The Impact of Hard-Decision Detection on the Energy Efficiency of Phase and Frequency Modulation

Mustafa Cenk Gursoy

Abstract— The central design challenge in next generation wireless systems is to have these systems operate at high bandwidths and provide high data rates while being cognizant of the energy consumption levels especially in mobile applications. Since communicating at very high data rates prohibits obtaining high bit resolutions from the analog-to-digital (A/D) converters, analysis of the energy efficiency under the assumption of hard-decision detection is called for to accurately predict the performance levels. In this paper, transmission over the additive white Gaussian noise (AWGN) channel, and coherent and noncoherent fading channels is considered, and the impact of hard-decision detection on the energy efficiency of phase and frequency modulations is investigated. Energy efficiency is analyzed by studying the capacity of these modulation schemes and the energy required to send one bit of information reliably in the low signal-to-noise ratio (SNR) regime. The capacity of hard-decision-detected phase and frequency modulations is characterized at low SNR levels through closed-form expressions for the first and second derivatives of the capacity at zero SNR. Subsequently, bit energy requirements in the low-SNR regime are identified. The increases in the bit energy incurred by hard-decision detection and channel fading are quantified. Moreover, practical design guidelines for the selection of the constellation size are drawn from the analysis of the spectral efficiency–bit energy tradeoff.

Index Terms: **Bit energy, spectral efficiency, AWGN channel, fading channels, phase-shift keying, frequency-shift keying, on-off keying, hard-decision detection.**

I. INTRODUCTION

Energy efficiency is of paramount importance in many communication systems and particularly in mobile wireless systems due to the scarcity of energy resources. Energy efficiency can be measured by the energy required to send one information bit reliably. It is well-known that for Gaussian channels subject to average power constraints, the minimum received bit energy normalized by the noise spectral level is $\frac{E_b}{N_0 \min} = -1.59$ dB regardless of the availability of channel side information (CSI) at the receiver (see e.g., [1] – [5], and [8]). Golay [1] showed that this minimum bit energy can be achieved in the additive white Gaussian noise (AWGN) channel by pulse position modulation (PPM) with

vanishing duty cycle when the receiver employs threshold detection. Indeed, Turin [2] proved that any orthogonal M-ary modulation scheme with envelope detection at the receiver achieves the normalized bit energy of -1.59 dB in the AWGN channel as $M \to \infty$. It is further shown in [3] and [4] that M-ary orthogonal frequency-shift keying (FSK) achieves this minimum bit energy asymptotically as $M \to \infty$ also in noncoherent fading channels where neither the receiver nor the transmitter knows the fading coefficients. These studies demonstrate the asymptotical high energy efficiency of orthogonal signaling even when the receiver performs hard-decision detection. As also well-known by now in the digital communications literature [23], these results are shown by proving that the error probabilities of orthogonal signaling can be made arbitrarily small as $M \to \infty$ as long as the normalized bit energy (or equivalently SNR per bit) is greater than -1.59 dB. As indicated by the unbounded growth of M, the minimum bit energy is in general achieved at infinite bandwidth or equivalently as the spectral efficiency (rate in bits per second divided by bandwidth in Hertz) goes to zero.

Indeed for average power limited channels, the bit energy required for reliable communication decreases monotonically with increasing bandwidth [6], [8]. This is the fundamental bandwidth-power tradeoff. Recently, Verdú [8] has offered a more subtle analysis of the tradeoff of bit energy versus spectral efficiency. In this work, the wideband slope, which is the slope of the spectral efficiency curve at zero spectral efficiency, has emerged as a new analysis tool to measure energy and bandwidth efficiency in the low-power regime. It is shown that if the receiver has perfect knowledge of the fading coefficients, quaternary phase-shift keying (QPSK) is an optimally efficient modulation scheme achieving both the minimum bit energy of -1.59 dB and the optimal wideband slope in the low SNR regime. This indicates that besides orthogonal signaling, phase modulation is also well-suited for energy efficient operation. However, it should be noted that asymptotic efficiency of QPSK holds under the assumption that the receiver performs soft detection. Verdú [8] has also provided expressions for the minimum bit energy and wideband slope of the quantized QPSK. We note that phase modulation is a widely used technique for information transmission, and the performance of coded phase modulation has been of interest in the research community since the 1960s. One of the early works was conducted in [10] where the capacity and error exponents of a continuous-phase modulated system, in which the transmitted phase can assume any value in $[-\pi, \pi)$, is studied. More recent

studies include [4], [9], and [11]–[14].

As discussed above, high energy efficiency requires operation in the wideband regime in which the spectral efficiencies are low. This is is achieved by either decreasing the data rates or increasing the bandwidth. If the system has large bandwidth, then the data rates are high. For instance, if the total signal power is $P = 1$ mW and the bandwidth is $B = 1$ GHz, then the capacity of the AWGN channel is $C = B \log_2 \left(1 + \frac{P}{N_0 B}\right) \approx 27.9$ Gbits/s[1]. If the bandwidth is increased to $B = 10$ GHz, the capacity becomes 245.7 Gbits/s. Similarly, high rates are also achieved in fading channels when the available bandwidth is large. For instance, in current practical applications, wideband CDMA and ultrawideband systems offer high data rate services by using large bandwidths [27]. Additionally, operating at high bandwidths and providing high data rates while conserving the energy in mobile applications are the key features of next generation wireless systems which have the goal of offering mobile multimedia access. For instance, one of the features of fourth generation (4G) systems will be the ability to support multimedia services at low transmission cost [27, Chap. 23, available online]. On the other hand, at these very high transmission rates, obtaining high bit resolutions from A/D converters may either be not possible or prohibitively expensive. Therefore, in such cases, the performance of soft detection will be a loose upper bound on the actual system performance, and analysis under the assumption of hard-decision detection will provide more faithful results. Moreover, even if the data rates are not high, hard-decision detection of the received signals might be preferred when reduction in the computational burden is required [23]. Such a requirement, for instance, may be enforced in sensor networks that consist of low-cost, low-power, small-sized sensor nodes [15]. Therefore, it is timely and practically relevant to study the energy efficiency of phase and frequency modulations in the wideband regime when the receiver performs hard-decision detection.

The contributions of this paper are the following:

1) We obtain closed-form expressions for the first and second derivatives at zero SNR of the hard-decision-detected PSK capacity for arbitrary modulation size M.
2) We find the bit energy required at zero spectral efficiency and wideband slope when PSK is employed at the transmitter. The analysis is initially performed for non-coherent fading channels, and subsequently specialized to the AWGN and coherent fading channels. We quantify the increase in the bit energy requirements due to hard-decision detection and channel fading.
3) We study the energy efficiency of hard-decision-detected on-off frequency-shift keying (OOFSK) modulation which is a general orthogonal signaling scheme that combines orthogonal FSK and on-off keying (OOK) and introduces peakedness in both time and frequency. We show that the bit energy requirements grow without bound with decreasing SNR if the peakedness in both time and frequency is limited. We identify the impact upon the energy efficiency of the number of orthogonal

frequencies, M, and the duty cycle of OOK. We prove a sufficient condition on how fast the duty cycle has to vanish with decreasing SNR in order to approach the fundamental bit energy limit of -1.59dB.

The organization of the rest of the paper is as follows. In Section II, we describe the channel model. The energy efficiency of phase modulation is investigated in Section III. M-ary OOFSK modulation and its special case M-ary FSK modulation are considered in Section IV. Section V includes our conclusions.

II. CHANNEL MODEL

We consider the following channel model

$$\mathbf{r}_k = h_k \mathbf{s}_{x_k} + \mathbf{n}_k \quad k = 1, 2, 3 \ldots \quad (1)$$

where x_k is the discrete input, $\mathbf{s}_{x_k}$ is the transmitted signal when the input is x_k, and $\mathbf{r}_k$ is the received signal during the the kth symbol duration. h_k is the channel gain. h_k is a fixed constant in unfaded AWGN channels, while in flat fading channels, h_k denotes the fading coefficient. $\{\mathbf{n}_k\}$ is a sequence of independent and identically distributed (i.i.d.) zero-mean circularly symmetric Gaussian random vectors with covariance matrix $E\{\mathbf{n}\mathbf{n}^\dagger\} = N_0 \mathbf{I}$ where $\mathbf{I}$ denotes the identity matrix. We assume that the system has an average energy constraint of $E\{\|\mathbf{s}_{x_k}\|^2\} \leq \mathcal{E} \quad \forall k$.

At the transmitter, if M-ary PSK modulation is employed for transmission, the discrete input, x_k, takes values from $\{1, \ldots, M\}$. If $x_k = m$, then the transmitted signal in the kth symbol duration is

$$s_{x_k} = s_m = \sqrt{\mathcal{E}} e^{j\theta_m} \quad (2)$$

where $\theta_m = \frac{2\pi(m-1)}{M}$ for $m = 1, \ldots, M$, is one of the M phases available in the constellation. In the case of PSK modulation, since s_{x_k} is a one-dimensional complex point, we opted to not use the boldface representation. Accordingly, the output r_k and the noise n_k are one complex-dimensional points. The receiver is assumed to perform hard-decision detection. Therefore, each received signal r_k is mapped to one of the points in the constellation set $\{\sqrt{\mathcal{E}} e^{j2\pi(m-1)/M}, m = 1, \ldots, M\}$ before going through the decoder. We assume that maximum likelihood decision rule is used at the detector.

