conditioned on the choice of M .

IV. BANDWIDTH-CONSTRAINED CAPACITY

Up to this point, the discussion has focused on minimizing \mathcal{E}_b/N_0 without respect to bandwidth. However, the goal of this paper is to determine the fundamental tradeoff between \mathcal{E}_b/N_0 and spectral efficiency. To quantify this tradeoff, the bandwidth of the CPFSK signal must be computed. The power spectral density (PSD) $\Phi_s(f)$ of the CPFSK signal $s(t)$ is given in Section 4.4.2 of [7]. From the PSD, the 99% power bandwidth B_{99} of $s(t)$ is defined as

$$\int_{-B_{99}/2}^{B_{99}/2} \Phi_s(f) df = 0.99 \int_{-\infty}^{\infty} \Phi_s(f) df. \quad (22)$$

This bandwidth is a function of M , h , and the symbol rate $R_s = 1/T_s$. Given that $s(t)$ with parameters M and h is transmitted at a rate of R_s baud, we can define the *normalized bandwidth* to be $B(M, h) = B_{99}T_s$ Hz/baud. We can then define the *spectral efficiency* $\eta = r/B(M, h)$, which has units of bits-per-second-per-Hz (bps/Hz).

To determine the fundamental tradeoff between η and \mathcal{E}_b/N_0 , one must determine the minimum value of \mathcal{E}_b/N_0 for a particular desired spectral efficiency η . This procedure is identical to the procedure for the bandwidth unlimited case, except now the range of r that may be considered is restricted. More specifically, for each choice of η , h , and M , there will be a threshold r' on code rate

$$r' = \eta B(M, h) \quad (23)$$

such that $r \in [r', \log_2 M]$. Rates $r < r'$ cannot be considered because for the particular h and M , the spectral efficiency will be lower than η . Under tight bandwidth constraints, the

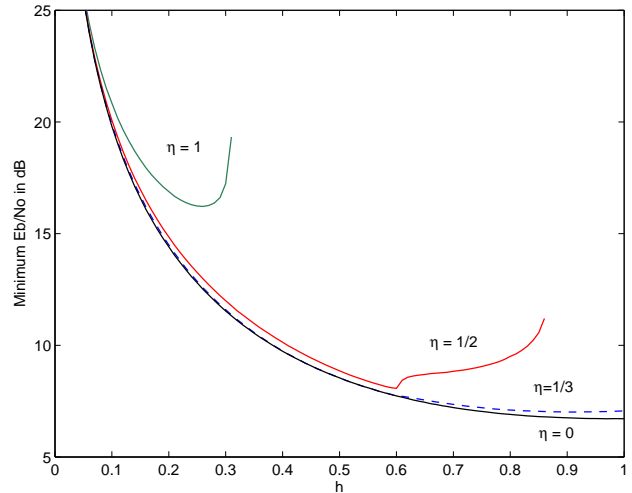


Fig. 3. Minimum \mathcal{E}_b/N_0 required for noncoherent CPFSK to achieve an arbitrarily low error rate versus modulation index h in AWGN with $M = 2$ for several spectral efficiencies $\eta = \{0, 1/3, 1/2, 1\}$. For fixed h , the minimum \mathcal{E}_b/N_0 increases with η .

optimal r is typically equal to its minimum value $\eta B(M, h)$, but in looser bandwidth constraints the optimal r might be higher due to the noncoherent combining penalty.

The determination of the minimum \mathcal{E}_b/N_0 for each choice of M , h , and η requires that the curve showing \mathcal{E}_b/N_0 versus r be generated, as it was in Fig. 2. Next, the minimum rate r' is determined. For example, when $M = 2$ and $\eta = 1/2$ bps/Hz, the minimum values of r are 0.39, 0.55, 0.64, and 0.96 for $h = 0.2, 0.4, 0.6$, and 0.8, respectively. Since $B(M = 2, h = 1) = 2.1309 > 1/\eta$, no code of rate $r \leq 1$ can be used at this η when $h = 1$ and thus orthogonal modulation cannot be considered. Next, the minimum \mathcal{E}_b/N_0 is found by inspecting the curve over the range of possible rates $r \in [\eta B(M, h), \log_2 M]$. For a given η and M , this procedure is repeated for each value of h over a range $(0, h')$, where $h' = \max_{h \leq 1} : B(M, h) \leq (\log_2 M)/\eta$ 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 $r \leq \log_2 M$. The minimum \mathcal{E}_b/N_0 for each possible h can then be plotted as a function of h . An example is shown in Fig. 3 for $M = 2$ in AWGN and several values of η (the $\eta = 0$ case corresponds to having no bandwidth constraint).

