Capacity investigation of on–off keying in noncoherent channel settings at low SNR

Citation for published version (APA):

DOI:
10.1002/ett.2921

Document status and date:
Published: 01/11/2015

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
ABSTRACT

On–off keying (OOK) has repossessed much new research interest to realize green communication for establishing autonomous sensor networks. To realize ultra-low power wireless design, we investigate the minimum energy per bit required for reliable communication of using OOK in a noncoherent channel setting where envelope detection is applied at the receiver. By defining different OOK channels with average transit power constraints, the achievability of the Shannon limit for both cases of using soft and hard decisions at the channel output is evaluated based on the analysis of the capacity per unit-cost at low signal-to-noise ratio. We demonstrate that in phase fading using hard decisions cannot destroy the capacity only if extremely asymmetric OOK inputs are used with a properly chosen threshold. The corresponding pulse-position modulation scheme is explicitly studied and demonstrated to be a Shannon-type solution. Moreover, we also consider a slow Rayleigh fading scenario where the transmitter and receiver have no access to channel realizations. Throughput per unit-cost results are developed to explore the tradeoff between power efficiency and channel quality for noncoherent OOK using soft and hard decisions. Copyright © 2015 John Wiley & Sons, Ltd.

1. INTRODUCTION

Being perhaps the most basic modulation scheme, on–off keying (OOK) has recently experienced a renewed research interest, mainly from the perspective of green communication of wireless sensor networks based on ultra-low power (ULP) wireless design [1]. ULP concepts play an important role in the design of energy-autonomous systems, such as body area networks [2]. In such systems, wireless sensor nodes are supposed to be driven by energy scavenging, which results in a critical power budget [3, 4]. Therefore, constrained by transmit power, autonomous wireless systems typically work in a low signal-to-noise ratio (SNR) environment, where accurate estimation of carrier phases is not so easy to obtain, because these methods are typically relatively power-hungry. Hence, simple but robust transmission schemes are highly demanded for realizing green communication for sensor networks in noncoherent channel settings.

Because of the presence of phase noise, some basic modulation schemes that are often used at low SNR, e.g. binary phase-shift keying (BPSK), can not be reliably used. Owing to the good balance between performance and complexity, OOK modulation then becomes an interesting candidate for making low-rate and high-reliability communications possible at ultra-low power. Noncoherent detection of OOK only based on received pulse energy that can result in very low-power hardware realizations [5], therefore, becomes a promising technique for applying in practice. In order to achieve optimal energy efficiency, we are motivated to investigate the capacity behaviour of noncoherent OOK within Gaussian background noise in the low-SNR regime. Because receivers can be based either on hard or soft decisions, we need to investigate capacity behaviour for both these alternatives.

As a fundamental measure of power efficiency at low SNR, the minimum received energy per information bit required for reliable transmission, which is denoted as $E_{b,\text{min}}^r$, becomes more meaningful in ULP design for green communication. For a complex-valued additive white noise channel,
where $E_r^n$ is the received symbol energy, and $N_0$ denotes the noise power spectral density of the corresponding continuous-time channel. Because $C$ is concave in $E_r^n$, $E_{b_{\text{min}}} = E_r^n/C$ when $E_r^n$ vanishes, see [6, eqn. 34]. This gives the well-known result that the minimum required received bit energy normalized to the background noise is

$$\frac{E_b}{N_{0_{\text{min}}}} = \lim_{E_r^n \to 0} \frac{E_r^n}{C N_0} = \ln 2,$$

which is approximately $-1.59$ dB. For convenience, we refer to this $\frac{E_b}{N_{0_{\text{min}}}}$, namely $\ln 2$, as the Shannon limit.

Investigations of OOK for $E_{b_{\text{min}}}$ relate to some classical results in information theory. For the coherent case where the channel phase is known at the receiver, Golay [7] and Wozencraft and Jacobs [8] demonstrated that pulse-position modulation (PPM), which essentially is a fixed-pattern OOK scheme, approaches the Shannon limit as the duty cycle vanishes. This holds both for hard (Golay) and soft decisions (Wozencraft and Jacobs, although the proof involves threshold decoding). Very recently, Koch and Lapidoth [9] found out that if one-bit quantization is used, the capacity per unit-cost can only be achieved by flash signalling [10, Def. 2] on the real-valued Gaussian channel, which implies that the Shannon limit is achievable by using coherent OOK with making binary decisions.

For the noncoherent case, Jacobs [11] showed that the Shannon limit can be approached either by applying $m$-ary frequency-shift keying with energy detection based on soft decisions when the occupied bandwidth goes to infinity. So, this implies that the minimum $E_b^n/N_0$ of noncoherent OOK with soft decisions can also reach $\ln 2$. This result is implicitly pointed out by Verdú [6] based on the capacity per unit-cost study of additive complex-valued Gaussian channels with soft channel outputs (soft decisions).

Previous results only relate to phase fading scenarios where the fixed channel magnitude is known whilst the random channel phase is unknown to the receiver. Previous results relate to the scenario where only phase fading is considered. Concerning magnitude fading, [6] and [9] also studied the Shannon limit achievement of fast fading channels with soft and hard outputs, respectively. It is shown that the Shannon limit can be achieved in fast fading except for the case of noncoherent reception with a binary quantizer.

Consequently, we are encouraged to explicitly investigate the capacity behaviour for noncoherent channel settings at low SNR, by constraining ourselves to using only one-off symbols as channel inputs. Although first the channel outputs are assumed to be soft, it is also interesting and important to study noncoherent OOK with hard decisions, because high-resolution quantization is not always available for ULP design. Moreover, because ULP design implicitly means targeting at the applications with low mobility [12], it is more reasonable to model the wireless environment as a slow fading channel when channel magnitude effect is involved. Therefore, it makes sense to investigate the behaviour of noncoherent OOK in slow magnitude-fading and varying-phase settings.

In this paper, a binary-in soft-out (BISO) channel and a binary-in binary-out (BIBO) channel subject to constrained average transmitted power are used as models to explore the minimum $E_r^n$ for noncoherent OOK using soft and hard decisions. By deriving and investigating the capacity per unit-cost [10] and the wideband slope [6] of noncoherent OOK channels, the main findings of the paper include the following:

- In phase fading, similar to the case of using soft decisions, noncoherent OOK with hard decisions can also achieve the Shannon limit only by using asymmetric inputs and unbounded thresholds. This result demonstrates that in a noncoherent channel using hard decisions will not destroy the capacity per unit-cost when only phase fading is present.
- In phase fading, similar to the case with noncoherent PPM with soft decisions, noncoherent PPM with hard decisions achieves the Shannon limit by setting a proper threshold.
- In slow (quasi-static) Rayleigh fading, throughput per unit-cost results are derived for both cases of using soft and hard decisions. It turns out that using hard decisions will not degrade the capacity when the system is not in outage if no channel estimates are available at both transmitter and receiver.