In [21], we have introduced the on-off frequency-shift keying (OOFSK) modulation by overlaying frequency-shift keying (FSK) on on-off keying (OOK). In M-ary OOFSK modulation, the transmitter either sends no signal with probability $1 - \nu$ or sends one of M orthogonal FSK signals each with probability ν/M. Hence, $\nu \in (0, 1]$ can be seen as the duty cycle of the transmission. In this case, the discrete input takes values from $x_k \in \{0, 1, 2, \ldots, M\}$. If $x_k = 0$, then there is no transmission and the geometric representation of the transmitted signal is the M-complex dimensional vector $\mathbf{s}_0 = (0, 0, \ldots, 0)$. On the other hand, if $x_k = m \neq 0$, an FSK signal is sent and the geometric representation is given by

$$\mathbf{s}_{x_k} = \mathbf{s}_m = (s_{m,1}, s_{m,2}, \ldots, s_{m,M}) \quad m = 1, 2, \ldots, M, \quad (3)$$

where $s_{m,m} = \sqrt{\mathcal{E}/\nu} e^{j\theta_m}$ and $s_{m,i} = 0$ for $i \neq m$. The phases θ_m can be arbitrary. Note that in M-ary OOFSK

---

[1]We have assumed that $N_0 = 4 \times 10^{-21}$ W/Hz [24].

modulation, we have $M + 1$ possible input signals including the no signal transmission. Therefore, no signal transmission being a part of the modulation also conveys a message to the receiver. While FSK signals have energy $\mathcal{E}/\nu$, the average energy of OOFSK modulation is $\mathcal{E}$. Hence, the peak-to-average power ratio of signaling is $1/\nu$. In the OOFSK transmission and reception model, the received signal $\mathbf{r}_k$ and noise $\mathbf{n}_k$ are also M-dimensional. It is assumed that the receiver performs energy detection on the received vector $\mathbf{r}_k$. Finally, note that OOFSK is a general orthogonal signaling format and specializes to regular orthogonal FSK if $\nu = 1$, and to OOK if $M = 1$ and $\nu \neq 1$.

We remark that in both PSK and OOFSK cases, the channel, after hard-decision detection, can be regarded as a discrete channel with finitely many inputs and outputs. Henceforth, capacity and achievable rate expressions throughout the paper will be obtained considering these discrete channels.

III. Energy Efficiency of Phase Modulation

A. Noncoherent Rician Fading Channels

In this section, we study the performance of phase modulated signals when they are hard-decision detected. We initially consider transmission of PSK signals over noncoherent Rician channels in which neither receiver nor transmitter knows the fading coefficients. Results for this channel are subsequently specialized to obtain the performance results of PSK in unfaded AWGN channels and coherent fading channels. Hence, we first assume that the fading coefficients $\{h_k\}$, whose realizations are unknown at the transmitter and receiver due to the noncoherence assumption, are i.i.d. proper complex Gaussian random variables with mean $E\{h_k\} = d \neq 0$ and variance $E\{|h_k - d|^2\} = \gamma^2$. We further assume that the channel statistics, and hence d and γ^2, are known both at the transmitter and receiver. Note that $d \neq 0$ is required because phase cannot be used to transmit information in a noncoherent Rayleigh fading channel where $d = 0$.

In the noncoherent Rician channel model, the conditional probability density function (pdf) of the channel output given the input is a conditionally complex Gaussian pdf and is given by[2]

$$f_{r|s_m}(r|s_m) = \frac{1}{\pi(\gamma^2|s_m|^2 + N_0)} e^{-\frac{|r - ds_m|^2}{\gamma^2|s_m|^2 + N_0}}. \quad (4)$$

Recall that $\{s_m = \sqrt{\mathcal{E}}e^{j\theta_m}\}$ are the PSK signals and hence $|s_m| = \sqrt{\mathcal{E}}$ for all $m = 1, \ldots, M$. Due to this constant magnitude property, it can be easily shown that the maximum likelihood detector selects s_k as the transmitted signal if[3]

$$\text{Re}(rs_k^*) > \text{Re}(rs_i^*) \quad \forall i \neq k \quad (5)$$

where s_k^* is the complex conjugate of s_k, and $\text{Re}()$ denotes the operation that selects the real part. We denote the signal at the output of the detector by y and assume that $y \in \{1, \ldots, M\}$. Note that $y = l$ for $l = 1, \ldots, M$ means that the detected

signal is $\sqrt{\mathcal{E}}e^{j2\pi(l-1)/M}$. Under the decision rule (5), the decision region for $y = l$ is the two-dimensional region

$$D_l = \left\{ r = |r|e^{j\theta} : \frac{(2l-3)\pi}{M} \leq \theta < \frac{(2l-1)\pi}{M} \right\} \quad (6)$$

for $l = 1, 2, \ldots, M$. With hard-decision detection at the receiver, the resulting channel is a symmetric, discrete, memoryless channel with input $x \in \{1, \ldots, M\}$ and output $y \in \{1, \ldots, M\}$. The transition probabilities are given by

$$P_{l,m} = P(y = l | x = m) \quad (7)$$

$$= P\left(\frac{(2l-3)\pi}{M} \leq \theta < \frac{(2l-1)\pi}{M} | x = m \right) \quad (8)$$

$$= \int_{\frac{(2l-3)\pi}{M}}^{\frac{(2l-1)\pi}{M}} f_{\theta|s_m}(\theta|s_m)\, d\theta \quad (9)$$

where $f_{\theta|s_m}(\theta|s_m)$ is the conditional probability density function of the phase of the received signal given that the input is $x = m$, and hence the transmitted signal is s_m. It is well-known that the capacity of this symmetric channel is achieved by equiprobable inputs and the resulting capacity expression [25] is[4]

$$C_{M,nc}(\text{SNR}) = \log M - H(y|x = 1) \quad (10)$$

$$= \log M + \sum_{l=1}^{M} P_{l,1} \log P_{l,1} \quad (11)$$

where $\text{SNR} = \frac{\mathcal{E}}{N_0}$, $H(\cdot)$ is the entropy function, and $P_{l,1} = P(y = l | x = 1)$. In order to evaluate the capacity of general M-ary PSK transmission with a hard-decision detector, the transition probabilities

$$P_{l,1} = P(y = l | x = 1) = \int_{\frac{(2l-3)\pi}{M}}^{\frac{(2l-1)\pi}{M}} f_{\theta|s_1}(\theta|s_1)\, d\theta \quad (12)$$

should be computed. Starting from (4) and noting that the conditional joint magnitude and phase distribution is given by

$$f_{|r|,\theta|s_1}(|r|,\theta|s_1) = \frac{|r|}{\pi(\gamma^2|s_1|^2 + N_0)} e^{-\frac{|r|^2 + |d|^2|s_1|^2 - 2|r||d||s_1|\cos\theta}{\gamma^2|s_1|^2 + N_0}} \quad (13)$$

where, without loss of generality, we have assumed that $d = |d|$, we can easily find that for $\theta \in [0, 2\pi)$, $f_{\theta|s_1}(\theta|s_1)$ is given by (14) on the next page where $Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$ is the Gaussian Q-function[5]. Since $f_{\theta|s_1}$ is rather complicated, closed-form capacity expressions in terms of Q-functions are available only for the special cases of $M = 2$ and 4:

$$C_{2,nc}(\text{SNR}) = \log 2 - h\left(Q\left(\sqrt{\frac{2|d|^2\text{SNR}}{\gamma^2\text{SNR} + 1}} \right) \right), \text{ and } \quad (15)$$

$$C_{4,nc}(\text{SNR}) = 2C_{2,nc}\left(\frac{\text{SNR}}{2} \right) \quad (16)$$

where $h(x) = -x \log x - (1-x) \log(1-x)$ is the binary

---

[2]Since the channel is memoryless, we henceforth, without loss of generality, drop the time index k in the equations for the sake of simplification.

[3]The decision rule is obtained when we assume, without loss of generality, that $d = |d|$.

[4]Throughout the paper, log is used to denote the logarithm to the base e, i.e., the natural logarithm. Additionally, the subscript "nc" in $C_{M,nc}$ signifies the noncoherent channel.

[5]See also [9] and references therein for a similar formula of the phase probability density function derived for the AWGN channel.

$$f_{\theta|s_1}(\theta|s_1) = \frac{1}{2\pi}\, e^{-\frac{|d|^2\text{SNR}}{\gamma^2\text{SNR}+1}} + \sqrt{\frac{|d|^2\text{SNR}}{\pi(\gamma^2\text{SNR}+1)}}\cos\theta\, e^{-\frac{|d|^2\text{SNR}}{\gamma^2\text{SNR}+1}\sin^2\theta}\left(1 - Q\left(\sqrt{2\frac{|d|^2\text{SNR}}{\gamma^2\text{SNR}+1}\cos^2\theta}\right)\right) \qquad (14)$$

entropy function. For the other cases, the channel capacity can only be found through numerical integration and computation.

On the other hand, the behavior of the capacity in the low-SNR regime can be accurately predicted through the second-order Taylor series expansion of the capacity[6]:

$$C_{M,nc}(\text{SNR}) = \dot{C}_{M,nc}(0)\text{SNR} + \ddot{C}_{M,nc}(0)\frac{\text{SNR}^2}{2} + o(\text{SNR}^2)$$

where $\dot{C}_{M,nc}(0)$ and $\ddot{C}_{M,nc}(0)$ denote the first and second derivatives, respectively, of the channel capacity (in nats/symbol) with respect to SNR at SNR $= 0$. In the following result, we provide closed-form expressions for these derivatives. Note that the wideband regime in which SNR per unit bandwidth is small can equivalently be regarded as the low-SNR regime.

Theorem 1: The first and second derivatives of $C_{M,nc}(\text{SNR})$ in nats per symbol at SNR $= 0$ are given by

$$\dot{C}_{M,nc}(0) = \begin{cases} \frac{2|d|^2}{\pi} & M = 2 \\ \frac{M^2|d|^2}{4\pi}\sin^2\frac{\pi}{M} & M \ge 3 \end{cases}, \quad \text{and}$$

$$\ddot{C}_{M,nc}(0) = \begin{cases} \frac{8}{3\pi}\left(\frac{1}{\pi}-1\right)|d|^4 - \frac{4|d|^2\gamma^2}{\pi} & M = 2 \\ \infty & M = 3 \\ \frac{4}{3\pi}\left(\frac{1}{\pi}-1\right)|d|^4 - \frac{4|d|^2\gamma^2}{\pi} & M = 4 \\ \psi(M)|d|^4 - \frac{|d|^2\gamma^2}{2\pi}M^2\sin^2\frac{\pi}{M} & M \ge 5 \end{cases} \quad (17)$$

respectively, where

$$\psi(M) = \frac{M^2}{16\pi^2}\Bigg((2-\pi)\sin^2\frac{2\pi}{M} + (M^2-4\pi)\sin^4\frac{\pi}{M}$$
$$- 2M\sin^2\frac{\pi}{M}\sin\frac{2\pi}{M}\Bigg). \qquad (18)$$

Proof: See Appendix A.