As can be seen in Fig. 3, for each value of η there is an optimal choice of h that minimizes \mathcal{E}_b/N_0 . For the unlimited bandwidth case ($\eta = 0$), the optimal $h = 1$, but as η increases, the optimal value of h decreases. The combination of η and the \mathcal{E}_b/N_0 minimized over h is the constrained channel capacity for that value of M , channel (AWGN), and noncoherent detection.

A plot of minimum \mathcal{E}_b/N_0 versus h for all $M \leq 64$ and $\eta = \{0, 1/2\}$ is shown in Fig. 4 for the AWGN channel and in Fig. 5 for the Rayleigh fading channel. For each of the six values of M and two channel types, capacity curves were

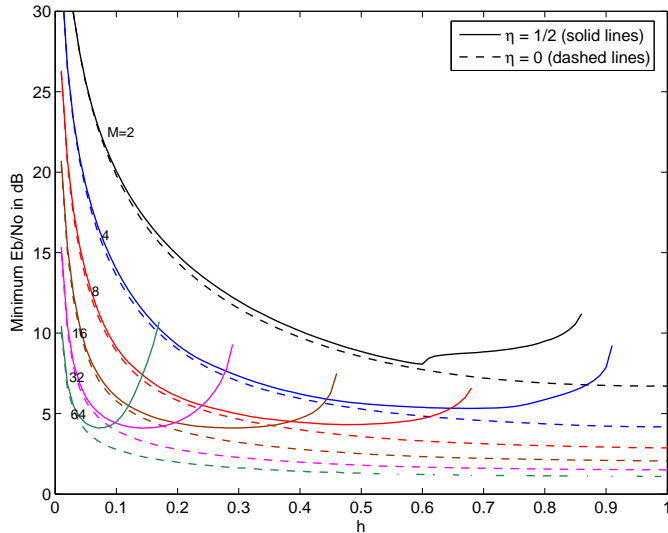


Fig. 4. Minimum \mathcal{E}_b/N_0 required for noncoherent CPFSK to achieve an arbitrarily low error rate versus modulation index h in AWGN for several modulation orders $M = \{2, 4, 8, 16, 32, 64\}$ and spectral efficiencies $\eta = \{0, 1/2\}$.

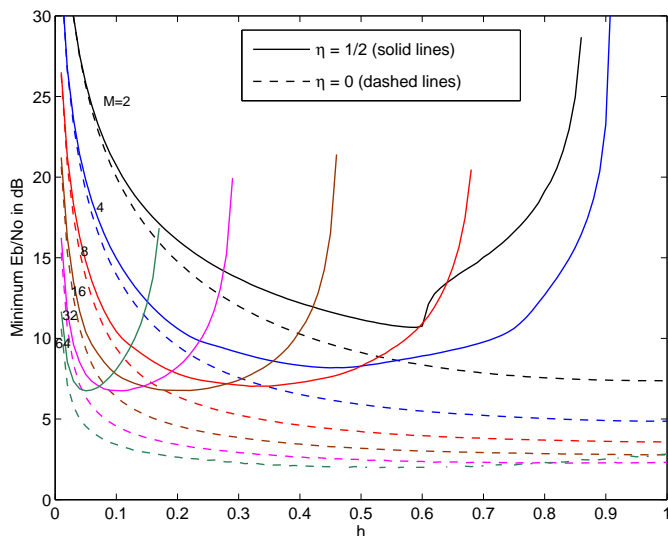


Fig. 5. Minimum \mathcal{E}_b/N_0 required for noncoherent CPFSK to achieve an arbitrarily low error rate versus modulation index h in Rayleigh fading for several modulation orders $M = \{2, 4, 8, 16, 32, 64\}$ and spectral efficiencies $\eta = \{0, 1/2\}$.

generated for values of h ranging from $h = 0.01$ to $h = 1$ in increments of 0.01. Thus a total of 1,200 capacity curves were generated and each curve was created using at least 2 million simulated symbols per SNR point in the range of interest. Altogether, over 1 trillion symbols were simulated, and it is estimated that this task would have taken about one year to complete on a single PC computer. To speed the run time, simulations were executed on a virtual private grid computer powered by the idle capacity of 30 workstations located in