Moreover, some other contributions are

- An equivalent noncoherent channel that consists of the concatenation of a coherent channel and an ‘envelop detection channel’ is developed to serve the basis for the capacity analysis for noncoherent OOK with soft and hard decisions.
- By viewing the studied channel as a class of OOK channels parameterized by input amplitudes, a simple capacity analysis is obtained for both soft-output and hard-output OOK channels.
- Noncoherent OOK with soft decisions in phase fading is explicitly investigated. By using the approach we developed, it is shown that this scheme can achieve the Shannon limit, which is consistent with the work of Verdú [6].

The rest of the paper is organized as follows. In Section 2, we describe the transmission and channel models based on noncoherent OOK for both soft and hard decisions. Then, we derive the capacity per unit-cost and wideband slope of the defined noncoherent OOK channels.
in Section 3 and apply them to investigate the capacity behaviour of phase fading channels in low-SNR regime in Section 4, where noncoherent PPM with hard decisions is also considered. Section 5 studies the throughput per unit-cost for the OOK channels in slow fading. In the end, Section 6 concludes the paper.

2. NONCOHERENT TRANSMISSION AND CHANNEL MODELS

In this section, we present OOK transmission models with soft and hard outputs (decisions) based on envelope detection. To find the best way of using on–off symbols to convey information (i.e. to approach the Shannon limit), we focus on constraining the average transmitted energy per symbol $E_s$ and do not limit the amplitude of transmitted ‘on’ symbols. When the amplitude of the ‘on’ symbol goes to infinity, such a scheme can be considered as a special type of flash signalling as defined in [6, Def. 2].

2.1. Noncoherent on–off keying transmission

We consider discrete-time memoryless transmission of real-valued ‘on–off’ symbols $X_i \in \{0, A\}$, where $i$ is the time index and where $A \in \mathbb{R}_+$, over a complex-valued AWGN channel. The receiver that observes the output of the channel performs envelope detection as is illustrated in Figure 1.

For the channel realizations $h_i = |h_i|e^{j\theta_i}$ for all $i$, the carrier phases $\theta_i$ are assumed to be independent of each other and uniformly distributed over $(0, 2\pi)$, throughout this paper.

In the analysis that follows, we mainly focus on the case where the fading magnitudes $|h_i|$ are assumed to be unity for all $i$. Slowly fading magnitudes are considered later in our study, in Section 5, where the magnitude is assumed to be fixed over a block of transmissions and drawn from a Rayleigh distribution for each block. First, however, we restrict ourselves to the case where $|h_i| = 1$.

For the additive noise, we assume that the variables $N_i$ are independent of each other and identically distributed (i.i.d.), and more specifically $N_i \sim \mathcal{CN}(0, 2\sigma^2)$, that is, the noise variables are circularly symmetric complex Gaussians with mean zero and variance $\sigma^2$ in both the real and imaginary part. If the corresponding continuous-time white Gaussian channel has a two-sided noise power spectral density $N_0/2$, the noise variance in our model is given by $\sigma^2 = N_0/2$.

Because the channel is memoryless, we will omit the time subscript $i$. Therefore, due to the phase fading, the received complex signal $R$ before envelope detection can be represented as

$$R = Xe^{j\theta} + N,$$

where the functions $\Re\{\cdot\}$ and $\Im\{\cdot\}$ represent the real and imaginary parts of their complex argument, respectively. Next, by applying envelope detection, the soft output of envelope detector $Y$ is

$$Y = \sqrt{\left(\cos \theta + \Re\{N\}\right)^2 + \left(\sin \theta + \Im\{N\}\right)^2},$$

which is real-valued. According to [13], for a certain fixed $A$, the possible probability densities of $Y$ are Rayleigh for $X = 0$, that is,

$$p(y|X = 0) = \frac{y}{\sigma^2} \exp\left(-\frac{y^2}{2\sigma^2}\right), \text{ for } y \geq 0,$$

and Rician for $X = A$, given by

$$p(y|X = A) = \frac{y}{\sigma^2} \exp\left(-\frac{y^2 + A^2}{2\sigma^2}\right) I_0\left(\frac{Ay}{\sigma^2}\right), \text{ for } y \geq 0,$$

respectively, where $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind.

In addition, if a deterministic threshold $\gamma \in (0, \infty)$ is set to perform hard decision based on $Y$, we denote the hard output by $Z$ where

$$Z = \begin{cases} 1 & \text{if } Y \geq \gamma, \\ 0 & \text{otherwise}. \end{cases}$$

Compared with [14] and the references therein where symmetric thresholds are taken on both real and imaginary parts of the channel output, the one-dimension threshold $\gamma$ in our paper is allowed to be asymmetric. According to [15], probabilities of error of hard decisions can be expressed as

$$P(Z = 1|X = 0, \gamma) = \exp\left(-\frac{\gamma^2}{2\sigma^2}\right),$$

$$P(Z = 0|X = A, \gamma) = 1 - Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right),$$

where $Q_1(\cdot, \cdot)$ is the Marcum Q-function of order one.

2.2. Soft-output and hard-output channels

We are constrained by using only on–off symbols without any restrictions on the amplitude $A$, hence the possible

\[
\begin{align*}
\text{Figure 1.} \quad & \quad \text{Discrete-time transmission model in a noncoherent channel setting.} \\
\end{align*}
\]
channel input sets are binary input sets \( \{0, A\} \) for all \( A \in (0, +\infty) \). We can now investigate two channel models for noncoherent OOK transmission, a first model with a soft output, and a second one with a hard output.

- **Binary-In Soft-Out (BISO) channel.** By considering the soft decisions at the envelope detector as realizations of the continuous random variable \( Y \), we can model the channel as a discrete-time memoryless soft-output noncoherent channel, where the inputs are on-off symbols. The amplitude of the ‘on’ symbol is parameterized by \( A \). The channel output \( Y \), for any fixed \( A \), is defined by the conditional densities (5) and (6).