The derivative expressions in (17) are in closed-form and can be computed easily. Therefore, the low-SNR approximation of the capacity of M-ary PSK can be readily obtained from

$$C_{M,nc}(\text{SNR}) \approx \dot{C}_{M,nc}(0)\text{SNR} + \ddot{C}_{M,nc}(0)\frac{\text{SNR}^2}{2}.$$

The following corollary provides the asymptotic behavior as $M \to \infty$. In this asymptotic regime, the transmitted signal is the continuous phase which can take any value in $[-\pi, \pi)$.

Corollary 1: In the limit as $M \to \infty$, the first and second derivatives of the capacity at zero SNR converge to

$$\lim_{M\to\infty}\dot{C}_{M,nc}(0) = \frac{\pi|d|^2}{4} \quad \text{and} \qquad (19)$$

$$\lim_{M\to\infty}\ddot{C}_{M,nc}(0) = \frac{(\pi^2-8\pi+8)|d|^4}{16} - \frac{|d|^2\gamma^2\pi}{2}. \qquad (20)$$

In the low-power regime, the tradeoff between bit energy and spectral efficiency is a key measure of performance [8]. The normalized energy per bit can be obtained from $\frac{E_b}{N_0} =$

$\frac{\text{SNR}\log 2}{C(\text{SNR})}$ where $C(\text{SNR})$ is the channel capacity in nats/symbol. The maximum achievable spectral efficiency in bits/s/Hz is given by $\mathsf{C}\left(\frac{E_b}{N_0}\right) = C(\text{SNR})\log_2 e$ bits/s/Hz if we, without loss of generality, assume that one symbol occupies a 1s $\times$ 1Hz time-frequency slot. Two important notions regarding the spectral-efficiency/bit-energy tradeoff in the low power regime are the bit-energy required at zero spectral efficiency and wideband slope, given by

$$\frac{E_b}{N_0}\bigg|_{\mathsf{C}=0} = \frac{\log 2}{\dot{C}(0)}, \quad \text{and } S_0 = \frac{2(\dot{C}(0))^2}{-\ddot{C}(0)}, \qquad (21)$$

respectively. The wideband slope, S_0, provides the slope of the spectral efficiency curve $\mathsf{C}(E_b/N_0)$ at zero spectral efficiency [8]. Therefore, $\frac{E_b}{N_0}\big|_{\mathsf{C}=0}$ and S_0 constitute a linear approximation of the spectral efficiency curve in the low-SNR regime, i.e.,

$$\mathsf{C}\left(\frac{E_b}{N_0}\right) = \frac{S_0}{10\log_{10}2}\left(\frac{E_b}{N_0}\bigg|_{dB} - \frac{E_b}{N_0}\bigg|_{\mathsf{C}=0,dB}\right) + \epsilon \quad (22)$$

where $\frac{E_b}{N_0}\big|_{dB} = 10\log_{10}\frac{E_b}{N_0}$ and $\epsilon = o\left(\frac{E_b}{N_0} - \frac{E_b}{N_0}\big|_{\mathsf{C}=0}\right)$, and characterize the spectral-efficiency/bit-energy tradeoff at low spectral efficiencies. Hence, these quantities enable us to analyze the energy efficiency and investigate the interactions between spectral and energy efficiencies in the low-SNR regime. Depending only on $\dot{C}(0)$ and $\ddot{C}(0)$, the bit energy at zero spectral efficiency and wideband slope achieved by M-ary PSK signals can be readily obtained by using the formulas in (17). Note that in the noncoherent Rician fading channel, the received bit energy is

$$\frac{E^r_{b,nc}}{N_0} = \frac{(|d|^2+\gamma^2)\text{SNR}\log 2}{C_{M,nc}(\text{SNR})}.$$

Corollary 2: The received bit energy at zero spectral efficiency and wideband slope achieved by M-ary PSK signaling in the noncoherent Rician fading channel are given by

$$\frac{E^r_{b,nc}}{N_0}\bigg|_{\mathsf{C}=0} = \begin{cases} \frac{\pi}{2}\left(1+\frac{1}{\mathsf{K}}\right)\log 2 & M = 2 \\ \frac{4\pi}{M^2\sin^2\frac{\pi}{M}}\left(1+\frac{1}{\mathsf{K}}\right)\log 2 & M \ge 3 \end{cases}$$

and

$$S_{0,nc} = \begin{cases} \frac{3}{\pi-1+\frac{3\pi}{2\mathsf{K}}} & M = 2 \\ 0 & M = 3 \\ \frac{6}{\pi-1+\frac{3\pi}{\mathsf{K}}} & M = 4 \\ \frac{\frac{M^4}{8\pi^2}\sin^4\frac{\pi}{M}}{-\psi(M)+\frac{1}{2\pi\mathsf{K}}M^2\sin^2\frac{\pi}{M}} & M \ge 5 \end{cases}, \qquad (23)$$

respectively, where $\psi(M)$ is given in (18), and $\mathsf{K} = \frac{|d|^2}{\gamma^2}$ is the Rician factor.

As it will be evident in numerical results, generally the

---

[6]It can be easily seen from the smoothness and boundedness of $f_{\theta|s_1}$ in (14) that $C_{M,nc}(\text{SNR})$ is continuous and differentiable in SNR.
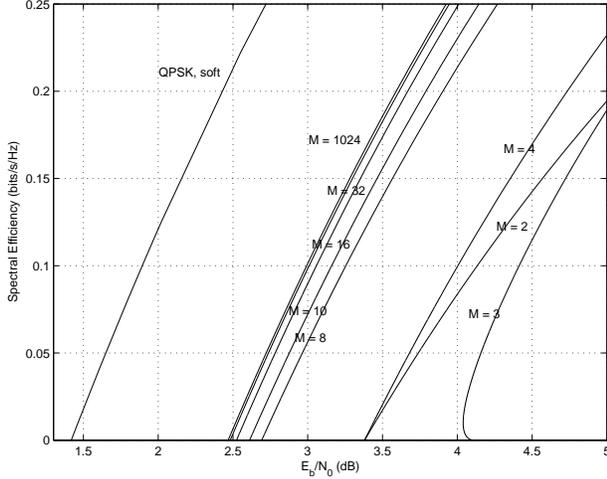
Fig. 1. Spectral efficiency $\mathsf{C}(E_b/N_0)$ vs. bit energy E_b/N_0 for hard-decision detected M-ary PSK with $M = 2, 3, 4, 8, 10, 16, 32, 1024$ and soft-detected QPSK in the noncoherent Rician fading channel with Rician factor $\mathsf{K} = \frac{|d|^2}{\gamma^2} = 1$.

$\left. \frac{E_{b,nc}^r}{N_0} \right|_{\mathsf{C}=0}$ is the minimum bit energy required for reliable transmission when $M \neq 3$. On the other hand, the minimum bit energy is achieved at a nonzero spectral efficiency when $M = 3$. Note that this behavior is not exhibited when 3-PSK signals are soft-detected [9]. Hence, this result is tightly linked to correct-detection and error probabilities which are in general functions of the distances in the signal constellation. Note further that at sufficiently low SNRs, 3-PSK performs worse than 2-PSK (i.e., BPSK), indicating that the decrease in the signal distance from $2\sqrt{\mathcal{E}}$ in 2-PSK to $\sqrt{3\mathcal{E}}$ in 3-PSK has a more dominating effect in the low-SNR regime than the increase in the constellation size M from 2 to 3.

Figure 1 plots the spectral efficiency curves as a function of the bit energy for hard-decision detected PSK with different constellation sizes in the noncoherent Rician fading channel with Rician factor $\mathsf{K} = 1$. As observed in this figure, the information-theoretic analysis conducted in this paper provides several practical design guidelines. We note that although hard-decision detected 2-PSK and 4-PSK achieve the same minimum bit energy of 3.38 dB at zero spectral efficiency, 4-PSK is more efficient at low but nonzero spectral efficiency values due to its wideband slope being twice that of 2-PSK. In the range of spectral efficiency values considered in the figure, 3-PSK performs worse than both 2-PSK and 4-PSK. 3-PSK achieves its minimum bit energy of 4.039 dB at 0.0101 bits/s/Hz. Operation below this level of spectral efficiency should be avoided as it only increases the energy requirements. We further observe that increasing the constellation size to 8 provides much improvement over 4-PSK. 8-PSK achieves a minimum bit energy of 2.692 dB. Note from (23) that $\left. \frac{E_{b,nc}^r}{N_0} \right|_{\mathsf{C}=0}$ is inversely proportional to $M^2 \sin^2 \frac{\pi}{M}$ for $M \geq 3$. Here, we see two competing factors. As M increases, the term M^2 increases and tends to decrease the bit energy requirement while the term $\sin^2 \frac{\pi}{M}$ decreases due to a decrease in the

minimum distance, which is proportional to $\sin \frac{\pi}{M}$ in M-PSK constellation. Hence, when we increase M from 4 to 8, M^2 is the dominant term and we note significant gains. As M is further increased, $\sin^2 \frac{\pi}{M}$ acts more strongly to offset the gains from M^2 and we see diminishing returns. For instance, there is little to be gained by increasing the constellation size more than 32 as 32-PSK achieves a minimum bit energy of 2.482 dB and the minimum bit energy as $M \to \infty$ is 2.468 dB. We find that the wideband slopes of hard-decision detected PSK with $M = 8,10,16,32,$ and 1024 are 0.571, 0.584, 0.599, 0.607, and 0.609, respectively. The similarity of the wideband slope values is also apparent in the figure. Note that the wideband slope of 3-PSK, as predicted, is 0.