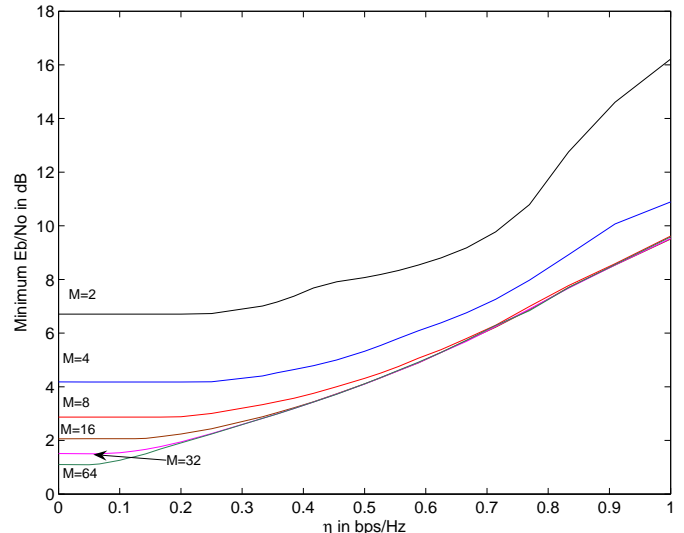


Fig. 6. Minimum \mathcal{E}_b/N_0 required for noncoherent CPFSK to achieve an arbitrarily low error rate versus spectral efficiency η in AWGN for several modulation orders $M = \{2, 4, 8, 16, 32, 64\}$. For fixed η the minimum \mathcal{E}_b/N_0 decreases with increasing M . The values at $\eta = 0$ correspond to the orthogonal FSK capacity.

the teaching laboratories at the first author's institution¹. The entire simulation scenario took just two weeks to complete on the grid computer.

From these curves, it can be seen that the minimum h decreases with increasing M . Interestingly, the minimum \mathcal{E}_b/N_0 decreases with increasing M even when the bandwidth is constrained. By finding the minimum value of \mathcal{E}_b/N_0 with respect to h for each M over a wide range of η , one can finally determine the capacity of CPFSK. Capacity can now be plotted in terms of spectral efficiency η versus the corresponding minimum \mathcal{E}_b/N_0 , as shown for several M and $\eta \leq 1$ in Fig. 6 for AWGN and Fig. 7 for Rayleigh fading. Note that the minimum \mathcal{E}_b/N_0 in dB increases roughly linearly with η . The minimum \mathcal{E}_b/N_0 decreases with increasing M . The minimum \mathcal{E}_b/N_0 at $\eta = 0$ is achieved with $h = 1$ for each M , and therefore these values are identical to the ones for orthogonal FSK modulation given in [2]. While there is a benefit to increasing M at very low η , these benefits begin to disappear as η is increased. For both the AWGN and Rayleigh fading channels, there is no benefit to using $M > 16$ for spectral efficiencies $\eta > 0.3$ since the curves for $M = \{16, 32, 64\}$ merge at these higher spectral efficiencies. Furthermore, in AWGN the curves for $M = \{8, 16, 32, 64\}$ begin to merge as $\eta > 0.5$, indicating that there is no benefit to using $M > 8$ in AWGN when $\eta > 0.5$.

V. CONCLUSION

This paper outlines a methodology for finding the coded modulation (CM) capacity of CPFSK with noncoherent detection. By optimizing over both the modulation index h and

¹Job scheduling was performed online via the *Global Grid Exchange* (g2ex.com) which runs the *Frontier Grid Platform* developed by *Parabon Computation* (parabon.com).

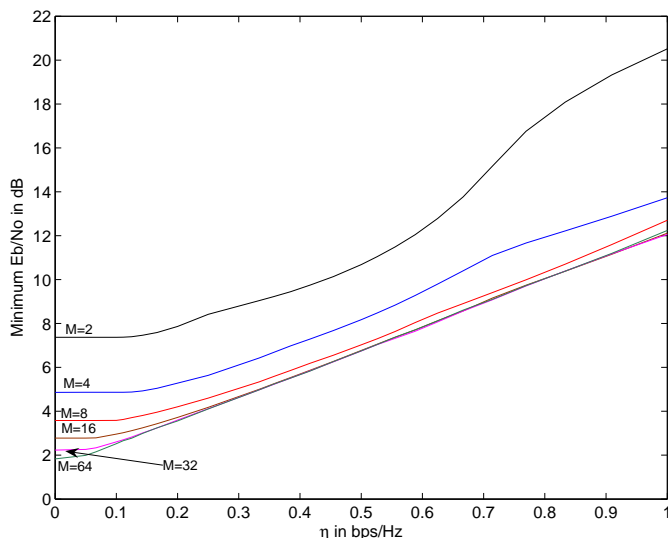


Fig. 7. Minimum \mathcal{E}_b/N_0 required for noncoherent CPFSK to achieve an arbitrarily low error rate versus spectral efficiency η in Rayleigh fading for several modulation orders $M = \{2, 4, 8, 16, 32, 64\}$. For fixed η the minimum \mathcal{E}_b/N_0 decreases with increasing M . The values at $\eta = 0$ correspond to the orthogonal FSK capacity.

code rate r , the capacity analysis is able to determine the minimum value of \mathcal{E}_b/N_0 as a function of spectral efficiency η and number of tones M . The values of the optimizing h and r and the resulting minimum \mathcal{E}_b/N_0 are channel dependent, and thus the values for AWGN will be different than the values for Rayleigh fading. For high spectral efficiency, it is generally better to use smaller values of h and allow the channel code to overcome the resulting increased inter-tone interference. While performance improves with increasing M , there is no significant benefit from using $M > 8$ when the spectral efficiency is sufficiently high.