- **Binary-In Binary-Out (BIBO) channel.** By taking the threshold-setting \( \gamma \) and the corresponding hard decisions into account, we obtain a discrete-time memoryless binary-input binary-output channel. The input of this channel is the same as that of the BISO channel whilst the hard output \( Z \) is as defined in (7). The amplitude of the ‘on’ symbol is again parameterized by \( A \). The binary output \( Z \) is then conditionally distributed as in (8) and (9). Note that such distributions imply that the BIBO channel is a binary asymmetric channel.

### 2.3. An equivalent transmission model

Based on the previous description, we can conclude that the output of the envelope detector for noncoherent transmission either has a Rayleigh distribution or a Rician distribution corresponding to amplitude \( A \). Note that the channel magnitude is \( |h| = 1 \).

Equivalent behaviour is obtained if we concatenate a real-valued Gaussian channel with a post-processing block that squares the output of this Gaussian channel \( \hat{Y} \), adds the square of a real-valued Gaussian variable \( N_1 \) to this square, and then takes the square-root of the resulting sum (Figure 2). Note that \( \hat{N} \sim \mathcal{N}(0, \sigma^2) \) and \( N_1 \sim \mathcal{N}(0, \sigma^2) \) are real-valued independent noise variables. In this way, the output of the real-valued Gaussian channel

\[
\hat{Y} = X + \hat{N}
\]

is transformed into the output of the equivalent ‘envelope detection’ channel \( Y \), which is given by

\[
X \xrightarrow{\text{Real-valued channel}} Y \xrightarrow{(\cdot)^2} \xrightarrow{\text{Post-processing block}} \sqrt{()} \xrightarrow{\text{'Envelope detection'}} Y
\]

Figure 2. Cascade of a real-valued channel and a post-processing block that is equivalent to a noncoherent channel followed by envelope detection.

\[
Y = \sqrt{(X + \hat{N})^2 + N_1^2}. \tag{11}
\]

According to [16, p. 191], for a fixed \( A \), this output \( Y \) has the same statistical behaviour as the output \( Y \) represented by (4) of the envelope detector in the noncoherent channel setting. Moreover, it is easy to see that \( X \rightarrow \hat{Y} \rightarrow Y \) form a Markov chain. If we now apply the data-processing inequality [17, Thm 4.3.3], we can conclude that the capacity of the noncoherent channel can never exceed that of the corresponding real-valued channel. This conclusion will be used in the capacity per unit-cost analysis for the noncoherent channel setting.

### 3. CAPACITY PER UNIT-COST AND WIDEBAND SLOPES

In this section, the channel capacities of two noncoherent OOK channels under an input energy constraint are defined. We then focus on the capacity behaviour of both channels in the low-SNR regime.

#### 3.1. Capacity-cost functions

We will first give the definition of the capacity-cost function for the BISO channel. Note that we focus on cost \( b[X] = X^2 \) for symbol \( X \). Therefore, our cost function specializes to energy. If now the expected transmitted (and received) energy \( E_s \) is constrained by \( \beta \), we obtain

\[
E_s = E[b(X)] = P_A A^2 \leq \beta, \tag{12}
\]

where \( P_A \) is the probability that the ‘on’ symbol occurs and \( A^2 \) is the cost (energy) of this ‘on’ symbol.

**Definition 1.** The capacity-cost function of a BISO channel denoted as \( C_{2, \infty}(\beta) \) is the supremum, over \( A \), of the capacity-cost function achieved by a noncoherent OOK channel with fixed amplitude \( A \), which is

\[
C_{2, \infty}(\beta) \triangleq \sup_A C_{2, \infty}(\beta, A), \text{ for } A > 0, \tag{13}
\]

where \( C_{2, \infty}(\beta, A) \) is the capacity-cost function of an OOK channel with a fixed \( A \), namely, see [18].

\[
C_{2, \infty}(\beta, A) = \max_{P_A} \{ I(X; Y|A) : E[b(X)] \leq \beta \}. \tag{14}
\]

Therefore, the capacity-cost curve of \( C_{2, \infty}(\beta) \) in (13) can be considered as the envelope of all capacity-cost curves as expressed in (14) for different \( A \).

Similarly, we can define the capacity-cost function of a BIBO channel. Besides the optimization over values of \( P_A \) and \( A \), there is another parameter, that is, threshold \( \gamma \), that needs to be optimized.
Definition 2. The capacity-cost function of a BIBO channel, which is denoted as \( C_{2,2}(\beta) \), is

\[
C_{2,2}(\beta) = \sup_{A, \gamma} C_{2,2}(\beta, A, \gamma), \quad A > 0,
\]

where \( C_{2,2}(\beta, A, \gamma) \) is the capacity-cost function of the OOK channel for fixed \( A \) and \( \gamma \), namely,

\[
C_{2,2}(\beta, A, \gamma) = \max_{P_x} \{ I(X; Z|A, \gamma) : E[b(X)] \leq \beta \}.
\]

3.2. Capacity per unit-cost

To reveal the capacity behaviour in the low-SNR regime, the capacity per unit-cost denoted as \( C_{uc} \) follows that from the capacity-cost function of a channel by

\[
C_{uc} = \sup_{\beta > 0} \frac{C(\beta)}{\beta} \tag{17}
\]

(Verdú [10, Thm. 2]) can be calculated for our channels. If \( \beta \) is defined as the transmitted symbol energy as presented in (12), the \( C_{uc} \) is intimately related with the required minimum transmit energy per bit for reliable communication \( E_{b,\text{min}} \) by

\[
E_{b,\text{min}} = \frac{1}{C_{uc}} \tag{18}
\]

[6, Sec. IV-A]. Note that for phase fading with unit channel magnitude, the transmitted energy per bit is same as the received energy, hence \( E_b = E_b^r \).

For soft-out and hard-out channel models, we now have the following theorems.

Theorem 1. Consider a certain BISO channel with capacity-cost functions \( C_{2,\infty}(\beta, A) \), parameterized by \( A \). Then the capacity per unit-cost of this channel satisfies \( C_{uc,2,\infty} = \sup_A C_{uc,2,\infty}(A) \), where \( C_{uc,2,\infty}(A) \) is the capacity per unit-cost of a channel with a fixed amplitude \( A \). Moreover \( C_{uc,2,\infty}(A) = C_{2,\infty}(0, A) \) where \( C_{2,\infty}(0, A) \) is the first derivative of the capacity cost function \( C_{2,\infty}(\beta, A) \) at \( \beta = 0 \).