For comparison, the spectral efficiency of soft-detected QPSK is also provided in Fig. 1. It has been shown in [13] that under the peak constraint $|x_i|^2 \leq \mathcal{E}$, the bit energy required at zero spectral efficiency and wideband slope in the noncoherent Rician fading channel with Rician factor K are $\left. \frac{E_b}{N_0} \right|_{\mathsf{C}=0} = \left(1 + \frac{1}{\mathsf{K}}\right) \log 2$ and $S_0 = \frac{2\mathsf{K}^2}{(1+\mathsf{K})^2}$, respectively. It is also proven that soft-detected QPSK is optimally efficient achieving these values. Note that when $\mathsf{K} = 1$, the bit energy at zero spectral efficiency is 1.418 dB which is also observed in Fig. 1. Note that even as $M \to \infty$, hard-decision detection presents a loss of 2.468 - 1.418 = 1.05 dB in the minimum bit energy.

B. AWGN Channels

Note that the noncoherent Rician fading channel, in which we have $E\{h_k\} = d$ and $E\{|h_k - d|^2\} = \gamma^2$, specializes to the AWGN channel if we assume $\gamma^2 = 0$. With this assumption, the fading coefficients become deterministic, i.e., $h_k = d$, and the channel model is now $r_k = ds_{x_k} + n_k$ where the channel gain is d. Note also that when we have $\gamma^2 = 0$, (4) becomes the conditional density function of the output given the input in the AWGN channel. Moreover, the maximum likelihood decision rule and decision regions for the AWGN channel are the same as in (5) and (6), respectively. Assuming further that $d = 1$ leads to the standard unfaded Gaussian channel with unit channel gain, i.e., the input-output relation becomes $r_k = s_{x_k} + n_k$. Based on the above observations, we immediately have the following Corollary.

Corollary 3: For the AWGN channel with channel gain d, the first and second derivatives of the PSK capacity at SNR $= 0$ are given by the expressions in (17) if we let $\gamma^2 = 0$. Furthermore, the bit energy, $\left. \frac{E_b^r}{N_0} \right|_{\mathsf{C}=0} = \left. \frac{|d|^2 \text{SNR} \log 2}{C_M(\text{SNR})} \right|_{\mathsf{C}=0}$, and wideband slope expressions are obtained if we let $\gamma^2 = 0$ and hence $\mathsf{K} \to \infty$ in the formulas in (23).

Remark: We should note that the first derivative of the capacity of PSK in the AWGN channel has previously been given in [7] through the bit energy expressions. In addition, Verdú in [8] has provided the second derivative expression for the special case of $M = 4$.

Fig. 2 plots the spectral efficiency curves as a function of the bit energy for hard-decision detected M-ary PSK for various values of M and soft-detected QPSK in the AWGN channel. Conclusions similar to those given for Fig. 1 also apply for Fig. 2. The main difference between the figures is
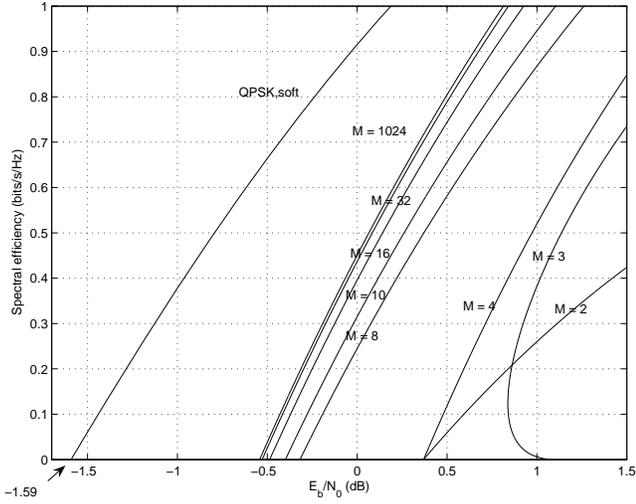
Fig. 2. Spectral efficiency $C(E_b/N_0)$ vs. bit energy E_b/N_0 for hard-decision detected M-ary PSK with $M = 2, 3, 4, 8, 10, 16, 32, 1024$ and soft detected QPSK in the AWGN channel.

that substantially lower bit energies are needed in the AWGN channel. For instance, 2- and 4-PSK now achieve a minimum bit energy of 0.369 dB while 8-PSK attains -0.318 dB. As $M \to \infty$, the minimum bit energy goes to -0.542 dB. We note that higher energy requirements in the noncoherent Rician channel is due to fading and not knowing the channel.

C. Coherent Fading Channels

In coherent fading channels, the fading coefficients $\{h_k\}$ are assumed to be perfectly known at the receiver. We assume that no such knowledge is available at the transmitter. The only requirements on the fading coefficients are that their variations are ergodic and they have finite second moments. Hence, independence of the random variables $\{h_k\}$ is no longer imposed. Due to the presence of receiver channel side information (CSI), maximum likelihood detection is the scaled nearest point detection. In this case, the average capacity is

$$C_{M,c}(\text{SNR}) = \log M + \sum_{l=1}^{M} E_h\{P_{l,1,h} \log P_{l,1,h}\} \qquad (24)$$

where

$$P_{l,1,h} = P(y = l | x = 1, h) = \int_{\frac{(2l-3)\pi}{M}}^{\frac{(2l-1)\pi}{M}} f_{\theta|s_1,h}(\theta|s_1, h)\, d\theta$$

and $f_{\theta|s_1,h}(\theta|s_1, h)$ is given in (25) on the next page with the definition $\text{SNR} = \mathcal{E}/N_0$. Note that if we assume $\gamma^2 = 0$ and replace d by the random channel gain h_k in the noncoherent Rician fading channel, we obtain the model for coherent fading channels. Hence, similarly as in Section III-B, results for coherent channels can be obtained easily by specializing those for the noncoherent Rician channel. Since we are interested in the average capacity (24), expressions will involve the expected values of the random gain h. Hence, we have the following Corollary to Theorem 1.

Corollary 4: The first and second derivatives of $C_{M,c}(\text{SNR})$ in nats per symbol at $\text{SNR} = 0$ are obtained by assuming in (17) $\gamma^2 = 0$, replacing d by h, and taking the expectation of the terms that involve h. The resulting expressions are

$$\dot{C}_{M,c}(0) = \begin{cases} \frac{2}{\pi} E\{|h|^2\} & M = 2 \\ \frac{M^2}{4\pi} \sin^2 \frac{\pi}{M} E\{|h|^2\} & M \geq 3 \end{cases},$$

and

$$\ddot{C}_{M,c}(0) = \begin{cases} \frac{8}{3\pi}\left(\frac{1}{\pi} - 1\right) E\{|h|^4\} & M = 2 \\ \infty & M = 3 \\ \frac{4}{3\pi}\left(\frac{1}{\pi} - 1\right) E\{|h|^4\} & M = 4 \\ \psi(M) E\{|h|^4\} & M \geq 5 \end{cases}$$
$$(26)$$

respectively, where $\psi(M)$ is given in (18).

Note that the first and second derivatives of the capacity at zero SNR are essentially equal to the scaled versions of those obtained in the AWGN channel with $d = 1$. The scale factors are $E\{|h|^2\}$ and $E\{|h|^4\}$ for the first and second derivatives, respectively. In the coherent fading case, we can define the received bit energy as $\frac{E_{b,c}^r}{N_0} = \frac{E\{|h|^2\}\text{SNR}\log 2}{C_M(\text{SNR})}$ since $E\{|h|^2\}\text{SNR}$ is the average received signal-to-noise ratio. It immediately follows from Corollary 4 that $E_b^r/N_0|_{\text{C}=0}$ in the coherent fading channel is the same as that in the AWGN channel. On the other hand, the wideband slope is scaled by $(E\{|h|^2\})^2/E\{|h|^4\}$. Fig. 3 plots the spectral efficiency curves as a function of bit energy for hard-decision detected M-ary PSK and soft detected QPSK in the coherent Rayleigh fading channel. Comparison of Fig. 2 and Fig. 3 reveals that the bit energy levels required at zero spectral efficiency are indeed the same for both cases. However, the presence of fading induces a performance penalty by reducing the wideband slope with a factor of $E\{|h|^2\}^2/E\{|h|^4\} = 1/2$. Therefore, at low but nonzero spectral efficiencies, the same bit energy as in the AWGN channel can be achieved at the cost of reduced spectral efficiency.

IV. ENERGY EFFICIENCY OF ORTHOGONAL SIGNALING

As discussed in Section I, orthogonal signaling is optimally energy efficient in the infinite bandwidth regime even if the receiver performs hard-decision detection. For instance, PPM with vanishing duty cycle or M-ary FSK as $M \to \infty$ achieves the minimum bit energy of -1.59 dB. In this section, we analyze the non-asymptotic energy efficiency of orthogonal signaling. We consider on-off FSK (OOFSK) modulation in which FSK is combined with on-off keying (or equivalently PPM) and peakedness is introduced in both time and frequency. The study of OOFSK modulation enables us to provide a general unified analysis of orthogonal signaling as OOFSK can be reduced to OOK and FSK with the appropriate choice of parameters.