There are several limitation of the system considered in this paper. First, since it is concerned with CM capacity, the code and modulation alphabets must be matched. Because many “off-the-shelf” codes are binary rather than M -ary, it may be more convenient to use bit-interleaved coded modulation (BICM) [3]. Extending the capacity analysis to encompass BICM is fairly straightforward and has already been considered for orthogonal FSK in [2]. The key issue when extending BICM to nonorthogonal FSK is that, unlike in the orthogonal case, performance will be sensitive to how bits are mapped to symbols. Thus the determination of BICM capacity requires an additional optimization over all symbol labellings.

Another limitation is that the demodulator is noncoherent in the sense that it is agnostic to the received phase offset, yet it requires knowledge of the channel state information (CSI) in the form of the fading amplitudes a and the average SNR \mathcal{E}_s/N_0 . Alternatively, the receiver could be designed to work without knowledge of the a 's by marginalizing the joint pdf (14) with respect to the pdf of a . The formulation of such a receiver and its capacity for orthogonal signalling is discussed in [1] and [2], though the value of \mathcal{E}_s/N_0 must still be known.

Another option is to estimate the value of $a\sqrt{\mathcal{E}_s}/N_0$ as we propose in [8]. Besides the noncoherent detection considered in this paper, the signal could be detected using coherent or differential techniques, and the corresponding capacities are open problems.

The CPFSK modulation in this paper is a special case of the more general class of continuous phase modulation (CPM) [9]. Because the signals defined in (1) are zero outside the symbol epoch $[0, T_s)$, the CPFSK modulation in this paper is full-response. More generally, partial response signaling may be used, as may more elaborate pulse shapes such as in Gaussian frequency shift keying (GFSK). When the signaling is partial response, there will be modulation induced intersymbol interference (ISI). If left unmitigated, the ISI is likely to severely degrade capacity and offset any advantages due to the more compact spectrum. However, more elaborate reception techniques can compensate for the ISI, albeit at the cost of increased complexity (e.g. a larger trellis).

While the CPFSK modulation described in this paper can be used for a variety of applications, it is especially attractive for frequency hopping spread spectrum (FHSS) systems [10]. Because FHSS systems are characterized by short dwell times, there is typically insufficient time to acquire the carrier and perform coherent detection, and thus its ability to be noncoherently detected is a key benefit of using CPFSK. On the other hand, FHSS systems should use modulation that is spectrally efficient so that more hopping frequencies can be supported by a fixed system bandwidth. Thus the combination of high spectral efficiency and ability to be noncoherently detected make nonorthogonal CPFSK an attractive choice for FHSS systems [11].

REFERENCES

- [1] W. E. Stark, “Capacity and cutoff rate of noncoherent FSK with nonselective Rician fading,” *IEEE Trans. Commun.*, vol. 33, pp. 1153–1159, Nov. 1985.
- [2] M. C. Valenti and S. Cheng, “Iterative demodulation and decoding of turbo coded M -ary noncoherent orthogonal modulation,” *IEEE J. Select. Areas Commun.*, vol. 23, pp. 1738–1747, Sept. 2005.
- [3] G. Caire, G. Taricco, and E. Biglieri, “Bit-interleaved coded modulation,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.
- [4] X. Li and J. A. Ritcey, “Bit-interleaved coded modulation with iterative decoding,” *IEEE Commun. Letters*, vol. 1, pp. 169–171, Nov. 1997.
- [5] A. Guillén i Fàbregas and A. Grant, “Capacity approaching codes for non-coherent orthogonal modulation,” *IEEE Trans. Wireless Comm.*, accepted 2006.
- [6] S. ten Brink, “Convergence of iterative decoding,” *Electronics Letters*, vol. 35, pp. 806–808, May 13, 1999.
- [7] J. Proakis, *Digital Communications*, 4th ed. New York, NY: McGraw-Hill, Inc., 2001.
- [8] S. Cheng, M. C. Valenti, and D. Torrieri, “Robust iterative noncoherent reception of coded FSK over block fading channels,” *IEEE Trans. Wireless Comm.*, 2006, accepted.
- [9] J. B. Anderson, T. Aulin, and C. E. Sundberg, *Digital Phase Modulation (Applications of Communications Theory)*. Springer, 1986.
- [10] D. Torrieri, *Principles of Spread-Spectrum Communication Systems*. Springer, 2004.
- [11] D. Torrieri, S. Cheng, and M. C. Valenti, “Robust frequency-hopping system for channels with interference and frequency-selective fading,” in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Glasgow, Scotland, 2007, to appear.