Proof. Observe that

\[
C_{uc,2,\infty}(A) = \frac{\sup_{\beta > 0} C_{2,\infty}(\beta)}{\beta} = \sup_{\beta > 0} \frac{C_{2,\infty}(\beta, A)}{\beta} = \sup_A \frac{C_{2,\infty}(\beta, A)}{\beta} = \sup_A C_{uc,2,\infty}(A).
\]

Moreover, because of [10, eqn. 7], we further have

\[
C_{uc,2,\infty}(A) = \dot{C}_{2,\infty}(0, A) \tag{20}
\]

as stated.

Using the approach for obtaining Theorem 1, we can obtain the following result.

Theorem 2. For a certain BISO channel with capacity-cost functions \( C_{2,2}(\beta, A, \gamma) \), parameterized by \( A \) and \( \gamma \), the capacity per unit-cost of this channel satisfies

\[
C_{uc,2,2} = \sup_{A, \gamma} C_{uc,2,2}(A, \gamma),
\]

where \( C_{uc,2,2}(A, \gamma) \) is the capacity per unit-cost of a channel with a fixed amplitude \( A \) and a fixed threshold \( \gamma \). In addition, \( C_{uc,2,2}(A, \gamma) = C_{2,2}(0, A, \gamma) \) where \( C_{2,2}(0, A, \gamma) \) is the first derivative of the capacity cost function \( C_{2,2}(\beta, A, \gamma) \) at \( \beta = 0 \).

In the next Theorem, we zoom in more on the BISO and BIBO channels.

Theorem 3. For fixed \( A \), the capacity per unit-cost of a BISO channel is given by

\[
C_{uc,2,\infty}(A) = \frac{D(p_{Y|X=A}||p_{Y|X=0})}{A^2}, \tag{21}
\]

and for fixed \( A \) and \( \gamma \), the capacity per unit-cost of a BIBO channel is

\[
C_{uc,2,2}(A, \gamma) = \frac{D(P_{Z|X=A, Y}||P_{Z|X=0, Y})}{A^2}. \tag{22}
\]

Here \( D(\cdot||\cdot) \) is the Kullback-Leibler divergence between the two densities or probability distributions [19, eqn. 9.46, eqn. 2.26].

Proof. First consider the BISO channel. Fix \( A \). For small enough \( \beta \), we get \( \beta = P_A A^2 \). Therefore \( C_{2,\infty}(\beta, A) = \dot{I}(P_A)/A^2 \), where \( \dot{I}(P_A) \) is the first derivative of the mutual information \( I(X; Y) \) with respect to \( P_A \) of a BISO channel evaluated at \( P_A \) for a certain fixed \( A \). According to the Leibniz’s rule for differentiating an integral [20, eqn. 1.1], \( \dot{I}(P_A) \) can be calculated as

\[
\frac{\partial I(P_A)}{\partial P_A} = \frac{\partial}{\partial P_A} \int_0^\infty (P_A p(y|X=A) \log_2 \left( \frac{p(y|X=A)}{p(y)} \right) + (1-P_A) p(y|X=0) \log_2 \left( \frac{p(y|X=0)}{p(y)} \right) dy
\]

\[
= \int_0^\infty p(y|X=A) \log_2 \left( \frac{p(y|X=A)}{p(y)} \right) dy - p(y|X=0) \log_2 \left( \frac{p(y|X=0)}{p(y)} \right) dy
\]

(23)

where \( p(y) = P_A p(y|X=A) + (1-P_A) p(y|X=0) \). For \( P_A \rightarrow 0 \), required because \( \beta \rightarrow 0 \), the output density \( p(y) \rightarrow p(y|X=0) \). Hence

\[
\dot{C}_{2,\infty}(0, A) = \frac{1}{A^2} \int p(y|X=A) \log_2 \left( \frac{p(y|X=A)}{p(y|X=0)} \right) dy.
\]

(24)
Similarly, for the BIBO channel specified by fixed $A$ and $\gamma$, we obtain

$$\dot{C}_{2,2}(0,A,\gamma) = \frac{1}{A^2} \sum_z p(z|X = A,\gamma) \log_2 \left( \frac{p(z|X = A,\gamma)}{p(z|X = 0,\gamma)} \right)$$

(25)

in bits per unit-cost.

### 3.3. Wideband slope

In addition to the minimum energy per bit, the wideband slope is also an important measure for evaluation of a signalling method in the low-SNR regime [6]. This slope is defined as the increase in capacity (in bits per second per Hz) observed when the energy per bit increases 3 dB relative to the minimum SNR per bit, $\left( E_b/N_0 \right)_{\text{min}}$. The slope can be interpreted as the growth speed of the capacity at the minimum bit energy. For a certain on–off input with fixed amplitude $A$, it also makes sense to parameterize the wideband slope here. According to [6], the wideband slope can be determined from the first and second derivatives, that is, for a BISO channel,

$$S_0(A) = -\frac{1}{2} \left[ \frac{\dot{C}_{2,\infty}(0,A)}{C_{2,\infty}(0,A)} \right]^2 \quad \text{(bits/s/Hz/3dB),}$$

(26)

and for a BIBO channel,

$$S_0(A,\gamma) = -\frac{1}{2} \left[ \frac{\dot{C}_{2,2}(0,A,\gamma)}{C_{2,2}(0,A,\gamma)} \right]^2 \quad \text{(bits/s/Hz/3 dB).}$$

(27)

Hence, it is also important to calculate the second derivative of channel capacity $\dot{C}(\beta)$ at $\beta = 0$ as a function of $A$ for the BISO channel and as a function of $A$ and $\gamma$ for the BIBO channel. Based on the previous analysis, we obtain the following results.

**Theorem 4.** For fixed $A$, the second derivative of the capacity cost function of a BISO channel evaluated at $\beta = 0$ is

$$\ddot{C}_{2,\infty}(0,A) = \frac{1}{A^4} \ln 2 \left[ 1 - \int_y \frac{[p(y|X = A)]^2}{p(y|X = 0)} \, dy \right],$$

(28)

and for fixed $A$ and $\gamma$, the second derivative of the capacity cost function of a BIBO channel at $\beta = 0$ is

$$\ddot{C}_{2,2}(0,A,\gamma) = \frac{1}{A^4} \ln 2 \left[ 1 - \sum_z \frac{[p(z|X = A,\gamma)]^2}{p(z|X = 0,\gamma)} \right].$$

(29)

**Proof.** First consider the BISO channel. Fix $A$. For small enough $\beta$, it follows that $\beta = P_A A^2$. Therefore $\ddot{C}_{2,\infty}(\beta,2) = \dot{I}(P_A)/A^4$, where $\dot{I}(P_A)$ is the second derivative of the mutual information $I(X;Y)$ of a BISO channel evaluated at $P_A$ for a fixed $A$. Now, based on (23) and applying the Leibniz’s rule, we have

$$\frac{\partial^2 I(P_A)}{\partial A} = \frac{\partial}{\partial A} \int_0^{\infty} p(y|X = A) \log_2 \left( \frac{p(y|X = A)}{p(y)} \right)$$

$$- p(y|X = 0) \log_2 \left( \frac{p(y|X = 0)}{p(y)} \right) \, dy$$

$$= \int_0^{\infty} \frac{[p(y|X = A) - p(y|X = 0)]^2}{\ln 2 \cdot p(y)} \, dy.$$  

(30)

If we observe that $p(y) = p(y|X = 0)$ for $P_A = 0$, we obtain the BISO result (28). In a similar way, we get the BIBO part (29) of the theorem.