A. OOFSK Modulation

1) AWGN Channels: In this section, we consider the transmission of OOFSK signals. We again assume that the received signal is hard-decision detected at the receiver. In [20] and

$$f_{\theta|s_1,h}(\theta|s_1,h) = \frac{1}{2\pi}\,e^{-|h|^2\text{SNR}} + \sqrt{\frac{|h|^2\text{SNR}}{\pi}}\cos\theta\,e^{-|h|^2\text{SNR}\sin^2\theta}\left(1 - Q(\sqrt{2|h|^2\text{SNR}\cos^2\theta})\right) \qquad (25)$$
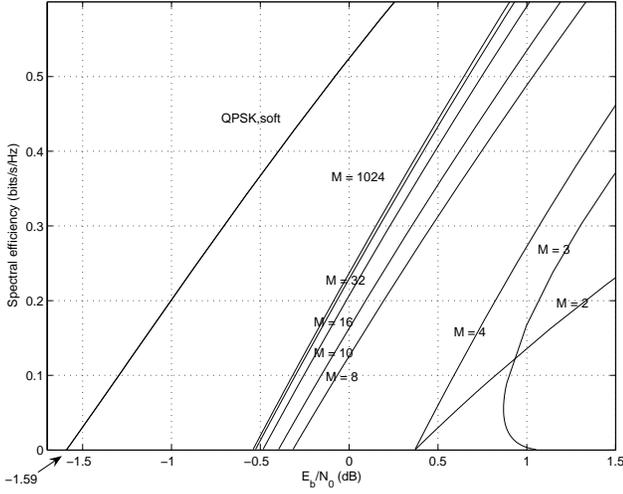
Fig. 3. Spectral efficiency $\mathsf{C}(E_b/N_0)$ vs. bit energy E_b/N_0 for hard-decision detected M-ary PSK with $M = 2, 3, 4, 8, 10, 16, 32, 1024$ and soft detected QPSK in the coherent Rayleigh fading channel.

[22], maximum a posteriori probability (MAP) detection rule for OOFSK modulation is identified and the error probability expressions are obtained. We initially consider the AWGN channel as the results for this channel will immediately imply similar conclusions for fading channels. The optimal detection rule in the AWGN channel is given by the following: $\mathbf{s}_i$ for $i \neq 0$ is detected if

$$|r_i|^2 > |r_j|^2 \quad \forall j \neq i \quad \text{and} \quad |r_i|^2 > \tau \qquad (27)$$

where $\tau = \begin{cases} \frac{[I_0^{-1}(\xi)]^2}{4\alpha^2} & \xi \geq 1 \\ 0 & \xi < 1 \end{cases}$, $\xi = \frac{M(1-\nu)\,e^{\alpha^2}}{\nu}$, and $\alpha^2 = \frac{\text{SNR}}{\nu}$. Above, r_i is the ith component of the received vector $\mathbf{r}$. $\mathbf{s}_0$ is detected if $|r_i|^2 < \tau\ \forall i$. Note that since $\mathbf{s}_0 = (0,\ldots,0)$, detection of $\mathbf{s}_0$ is essentially the detection of no transmission. Note further that the detection rule in (27) together with the rule for $\mathbf{s}_0$ can be regarded as energy detection. After detection, the channel can now be seen as a discrete channel with $M + 1$ inputs and $M + 1$ outputs. From the error probability analysis in [20] and [22], we have the expressions in (28) through (31) on the next page for the transition probabilities in the AWGN channel. In these expressions, $Q_1(\cdot,\cdot)$ is the Marcum Q-function [16], and I_0^{-1} is the functional inverse of the zeroth order modified Bessel function of the first kind. The rates achieved by the M-ary OOFSK modulation with duty cycle ν and equiprobable FSK signals is given by (33) on the next page. If M-ary OOFSK signals have a symbol duration of T, the bandwidth requirement is $\frac{M}{T}$ and the spectral efficiency is given by $\frac{I_M(\text{SNR},\nu)}{\frac{M}{T}} = \frac{I_M(\text{SNR},\nu)}{M}$ bits/s/Hz.

It is shown in [21] that in the AWGN channel, the first

derivative of the capacity of soft-detected OOFSK is zero at $\text{SNR} = 0$. For the sake of completeness, we provide this result below.

Theorem 2: The first derivative of the capacity at zero SNR achieved by soft-detected M-ary OOFSK signaling with a fixed duty factor $\nu \in (0,1]$ over the AWGN channel is zero, and hence the bit energy required at zero spectral efficiency is infinite.

Proof: See [21].

Since hard-decision detection does not increase the capacity, we immediately have the following Corollary to Theorem 2.

Corollary 5: The first derivative at zero SNR of the achievable rates of hard-decision-detected M-ary OOFSK transmission with a fixed duty factor $\nu \in (0,1]$ over the AWGN channel is zero i.e., $\dot{I}_M(0,\nu) = 0$, and hence the bit energy required at zero spectral efficiency is infinite, i.e.,

$$\left.\frac{E_b}{N_0}\right|_{I=0} = \frac{\log 2}{\dot{I}_M(0,\nu)} = \infty. \qquad (34)$$

On the other hand, we know from [1] and [8] that if the duty cycle ν vanishes simultaneously with SNR, the minimum bit energy of -1.59 dB can be achieved. The following result identifies the rate at which ν should decrease as SNR gets smaller.

Theorem 3: Assume that $\nu = \frac{\text{SNR}}{(1+\epsilon)\log\frac{1}{\text{SNR}}}$ for $\text{SNR} < 1$ and for some $\epsilon > 0$. Then, we have

$$\lim_{\epsilon\to 0}\lim_{\text{SNR}\to 0}\frac{I_M(\text{SNR},\nu)}{\text{SNR}} = 1 \qquad (35)$$

and hence

$$\lim_{\epsilon\to 0}\lim_{\text{SNR}\to 0}\frac{\text{SNR}\log 2}{I_M(\text{SNR},\nu)} = \log 2 = -1.59 \text{ dB}. \qquad (36)$$

Proof: Note that as $\text{SNR} \to 0$, $\nu \to 0$ and $\alpha^2 = \frac{\text{SNR}}{\nu} = (1+\epsilon)\log\frac{1}{\text{SNR}} \to \infty$. It can also be seen that $\xi \to \infty$ and $\tau \to \infty$ as SNR diminishes. From (28), we immediately note that $P_{0,0} \to 1$ and $P_{l,0} \to 0$ for $l = 2,\ldots,M$. In (29), all the terms in the summation other than for $n = 0$ vanishes because $\alpha^2 \to \infty$. Therefore, in order for $P_{l,l}$ for $l = 1,\ldots,M$ to approach 1, we need $Q_1(\sqrt{2}\,\alpha,\sqrt{2\tau}) \to 1$. Also note that if $Q_1(\sqrt{2}\,\alpha,\sqrt{2\tau}) \to 1$, then we can observe from (30) and (31) that $P_{0,l} \to 0$ and $P_{l,m} \to 0$. Hence, eventually all crossover error probabilities will vanish and correct detection probabilities will be 1.

In [16], it is shown that $Q_1(a, a\zeta) \geq 1 - \frac{\zeta}{1-\zeta}e^{-\frac{a^2(1-\zeta)^2}{2}}$ $0 \leq \zeta < 1$. From this lower bound we can immediately see that $\lim_{\text{SNR}\to 0}Q_1(\sqrt{2}\,\alpha,\sqrt{2\tau}) = 1$ if $\lim_{\text{SNR}\to 0}\frac{\tau}{\alpha^2} < 1$. Note that both α^2 and τ grow without bound as $\text{SNR} \to 0$. Recall that $\tau = \frac{[I_0^{-1}(\xi)]^2}{4\alpha^2}$. Equivalently, we have $I_0(\sqrt{4\alpha^2\tau}) = \xi = \frac{M(1-\nu)\,e^{\alpha^2}}{\nu}$. Using the asymptotic form $I_0(x) = \frac{1}{\sqrt{2\pi x}}\,e^x + \mathcal{O}\left(\frac{1}{x^{3/2}}\right)$ [26] for large x, we can easily show that $\lim_{\text{SNR}\to 0}\frac{\tau}{\alpha^2} = \left(\frac{1+\epsilon/2}{1+\epsilon}\right)^2 < 1$ $\forall \epsilon > 0$ if

$$P_{0,0} = (1 - e^{-\tau})^M, \quad \text{and} \quad P_{l,0} = \frac{1}{M}(1 - (1 - e^{-\tau})^M) \quad \text{for } l = 1, 2, \ldots, M, \tag{28}$$

$$P_{l,l} = \sum_{n=0}^{M-1} \frac{(-1)^n}{n+1} \binom{M-1}{n} e^{-\frac{n}{n+1}\alpha^2} Q_1\left(\sqrt{\frac{2}{n+1}}\,\alpha, \sqrt{2(n+1)\tau}\right) \quad \text{for } l = 1, 2, \ldots, M, \tag{29}$$

$$P_{0,l} = (1 - e^{-\tau})^{M-1}\left(1 - Q_1\left(\sqrt{2}\,\alpha, \sqrt{2\tau}\right)\right) \quad \text{for } l = 1, 2, \ldots, M, \tag{30}$$

$$P_{l,m} = \frac{1}{M-1}(1 - P_{m,m} - P_{0,m}) \quad \text{for all } l \neq 0, m \neq 0, \text{and } l \neq m \tag{31}$$

$$I_M(\text{SNR}, \nu) = H(y) - H(y|x) \tag{32}$$
$$= -\left((1 - \nu)P_{0,0} + \nu P_{0,1}\right)\log\left((1 - \nu)P_{0,0} + \nu P_{0,1}\right)$$
$$- M\left((1 - \nu)P_{1,0} + \frac{\nu}{M}P_{1,1} + \frac{(M-1)\nu}{M}P_{1,2}\right)\log\left((1 - \nu)P_{1,0} + \frac{\nu}{M}P_{1,1} + \frac{(M-1)\nu}{M}P_{1,2}\right)$$
$$+ (1 - \nu)\left(P_{0,0}\log P_{0,0} + M P_{1,0}\log P_{1,0}\right) + \nu\left(P_{0,1}\log P_{0,1} + P_{1,1}\log P_{1,1} + (M-1)P_{2,1}\log P_{2,1}\right). \tag{33}$$

$\nu = \frac{\text{SNR}}{(1+\epsilon)\log\frac{1}{\text{SNR}}}$. Therefore, if ν decays at this rate, the error probabilities go to zero. It can then be shown that $\lim_{\text{SNR}\to 0}\frac{I_M(\text{SNR},\nu)}{\text{SNR}} = \frac{1}{1+\epsilon}$. Since results hold for any $\epsilon > 0$, letting $\epsilon \to 0$ gives the desired result. $\square$

We note that Zheng *et al.* have shown in [19] that the low SNR capacity of unknown Rayleigh fading channel can be approached by on-off keying if $\frac{\log\frac{1}{\text{SNR}}}{\log\log\frac{1}{\text{SNR}}} \leq \alpha^2 \leq \log\frac{1}{\text{SNR}}$. We see a similar behavior here when FSK signals are sent over the AWGN channel and energy detected.

2) Fading Channels: In coherent fading channels where the receiver has perfect knowledge of the fading coefficients, the transition probabilities are the same as those in (28)-(31) with the only difference that we now have $\alpha^2 = \frac{\text{SNR}}{\nu}|h|^2$. As a result, the achievable rates $I_M(\text{SNR}, \nu, |h|^2)$ are also dependent on the fading coefficients and average achievable rates are obtained by finding the expected value $I_{M,c}(\text{SNR}, \nu) = E_{|h|^2}\{I_M(\text{SNR}, \nu, |h|^2)\}$.

In noncoherent Rician fading channels with $E\{h\} = d$ and $E\{|h - d|^2\} = \gamma^2$, the transition probabilities [20], [22] are given by (37) through (39) on the next page. In these expressions,

$$\tau = \begin{cases} \Phi^{-1}(\xi) & \xi \geq 1 \\ 0 & \xi < 1 \end{cases}$$

where

$$\Phi(x) = e^{\frac{\alpha^2\gamma^2 x}{1+\alpha^2\gamma^2}} I_0\left(\frac{2\sqrt{x\,\alpha^2|d|^2}}{1+\alpha^2\gamma^2}\right), \quad \text{and} \tag{40}$$

$$\xi = \frac{M(1-\nu)}{\nu}(1+\alpha^2\gamma^2)\, e^{\frac{\alpha^2|d|^2}{1+\alpha^2\gamma^2}}. \tag{41}$$

The achievable rates, $I_{M,nc}(\text{SNR}, \nu)$, can be obtained from (33).

Since the presence of fading unknown at the transmitter does not improve the performance, we readily conclude that the bit energy requirements in fading channels still grow without bound with vanishing SNR.

Corollary 6: The first derivatives at zero SNR of the achievable rates $I_{M,c}(\text{SNR}, \nu)$ and $I_{M,nc}(\text{SNR}, \nu)$ are equal to zero, i.e., $\dot{I}_{M,c}(0, \nu) = \dot{I}_{M,nc}(0, \nu) = 0$. Therefore, the bit energy required at zero spectral efficiency is infinite in both coherent and noncoherent fading channels, i.e., $\left.\frac{E_{b,c}}{N_0}\right|_{I=0} = \left.\frac{E_{b,nc}}{N_0}\right|_{I=0} = \infty$.

On the other hand, in noncoherent fading channels, if $|d| \geq 1$, then following the same steps as in the proof of Theorem 3, we can show that the minimum bit energy of -1.59 dB is achieved as SNR $\to 0$ if $\nu = \frac{\text{SNR}}{(1+\epsilon)\log\frac{1}{\text{SNR}}}$.

B. FSK Modulation

Recall that if we set $\nu = 1$ in OOFSK modulation, we recover the regular FSK modulation. Similarly, choosing $\nu = 1$ in the decision rules and transition probabilities leads to the corresponding expressions for FSK. For instance, when $\nu = 1$, $\tau = 0$ in the decision rule (27) of OOFSK modulation. Therefore, $\mathbf{s}_i$ is declared as the detected signal if the ith component of the received vector $\mathbf{r}$ has the largest energy, i.e., $|r_i|^2 > |r_j|^2 \quad \forall j \neq i$. This is the well-known noncoherent detection of FSK signals. Furthermore, Theorem 2 and Corollaries 5 and 6 are valid for all $\nu \in (0, 1]$ and hence for $\nu = 1$ as well. Therefore, the same conclusions are automatically drawn for FSK modulation. Hence, although FSK is energy efficient asymptotically as $M \to \infty$, operating at very low SNR levels with fixed M is extremely energy inefficient as the bit energy requirement increases without bound with decreasing SNR. As a result, the minimum bit energy is achieved at a nonzero spectral efficiency, the value of which can be found through numerical analysis. We finally note that when FSK modulation is considered, the achievable rates are indeed the capacity of FSK modulation as it is well-known that hard-decision detection capacity is achieved with equiprobable signals.

C. Numerical Results

In this section, we provide numerical results and initially concentrate on FSK modulation due to its widespread and

$$P_{0,0} = (1 - e^{-\tau})^M, \quad \text{and} \quad P_{l,0} = \frac{1}{M}(1 - (1 - e^{-\tau})^M) \quad \text{for } l = 1, 2, \ldots, M, \tag{37}$$

$$P_{l,l} = \sum_{n=0}^{M-1} (-1)^n \binom{M-1}{n} \frac{e^{-\frac{n\alpha^2|d|^2}{n(1+\gamma^2\alpha^2)+1}}}{n(1+\gamma^2\alpha^2)+1} Q_1\left(\sqrt{\frac{2\alpha^2|d|^2}{(1+\gamma^2\alpha^2)(n(1+\gamma^2\alpha^2)+1)}}, \sqrt{\frac{2(n(1+\gamma^2\alpha^2)+1)\tau}{(1+\gamma^2\alpha^2)}}\right) \quad \text{for } l \neq 0,$$

$$P_{0,l} = (1 - e^{-\tau})^{M-1}\left(1 - Q_1\left(\sqrt{\frac{2\alpha^2|d|^2}{1+\gamma^2\alpha^2}}, \sqrt{\frac{2\tau}{1+\gamma^2\alpha^2}}\right)\right) \quad \text{for } l = 1, 2, \ldots, M, \tag{38}$$

$$P_{l,m} = \frac{1}{M-1}(1 - P_{m,m} - P_{0,m}) \quad \text{for all } l \neq 0, m \neq 0, \text{and } l \neq m \tag{39}$$
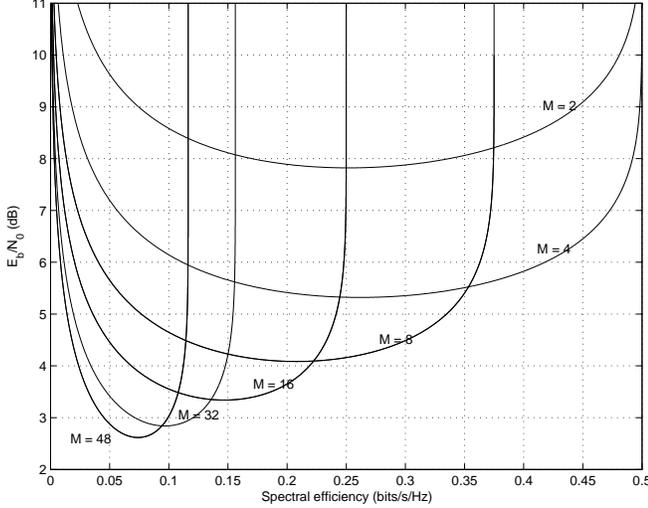
Fig. 4. Bit energy E_b/N_0 vs. Spectral efficiency $\mathsf{C}(E_b/N_0)$ for energy-detected M-ary FSK in the AWGN channel.
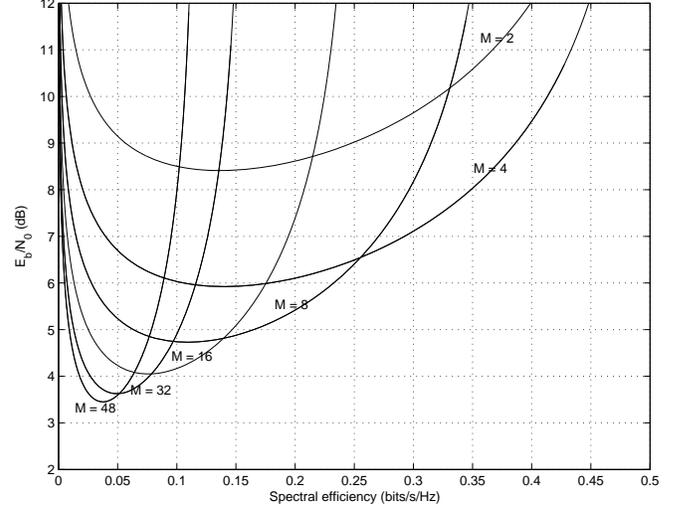
Fig. 5. Bit energy E_b/N_0 vs. Spectral efficiency $\mathsf{C}(E_b/N_0)$ for energy-detected M-ary FSK in the coherent Rician fading channel with Rician factor $\mathsf{K} = 1$.

frequent use. Fig. 4 plots the bit energy E_b/N_0 curves as a function of spectral efficiency for M-ary FSK in the AWGN channel for different values of M. In all cases, we observe that the minimum bit energy is achieved at a nonzero spectral efficiency C^*, and the bit energy requirements increase to infinity as spectral efficiency decreases to zero. Hence, operation below C^* should be avoided. Another observation is that the minimum bit energy and the spectral efficiency value at which the minimum is achieved decrease with increasing M. For instance, when $M = 2$, the minimum bit energy is 7.821 dB and is achieved at $\mathsf{C}^* = 0.251$ bits/s/Hz. If the value of M is increased to 48, the minimum bit energy decreases to 2.617 dB and is now attained at $\mathsf{C}^* = 0.074$ bits/s/Hz. Another fact is that as M increases, the minimum bit energy is achieved at a higher SNR value. Indeed, we can show that

$$\lim_{\substack{\epsilon \to 0 \\ M \to \infty}} \left.\frac{C_M(\text{SNR})}{\text{SNR}}\right|_{\text{SNR}=(1+\epsilon)\log M} = \lim_{\substack{\epsilon \to 0 \\ M \to \infty}} \frac{C_M((1+\epsilon)\log M)}{(1+\epsilon)\log M} \tag{42}$$

$$= \lim_{\epsilon \to 0} \frac{1}{1+\epsilon} \lim_{M \to \infty} P_{1,1} = 1. \tag{43}$$

Hence, if SNR grows logarithmically with increasing M, the bit energy $\frac{E_b}{N_0} = \frac{\text{SNR} \log 2}{C_M(\text{SNR})}$ approaches $\log 2 = -1.59$ dB. The

proof of (43) is omitted because Turin [2] has already shown that -1.59 dB is achieved if the signal duration increases as $\log M$, which in turn increases the SNR logarithmically in M.