**4. ACHIEVING SHANNON LIMITS**

In this section, we evaluate the capacity per unit-cost and wideband slope of the BISO and BIBO channels in noncoherent OOK settings for low SNR.

**4.1. The binary-in soft-out noncoherent channel**

In the following, we explicitly study the achievability of the Shannon limit for the BISO noncoherent OOK channel. Direct proofs give a better insight on using noncoherent OOK with soft channel outputs.

**4.1.1. Capacity per unit-cost result.**

Based on Theorem 3, the following result describes the achievability of the Shannon limit for the BISO channel.

**Corollary 1.** For the BISO noncoherent channel, the capacity per unit-cost is

$$C_{uc,\infty} = \frac{1}{N_0 \ln 2} \quad \text{(bits/Joule).}$$

(31)

It is achieved and only achieved when amplitude value $A$ goes to infinity.

To prove this result, the following steps are taken.

(i) We first show that using any finite $A$ cannot achieve the Shannon limit by applying the real-valued OOK channel in Figure 2 to provide a strict upper bound of the capacity of the noncoherent BISO channel.

(ii) We then show that letting $A$ go to zero cannot approach the Shannon limit by taking the limit of $C_{uc,\infty}(A)$ in (21) as $A \to 0$. 

(iii) We finally demonstrate only $A$ going to infinity can approach the Shannon limit by evaluating a lower bound of $C_{W, 2\infty}(A)$.

To simplify the presentation, the information is measured in nats in the proof and in the rest of the paper.

Proof.

(i) By considering the transmission model in Figure 2, we first prove that $I(X; \tilde{Y}|A)$ is strictly larger than $I(X; \tilde{Y}^2|A)$ in what follows. Because $I(X; \tilde{Y}^2) = I(X; |\tilde{Y}|)$ for any fixed $A$, we will equivalently demonstrate the inequality

$$H(X|\tilde{Y}) = \int_0^\infty p(\tilde{y}) \mathcal{H}(P(X = 0|\tilde{y})) d(\tilde{y})$$

$$> H(X; \tilde{Y}) = \int_0^\infty [p(-\tilde{y}) \mathcal{H}(P(X = 0) - \tilde{y}) + \mathcal{H}(P(X = 0|\tilde{y})) d(\tilde{y})$$

where $\mathcal{H}(\cdot)$ denotes the binary entropy function [19, eqn. 2.5]. To prove (32), we can focus on a fixed $|\tilde{y}|$ that corresponds to the pair $(\tilde{y}, \tilde{y})$. Due to the convexity of $\mathcal{H}(\cdot)$, we have

$$\frac{p(-\tilde{y})}{p(|\tilde{y}|)} \mathcal{H}(P(X = 0) - \tilde{y}) + \frac{p(\tilde{y})}{p(|\tilde{y}|)} \mathcal{H}(P(X = 0|\tilde{y}))$$

$$< \mathcal{H}(\frac{p(-\tilde{y}) P(X = 0) - \tilde{y}) + p(\tilde{y}) P(X = 0|\tilde{y}))}{p(\tilde{y})}$$

$$= \mathcal{H}(\frac{p(P(X = 0, -\tilde{y}) + p(0, \tilde{y}))}{p(\tilde{y})})$$

$$= \mathcal{H}(p(P(X = 0|\tilde{y}))$$

where $p(\tilde{y}) = p(-\tilde{y}) + p(\tilde{y})$. Because of the asymmetry of the on–off signal's constellation, it is easy to show that $P(X = 0) - \tilde{y} \neq P(X = 0|\tilde{y})$ for $\tilde{y} > 0$. Now, strict inequality in the second step of (33) follows. By integrating over all $|\tilde{y}|$, we obtain (32). Consequently, by the data-processing theorem, we have that

$$C_{W, 2\infty}(\beta, A) = I(X; Y|A) \leq I(X; \tilde{Y}^2|A) < I(X; \tilde{Y}|A).$$

(34)

Thus, we can conclude now that the capacity of the real-valued channel with on–off keying is always strictly larger than $C_{W, 2\infty}(\beta, A)$ for any $\beta > 0$ and finite $A$. This implies that $C_{W, 2\infty}(A)$ is strictly less than the capacity per unit-cost of the real-valued channel, which is equal to $1/(N_0 \ln 2)$ regardless of $A$ [6, Sec. IV. B]. Now, we consider the cases where $A$ goes to zero and infinity, respectively.

(ii) Based on Theorem 3, we can evaluate (21) at $A \to 0$ by applying L'Hôpital's rule and the Leibniz's rule as

$$C_{W, 2\infty}(\beta, A)_{A \to 0} = \frac{D(P_Y|X = A)^2}{\mathcal{H}(P_Y|X = 0)}$$

$$= \lim_{A \to 0} \frac{1}{2A} \frac{\partial}{\partial A} \int_0^\infty p(y|X = A) \mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)}) dy$$

$$= \lim_{A \to 0} \frac{1}{2A} \frac{\partial}{\partial A} \int_0^\infty \mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)}) dy$$

$$+ \frac{1}{2A} \int_0^\infty \mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)}) dy$$

(35)

where $p(y|X = A)$ and $p(y|X = 0)$ are the densities (5) and (6), respectively, and $\mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)})$ is the first derivative of (5) with respect to $A$. By further evaluating $\mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)})$, we can see that

$$\mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)}) = \frac{1}{\sigma^2} q(y) - q(y|X = A)),$$

(36)

where

$$q(y) = \frac{y^2}{\sigma^2} \exp \left(-\frac{y^2 + A^2}{2\sigma^2}\right) I_1 \left(\frac{Ay}{\sigma^2}\right),$$