Figures 5 and 6 plot the bit energy curves for M-ary FSK transmission over coherent and noncoherent Rician fading channels. As predicted, the bit energy levels for all values of M increase without bound as the spectral efficiency decreases to zero. Due to the presence of fading, the minimum bit energies have increased with respect to those achieved in the AWGN channel. For instance, when $M = 48$, the minimum bit energies are now $E_b/N_{0\min} = 3.45$ dB in the coherent Rician fading channel and $E_b/N_{0\min} = 4.23$ dB in the noncoherent Rician fading channel. We again observe that the minimum bit energy decreases with increasing M. Fig. 7 provides the minimum bit energy values as a function of M in the AWGN and noncoherent Rician fading channels with different Rician factors. In all cases, the minimum bit energy decreases with increasing M. However, Fig. 7 indicates that approaching -1.59 dB is very slow and demanding in M. In this figure, we also note the energy penalty due to the presence of unknown fading. But, as the Rician factor K increases, the noncoherent Rician channel approaches to the AWGN channel and so do the minimum bit energy requirements. Figures 8 and 9 plot the spectral efficiencies and average

9
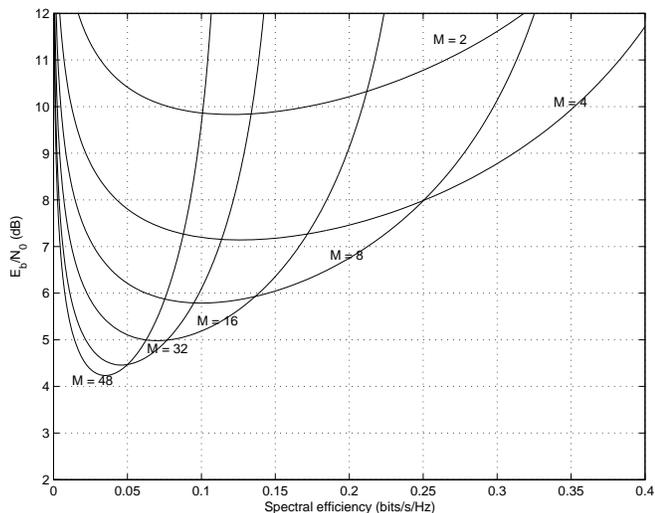
Fig. 6. Bit energy E_b/N_0 vs. Spectral efficiency $\mathsf{C}(E_b/N_0)$ for energy-detected M-ary FSK in the noncoherent Rician fading channel with Rician factor $\mathsf{K} = 1$.

Fig. 8. Spectral efficiency at which $E_b/N_{0\min}$ is achieved vs. M for M-ary FSK in the AWGN channel and noncoherent Rician fading channels with Rician factors $\mathsf{K} = 0, 1, 4, 9, 16$.
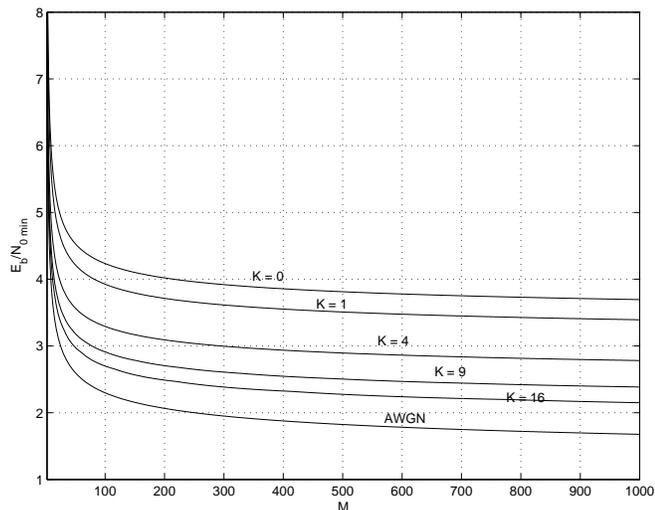
Fig. 7. Minimum bit energy $E_b/N_{0\min}$ vs. M for M-ary FSK in the AWGN channel and noncoherent Rician fading channels with Rician factors $\mathsf{K} = 0, 1, 4, 9, 16$.

Fig. 9. SNR at which $E_b/N_{0\min}$ is achieved vs. M for M-ary FSK in the AWGN channel and noncoherent Rician fading channels with Rician factors $\mathsf{K} = 0, 1, 4, 9, 16$.

received SNR values at which $E_b/N_{0\min}$ is achieved as a function of M. As we have also observed in Figs. 4 and 6, we see in Fig. 8 that the spectral efficiency at which $E_b/N_{0\min}$ is achieved decreases with increasing M. From Fig. 8, we further note that the required spectral efficiencies are lower and hence the bandwidth requirements are higher in noncoherent fading channels. In Fig. 9, we observe that the SNR levels at which $E_b/N_{0\min}$ is achieved increases with increasing M. As predicted by (43), SNR increases logarithmically with M in the AWGN channel. Similar rates of increase are also noted for the noncoherent fading channel.

Figs. 10 and 11 plot the bit energies as a function of spectral efficiency of 8-OOFSK with different duty cycle factors in the AWGN and noncoherent Rayleigh fading channels. We immediately observe that decreasing the duty cycle ν lowers

the minimum bit energy. Hence, increasing the signal peakedness in the time domain improves the energy efficiency. In the AWGN channel, while regular 8-FSK (8-OOFSK with $\nu = 1$) has $E_b/N_{0\min} = 4.08$ dB, 8-OOFSK with $\nu = 0.01$ has $E_b/N_{0\min} = 2.017$ dB. However, this energy gain is obtained at the cost of increased peak-to-average ratio. We also note that unknown fading again induces a energy penalty with respect to that achieved in the AWGN channel as observed by comparing Figs. 10 and 11.

V. CONCLUSION

In this paper, we have analyzed the impact of hard-decision detection on the energy efficiency of phase modulation and frequency modulation together with on-off keying. We have obtained closed-form expressions for the first and second
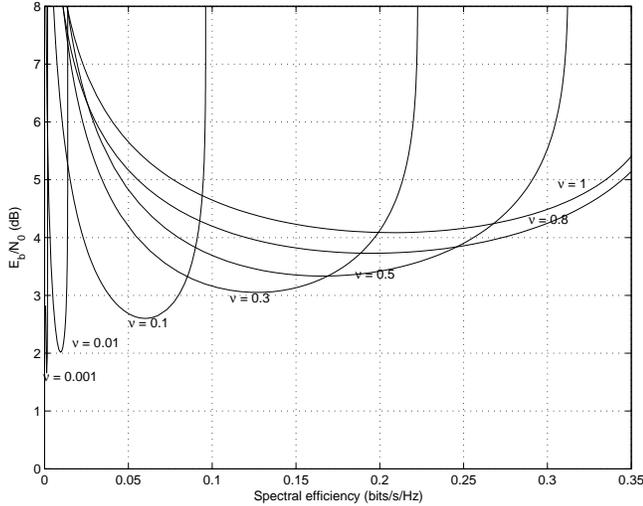
Fig. 10. Bit energy E_b/N_0 vs. Spectral efficiency $\mathsf{C}(E_b/N_0)$ for 8-OOFSK in the AWGN channel. The duty cycle values are $\nu = 1, 0.8, 0.5, 0.3, 0, 1, 0.01$ and 0.001.
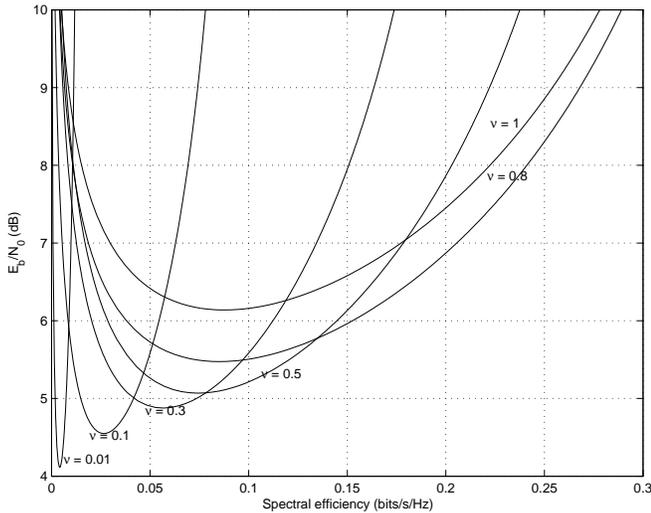
Fig. 11. Bit energy E_b/N_0 vs. Spectral efficiency $\mathsf{C}(E_b/N_0)$ for 8-OOFSK in the noncoherent Rayleigh fading channel. The duty cycle values are $\nu = 1, 0.8, 0.5, 0.3, 0, 1, 0.01$ and 0.001.

derivatives of the M-ary PSK capacity in AWGN, coherent fading, and noncoherent Rician fading channels. Subsequently, we have found closed-form expressions for the bit energy required at zero spectral efficiency and wideband slope, and quantified the loss in energy efficiency incurred by hard-decision detection and channel fading. The inefficiency of 3-PSK at very low SNRs is noted. We have also considered energy detected M-ary OOFSK transmission over the AWGN and fading channels. We have shown that bit energy requirements grow without bound as SNR vanishes for any fixed duty cycle value. Results are easily specialized to FSK modulation as well. Through numerical results, we have investigated the value of the minimum bit energy for different values of M in various channels. We have shown through numerical results that the minimum bit energy decreases with decreasing duty

cycle and increasing M. We have proved that if the duty cycle decreases as $\frac{\text{SNR}}{\log \frac{1}{\text{SNR}}}$, the minimum bit energy of -1.59 dB can be approached.