(37)

where $I_1(\cdot)$ is the first-order modified Bessel function of the first kind. By normalizing $Y$ and $A$ to $\sigma$, function $q(y)$ becomes the density function of the non-central chi distribution with four degrees of freedom [16]. Meanwhile, $p(y|X = A)$ and $p(y|X = 0)$ become Rician and Rayleigh densities with $\sigma^2 = 1$, respectively. Now, it can be immediately seen that the second integral of (35) is zero as $A \to 0$. By writing out $\mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)})$, the first term of (35) can be further upperbounded as

$$\lim_{A \to 0} \frac{1}{2A} \int_0^\infty \mathcal{H}(\frac{p(y|X = A)}{p(y|X = 0)}) dy$$

$$= \lim_{A \to 0} \frac{1}{2A^2} \int_0^\infty (q(y) - q(y|X = A)) \mathcal{H}(I_1(Ay)) dy$$

$$\leq \frac{A^2}{2A^2} \int_0^\infty (q(y) - q(y|X = 0)) dy$$

$$= \frac{A}{\sigma^2} \int_0^\infty y q(y) - q(y|X = 0) dy$$

$$= 0,$$

(38)

where the inequality follows from $I_1(x) \leq \exp(x)$. The last step follows from the fact that the difference of two expected values of noncentral chi and Rician distributions are bounded [21]. Because the capacity per unit-cost is nonnegative, we thus have $C_{W, 2\infty}(\beta, A)_{A \to 0} = 0$, which shows that the Shannon limit cannot be achieved by letting $A$ go to zero.
(iii) Finally, we investigate the case where $A$ goes to infinity. Because directly evaluating $C_{\text{uc},2,\infty}(\beta,A)|_{A\to\infty}$ is nontrivial, a lower bound of $C_{\text{uc},2,\infty}(\beta,A)$ is applied as follows. We first have

$$D \frac{(P_{y|X=A})}{A^2} [P_{y|X=0}] = \int_0^\infty p(y|X=A) \ln \left( \frac{A}{\sigma^2} \right) dy$$

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \frac{A^2}{2\sigma^2} + \ln \left( \frac{A}{\sigma^2} \right) dy$$

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{A}{\sigma^2} \right) dy - \frac{1}{2\sigma^2}.$$  

(39)

Next, by using the inequalities [22],

$$I_0(Ay/\sigma^2) > \frac{(x_0(1))}{e} exp(\frac{Ay/\sigma^2}{3})$$

for $0 < Ay/\sigma^2 < 1$, for $Ay/\sigma^2 \geq 1$,

(40)

we can then write

$$\int_0^\infty p(y|X=A) \ln \left( \frac{A}{\sigma^2} \right) dy$$

$$\geq \int_0^\infty p(y|X=A) \ln \left( \frac{(h_0(1))}{e} + \frac{Ay}{\sigma^2} \right) dy$$

$$+ \int_0^\infty p(y|X=A) \ln \left( \frac{(h_0(1))}{e} + \frac{Ay}{\sigma^2} - \ln \frac{A}{\sigma^2} \right) dy$$

$$> \ln \frac{(h_0(1))}{e} + E_A \frac{Ay}{\sigma^2} - \int_0^\infty p(y|X=A) \frac{A}{\sigma^2} + \ln \left( \frac{A}{\sigma^2} \right) dy$$

$$> \ln \frac{(h_0(1))}{e} + E_A \frac{Ay}{\sigma^2} - \int_0^\infty p(y|X=A) \frac{A}{\sigma^2} + y dy$$

(41)

where the second step follows from $\ln(Ay/\sigma^2) = \ln(A/\sigma^2) + \ln(y) < A/\sigma^2 + y$ due to the fact $\ln(x) < x$, and where we denote expectation with respect to $p(Y|X = A)$ by $E_A[\cdot]$.

By observing that the summation in (42) is positive and goes to zero fast when $x \to \infty$, we can bound $I_0(x)$ in a tighter way as

$$I_0(x) \sim \frac{\exp(x)}{\sqrt{2\pi x}} \left( 1 + \sum_{k=1}^{\infty} \frac{(2j - 1)^2}{k!(8x)^k} \right).$$

(42)

By observing that the summation in (42) is positive and goes to zero fast when $x \to \infty$, we can bound $I_0(x)$ in a tighter way as

$$I_0(x) \sim \frac{\exp(x)}{\sqrt{2\pi x}} \left( 1 + \sum_{k=1}^{\infty} \frac{(2j - 1)^2}{k!(8x)^k} \right).$$

(43)

which holds for any $x \geq \frac{1}{2}$. Now by analysing (39) again based on (43), the following proposition gives a tight lower bound of $C_{\text{uc},2,\infty}(A)$ in closed form when $A$ is large.

**Proposition 1.** For large fixed $A$ and $\sigma^2 = 1$, we can lower bound $C_{\text{uc},2,\infty}(A)$ in nats as

$$C_{\text{uc},2,\infty}(A) > \frac{E_A[Y]}{A} - \frac{\ln(2\pi E_A[Y])}{2A^2} = \frac{1}{2}.$$  

(44)

Note that we normalize the power of the background noise without loss of generality.

**Proof.** For any $y \geq \frac{1}{2\pi}$, we have $I_0(Ay) > \exp(Ay) / \sqrt{2\pi Ay}$. Hence, we can obtain

$$D \frac{(P_{y|X=A})}{A^2} [P_{y|X=0}] = \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy - \frac{1}{2}$$

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy - \frac{1}{2}$$

(45)

$$= \frac{E_A[Y]}{A} - \frac{\ln(2\pi E_A[Y])}{2A^2} = \frac{1}{2}.$$  

$$= \frac{E_A[Y]}{A} - \frac{\ln(2\pi E_A[Y])}{2A^2} = \frac{1}{2}.$$  

$$- \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy - \frac{1}{2}.$$  

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy.$$  

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy.$$  

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy.$$  

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy.$$  

$$= \frac{1}{A^2} \int_0^\infty p(y|X=A) \ln \left( \frac{\exp(Ay)}{\sqrt{2\pi Ay}} \right) dy.$$  

Although it is possible to consider this capacity per unit-cost result as a special case of [6, Thm. 1], where a fast fading channel with complex-valued soft outputs is studied, our proof provides a better understanding of the capacity behaviour of noncoherent OOK parameterized by an amplitude $A$ in the low-SNR regime. Also note that the previous procedure of the proof is same to the one used [9, Thm. 2, Thm. 3], which proves that any input distribution achieving the Shannon limit over real-valued Gaussian channel with hard decisions must be flash signalling.

To lower bound $C_{\text{uc},2,\infty}(A)$ for $\forall A \in (0, \infty)$ in our proof, bounds (40) and $\ln(x) < x$ were used. Although these bounds successfully revealed the capacity behaviour as $A$ goes to infinity, the resulting lower bound (41) is not very tight if it is compared with numerical evaluation of (39). To better demonstrate the behaviour of $C_{\text{uc},2,\infty}(A)$ as a function of $A$, we consider an asymptotic expansion of $I_0(x)$ [24, eqn. 9.7.1]

$$I_0(x) \sim \frac{\exp(x)}{\sqrt{2\pi x}} \left( 1 + \sum_{k=1}^{\infty} \frac{(2j - 1)^2}{k!(8x)^k} \right).$$

(42)
The second derivative \( \frac{d^2}{dA^2} C_{\text{uc},2,\infty}(A) \) is \(-\infty\) for \( A \to \infty \). This means that the BISO noncoherent channel sideband slope \( \lim_{A \to \infty} \tilde{S}_0(A) = 0 \).

Proof. It follows from the proof of \( C_{\text{uc},2,\infty} \), that the Shannon limit is achieved only for \( A \to \infty \). Therefore, based on (28), we have

\[
\lim_{A \to \infty} \tilde{C}_{2,\infty}(0,A) = \lim_{A \to \infty} \frac{1}{A^2} \int_0^1 \exp\left(\frac{A y}{\sqrt{2 \pi A y}}\right) dy \exp\left(-\frac{\exp(A y)}{\sqrt{2 \pi A y}}\right) dy
\]

where the last step follows from Jensen’s inequality because \( I_0(x) \) is convex in \( x \). Now, observing that \( E_A[Y] \to A \) for \( A \to \infty \) and using the lower bound (43) for large \( A \), we get

\[
\lim_{A \to \infty} \tilde{C}_{2,\infty}(0,A) < \lim_{A \to \infty} \frac{-\exp\left(\frac{A y}{\sigma^2}\right)}{A^4 \sqrt{2 \pi A^2 / \sigma^2}} = -\infty.
\]

Now, because the numerator of (26) is finite for \( A \to \infty \), we obtain that \( \lim_{A \to \infty} \tilde{S}_0(A) = 0 \).

Note that [6, Thm. 16], which discusses the wideband slope achieved by OOK with \( A \to \infty \), implies aforementioned result. However, our approach presented in the previous proof provides a relatively simple solution to our signalling scheme compared with the one used in [6].

4.2. The binary-in binary-out noncoherent channel

Next, we concentrate on using hard decisions for OOK. We evaluate the \( C_{\text{uc},2,2} \) and the wideband slope for the BIBO channel to investigate the capacity behaviour in low-SNR regime.

4.2.1. Capacity per unit-cost result.

Corollary 3. For the BIBO noncoherent channel, the capacity per unit-cost also achieves the Shannon limit, namely

\[
C_{\text{uc},2,2} = \frac{1}{N_0 \ln 2} \quad \text{(bits/joule)}.
\]

It is achieved if and only if \( A \to \infty \) and \( \gamma < A \). Only if the threshold \( \gamma \) is chosen such that both \( (A - \gamma) \to \infty \) and \( \gamma / A \to 1 \) the Shannon limit is obtained. A fixed threshold value \( \gamma \) does not achieve the Shannon limit.

The following steps are taken to prove this result.

(i) We first apply the BISO channel to upper bound the capacity of the BIBO channel to show that the...
Shannon limit can only be obtained by letting $A$ be infinity.

(ii) We then derive the expression of $C_{uc,2,2}$ under $A \to \infty$.

(iii) We next demonstrate that the Shannon limit cannot be achieved by any limited threshold $\gamma$.

(iv) We then study the case where the unlimited $\gamma \geq A$ and show that the Shannon limit cannot be achieved in this setting.

(v) We finally study the case where the unlimited $\gamma < A$ to show that the Shannon limit can be achieved only if $A - \gamma$ is going to infinity whilst $\gamma / A$ is going to one are simultaneously satisfied.

Proof.

(i) For any finite and fixed $A$, we can first upper bound the $C_{uc,2,2}(\beta, A)$ by

$$C_{uc,2,2}(\beta, A) = \sup_{\gamma} I(X; Z|A, \gamma) \leq I(X; Y|A),$$

which follows from the Markovity of $X \to Y \to Z$ for any possible $\gamma$. Because $A$ going to infinity is required for the BISO channel if we want to achieve the Shannon limit (Corollary 1), it is indicated that for any finite $A$,

$$C_{uc,2,2}(A) \leq C_{uc,2,\infty}(A) < \frac{1}{N_0 \ln 2}. \tag{52}$$

(ii) Now, we evaluate $C_{uc,2,2}(A, \gamma)$ given by (22) in the case where $A \to \infty$. By applying distributions of (8) and (9), we have that

$$\frac{D(P_{Z|X=A,\gamma}|P_{Z|X=0,\gamma})}{A^2} \bigg|_{A \to \infty} = \lim_{A \to \infty} \frac{1}{A^2} \sum_{z=1,0} P(z|X=A) \ln \frac{P(z|X=A)}{P(z|X=0)}$$

$$= \lim_{A \to \infty} \frac{1}{A^2} \left( Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \ln \frac{Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right)}{\exp\left(-\frac{\gamma^2}{2\sigma^2}\right)} \right.$$  

$$+ \left( 1 - Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \right) \ln \frac{1 - Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right)}{1 - \exp\left(-\frac{\gamma^2}{2\sigma^2}\right)} \right)$$

$$= \lim_{A \to \infty} \left( -\mathcal{H}\left(Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right)\right) + Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \frac{\gamma^2}{2\sigma^2 A^2} \right.$$  

$$- \left( 1 - Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \right) \ln \left( 1 - \exp\left(-\frac{\gamma^2}{2\sigma^2}\right) \right) \bigg)$$

$$= Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \frac{\gamma^2}{2\sigma^2 A^2} \bigg|_{A \to \infty}, \tag{53}$$

where the last step follows from the fact that $1 - \exp\left(-\frac{\gamma^2}{2\sigma^2}\right)$. $Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right)$, and $\mathcal{H}\left(Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right)\right)$ are bounded function regardless of the value of threshold $\gamma$. In what follows, we discuss all possible $\gamma$ setting scenarios.