APPENDIX

A. Proof of Theorem 1

The main approach is to obtain $\dot{C}_{M,nc}(0)$ and $\ddot{C}_{M,nc}(0)$ by first finding the derivatives of the transition probabilities $\{P_{l,1}\}$. This can be accomplished by finding the first and second derivatives of $f_{\theta|s_1}$ with respect to SNR. However, the presence of $\sqrt{\frac{|d|^2 \text{SNR}}{\pi(\gamma^2 \text{SNR}+1)}}$ in the second term of (14) complicates this approach because it leads to the result that $\left.\frac{df_{\theta|s_1}}{d\text{SNR}}\right|_{\text{SNR}=0} = \infty$. In order to circumvent this problem, we define the new variable $a = \sqrt{\text{SNR}}$ and consider the conditional density expression in (44) on the next page. Now, the derivative expressions in (45) and (46) on the next page evaluated at $a = 0$ can easily be verified. Using the derivatives of $P_{l,1}$ and performing several algebraic operations, we arrive to the following Taylor expansion of $C_{M,nc}(a)$ at $a = 0$:

$$C_{M,nc}(a) = \phi_1(M)\,a^2 + \phi_2(M)\,a^3 + \phi_3(M)\,a^4 + o(a^4) \quad (47)$$
$$= \phi_1(M)\text{SNR} + \phi_2(M)\text{SNR}^{3/2} + \phi_3(M)\text{SNR}^2 + o(\text{SNR}^2) \quad (48)$$

where (48) follows due to the fact that $a = \sqrt{\text{SNR}}$. In the above expansion, $\phi_1(M)$, $\phi_2(M)$, and $\phi_3(M)$ are given by (49)–(51) on the next page. We immediately conclude from (48) that $\dot{C}_{M,nc}(0) = \phi_1(M)$. Note that the expansion includes the term $\text{SNR}^{3/2}$ which implies that $\ddot{C}_{M,nc}(0) = \pm\infty$ for all M. However, it can be easily seen that $\phi_2(M) = 0$ for all $M \neq 3$, and at $M = 3$, $\phi_2(3) = 0.1718|d|^3$. Therefore, while $\ddot{C}_{3,nc}(0) = \infty$, $\ddot{C}_{M,nc}(0) = 2\phi_3(M)$ for $M \neq 3$. Further algebraic steps and simplification yield (17). $\square$

REFERENCES

[1] M. J. E. Golay, "Note on the theoretical efficiency of information reception with PPM," *Proc. IRE*, vol. 37, pp. 1031, Sept. 1949.
[2] G. L. Turin, "The asymptotic behavior of ideal M-ary systems," *Proc. IRE*, vol. 47, pp. 93-94, Jan. 1959.
[3] I. Jacobs, "The asymptotic behavior of incoherent M-ary communications systems," *Proc. IEEE*, vol. 51, pp. 251-252, Jan. 1963.
[4] J. N. Pierce, "Ultimate performance of M-ary transmissions on fading channels," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 2-5, Jan. 1966.
[5] A. Lapidoth and S. Shamai (Shitz), "Fading channels: How perfect need 'perfect side information' be?," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1118-1134, May 2002.
[6] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1019-1030, Sep. 1990
[7] J. R. Pierce, "Comparison of three-phase modulation with two-phase and four-phase modulation," *IEEE Trans. Commun*, vol. 28, pp. 1098-1099, July 1980.
[8] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1319-1343, June 2002.
[9] G. Kramer, A. Ashikhmin, A. J. van Wijngaarden, and X. Wei, "Spectral efficiency of coded phase-shift keying for fiber-optic communication," *IEEE/OSA J. Lightwave Technol.*, vol. 21, pp. 2438-2445, Oct. 2003.
[10] A. D. Wyner, "Bounds on communication with polyphase coding," *Bell Syst. Tech. J.,* vol. XLV, pp. 523-559, Apr. 1966.
[11] G. Kaplan and S. Shamai (Shitz), "On the achievable information rates of DPSK," *IEE Proceesings*, vol. 139, pp. 311-318, June 1992.
[12] M. Peleg and Shlomo Shamai (Shitz), "On the capacity of the blockwise incoherent MPSK channel," *IEEE Trans. Commun*, vol. 46, pp. 603-609, May 1998.

$$f_{\theta|s_1}(\theta|s_1) = \frac{1}{2\pi} e^{-\frac{|d|^2 a^2}{\gamma^2 a^2+1}} + \frac{|d|a}{\sqrt{\pi(\gamma^2 a^2+1)}} \cos\theta \, e^{-\frac{|d|^2 a^2}{\gamma^2 a^2+1}\sin^2\theta}\left(1 - Q\left(\sqrt{2\frac{|d|^2 a^2}{\gamma^2 a^2+1}\cos^2\theta}\right)\right). \tag{44}$$

$$f_{\theta|s_1}(\theta|s_1)|_{a=0} = \frac{1}{2\pi}, \qquad \left.\frac{df_{\theta|s_1}}{da}\right|_{a=0} = \frac{|d|\cos\theta}{2\sqrt{\pi}}, \qquad \left.\frac{d^2 f_{\theta|s_1}}{da^2}\right|_{a=0} = \frac{|d|^2\cos 2\theta}{\pi} \tag{45}$$

$$\left.\frac{df_{\theta|s_1}^3}{da^3}\right|_{a=0} = -\frac{3|d|\gamma^2\cos\theta}{2\sqrt{\pi}} - \frac{3|d|^3\cos\theta\sin^2\theta}{\sqrt{\pi}}, \qquad \left.\frac{df_{\theta|s_1}^4}{da^4}\right|_{a=0} = -\frac{12|d|^2\gamma^2\cos 2\theta}{\pi} + \frac{6|d|^4\cos^2 2\theta}{\pi} - \frac{8|d|^4\cos^4\theta}{\pi}. \tag{46}$$

$$\phi_1(M) = \frac{M|d|^2}{2\pi}\sin^2\frac{\pi}{M}\sum_{i=1}^{M}\cos^2\frac{2\pi i}{M}, \tag{49}$$

$$\phi_2(M) = \frac{M|d|^3}{\pi\sqrt{\pi}}\left(\sin\frac{\pi}{M}\sin\frac{2\pi}{M} - \frac{M}{6}\sin^3\frac{\pi}{M}\right)\sum_{i=1}^{M}\cos^3\frac{2\pi i}{M}, \tag{50}$$

$$\phi_3(M) = -\frac{M^2|d|^4}{16\pi}\sin^2\frac{2\pi}{M} + \frac{M|d|^4(\pi+2)}{16\pi^2}\sin^2\frac{2\pi}{M}\sum_{i=1}^{M}\cos^2\frac{4\pi i}{M}$$

$$+ |d|^4\left(\left(\frac{M^3}{12\pi^2} - \frac{M}{3\pi}\right)\sin^4\frac{\pi}{M} - \frac{M^2}{2\pi^2}\sin^2\frac{\pi}{M}\sin\frac{2\pi}{M}\right)\sum_{i=1}^{M}\cos^4\frac{2\pi i}{M} + \frac{M^2}{4\pi^2}\sin^2\frac{\pi}{M}\sin\frac{2\pi}{M}\sum_{i=1}^{M}\cos^2\frac{2\pi i}{M}$$

$$- \frac{|d|^2\gamma^2}{2\pi}M\sin^2\frac{\pi}{M}\sum_{i=1}^{M}\cos^2\frac{2\pi i}{M}. \tag{51}$$

[13] M. C. Gursoy, H. V. Poor, and S. Verdú, "The noncoherent Rician fading channel – Part II : Spectral efficiency in the low power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2207-2221, Sept. 2005.

[14] W. Zhang and J. N. Laneman,"How good is phase-shift keying for peak-limited Rayleigh fading channels in the low-SNR regime?," *IEEE Trans. Inform. Theory*, vol. 53, pp. 236 - 251, Jan. 2007.

[15] X. Luo and G. B. Giannakis, "Energy-constrained optimal quantization for wireless sensor networks," IEEE SECON, pp. 272-278, 4-7 Oct. 2004.

[16] M. K. Simon and M.-S. Alouni, "A unified approach to the performance analysis of digital communication over generalized fading channels", *Proc. of the IEEE*, vol. 86, no. 9, pp. 1860-1877, Sept. 1998

[17] W. C. Lindsey, "Error probabilities for Rician fading multichannel reception of binary and N-ary Signals", *IEEE Trans. Inform. Theory*, vol.10, pp. 339-350, Oct. 1964.

[18] W. E. Stark, "Capacity and cutoff rate of noncoherent FSK with nonselective Rician fading," *IEEE Trans. Commun*, vol. 33, pp. 1153-1159, Nov. 1985.

[19] L. Zheng, D. N. C. Tse, and M. Médard "Channel coherence in the low SNR regime," *IEEE Trans. Inform. Theory*, vol. 53, pp. 976-997, March 2007.

[20] Q. Wang and M.C. Gursoy, "Error Performance of OOFSK Signaling over Fading Channels," Proceedings of the 40th Annual Conference on Information Sciences and Systems, Princeton University, Princeton, NJ, March, 22-24, 2006.

[21] M. C. Gursoy, H. V. Poor, S. Verdú, "On-Off frequency-shift keying for wideband fading channels" *EURASIP Journal on Wireless Communications and Networking*, 2006.

[22] M. C. Gursoy, "Error rate analysis for peaky signaling over fading channels," accepted to the IEEE Transactions on Communications, 2007 (available at http://www.ee.unl.edu/faculty-staff/gursoy.shtml and also at http://arxiv.org/abs/0712.3286).

[23] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[24] J. G. Proakis and M. Salehi, *Fundamentals of Communication Systems*. New York: Pearson Prentice Hall, 2005.

[25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[26] E. Butkov, *Mathematical Physics*. Addison-Wesley, 1968.

[27] V. K. Garg, *Wireless Communications and Networking*, Elsevier, 2007.