(iii) By first considering the case where threshold $\gamma$ is bounded, we have that

$$Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \bigg|_{A \to \infty} \to 1, \tag{54}$$

which yields that

$$C_{uc,2,2} = Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \frac{\gamma^2}{2\sigma^2 A^2} \bigg|_{A \to \infty} = 0. \tag{55}$$

This shows that the Shannon limit cannot be achieved. Therefore, the threshold $\gamma$ should not be limited. For unbounded $\gamma$, we can further distinguish between $\gamma \geq A$ and $\gamma < A$.

(iv) For the case where the unbounded $\gamma \geq A$, we can set $\gamma = cA$ where $c \geq 1$ such that the capacity per unit-cost turns out to be

$$C_{uc,2,2} = Q_1\left(\frac{A}{\sigma}, \frac{cA}{\sigma}\right) \frac{c^2}{2\sigma^2 A^2} \bigg|_{A \to \infty}. \tag{56}$$

According to [22, eqn. 5], we can further have

$$Q_1\left(\frac{A}{\sigma}, \frac{cA}{\sigma}\right) c^2 \leq c^2 \exp\left(-\frac{(c-1)^2 A^2}{2\sigma^2}\right) + \frac{\sqrt{2\pi c^2 A}}{\sigma} \mathcal{Q}\left(\frac{(c-1)A}{\sigma}\right)$$

$$\leq c^2 \left( 1 + \frac{\sqrt{2\pi c A}}{2\sigma} \right) \exp\left(-\frac{(c-1)^2 A^2}{2\sigma^2}\right), \tag{57}$$

where $Q(\cdot)$ is the standard $Q$-function, and where the Chernoff bound $Q(\cdot) \leq \exp(-x^2/2)/2$ is used to obtain the second inequality. When $A \to \infty$, (57) vanishes for any $c > 1$ so that

$$Q_1\left(\frac{A}{\sigma}, \frac{cA}{\sigma}\right) c^2 \bigg|_{A \to \infty} < 1. \tag{58}$$

On the other hand, for $c = 1$, it is easy to see that (58) is also valid because $Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) < 1$. Hence, by (56), the Shannon limit cannot be achieved by letting $\gamma \geq A$.

(v) Finally, consider $\gamma < A$ with $A \to \infty$ and $\gamma \to \infty$. Based on (53), we have that

$$C_{uc,2,2} = Q_1\left(\frac{A}{\sigma}, \frac{\gamma}{\sigma}\right) \frac{\gamma^2}{2\sigma^2 A^2} \bigg|_{A,\gamma \to \infty}. \tag{59}$$
with \( y < A \). Therefore, the capacity per unit-cost can be achieved only if

\[
\left\{ \begin{array}{l}
Q_1\left(\frac{A}{\sigma}, \frac{y}{\sigma}\right) \to 1 \\
\gamma/A \to 1.
\end{array} \right.
\]

(60)

Thus, the threshold is required to be chosen in the way such that \((A - y)\) goes to infinity and \(y/A\) goes to one at the same time. To accomplish this, we can take the assignment

\[
\gamma/\sigma = A/\sigma - \sqrt{A/\sigma} \text{ when } A/\sigma \to \infty
\]

(61)

as one way to achieve the Shannon limit. \( \square \)

Based on this result, we can see that the Shannon limit can also be achieved by using hard decisions with noncoherent detection if we just use extreme asymmetric signalling. It is demonstrated that an unbounded threshold is also needed, which is the same as in the case of a real-valued Gaussian channel [9, Corr. 1]. Given these positive conclusions, it is time to also investigate the wideband slope for the BIBO channel for \( A \to \infty \).

4.2.2. Wideband result.

Corollary 4. For \( A \to \infty \), the second derivative \( \tilde{C}_{2,2}(0, A, y) = -\infty \). Hence, also the BIBO noncoherent channel has \( \lim_{A \to \infty} S_0(A) = 0 \).

Proof. From the derivation of \( C_{CC,2,2} \), we see that the Shannon limit is achieved only by simultaneously letting \((A - y) \to \infty\) and \(y/A \to 1\) with \( A \to \infty \) and \( y \to \infty \). So we must have \( y > \frac{A}{2} \); otherwise, the condition (60) is not satisfied. Hence,

\[
\begin{align*}
\lim_{A \to \infty} \tilde{C}_{2,2}(0, A, y) &= \lim_{A \to \infty} -1\left[ \frac{2}{A^4} \left( Q_1\left(\frac{A}{\sigma}, \frac{y}{\sigma}\right) \right)^2 \right. \\
&\quad \left. - \left( 1 - Q_1\left(\frac{A}{\sigma}, \frac{y}{\sigma}\right) \right)^2 \exp\left(-\frac{y^2}{2\sigma^2}\right) \right. \\
&\quad \left. - \exp\left(-\frac{y^2}{2\sigma^2}\right) \right] \\
&= \lim_{A \to \infty} -\frac{\exp\left(-\frac{A^2}{8\sigma^2}\right)}{A^4} \\
&< \lim_{A \to \infty} -\frac{\exp\left(-\frac{A^2}{8\sigma^2}\right)}{A^4} = -\infty,
\end{align*}
\]

(62)

which indicates that \( S_0(A) = 0 \). \( \square \)

4.3. Capacity comparison

Based on the aforementioned results, we can compare capacity-cost behaviours of the BISO channel and the

![Figure 4. Capacity comparison.](image)

BIBO channel in phase fading with that of the Gaussian-in soft-out complex AWGN channel. The capacity-cost function of complex-valued additive Gaussian channel is

\[
C(\beta) = \log(2) \left( 1 + \frac{\beta}{2\sigma^2} \right).
\]

(63)

Figure 4 illustrates the numerically derived capacity-cost functions in the capacity-versus-\( E_b/N_0 \) plane. It is shown that both curves corresponding to the noncoherent OOK channels have zero \( S_0 \) for reaching the Shannon limit \(-1.59 \text{ dB} \), which is consistent with previous analysis. By comparing these findings with the result for the AWGN channel, we can see that a lot of extra energy per bit has to be invested to increase the capacity for the investigated noncoherent OOK channels. In addition, the figure indicates that we can get roughly 1~dB of \( E_b/N_0 \) improvement if we use soft instead of hard decision for the same capacity.

5. NONCOHERENT PULSE-POSITION MODULATION WITH HARD DECISION

As presented, a PPM scheme can be considered as a special and natural realization of OOK with an asymmetric on-off input distribution. It has also been shown that noncoherent PPM is a promising technique for achieving ULP wireless design in practice [25, 26]. In a PPM scheme, every group of \( k \) information bits is presented by a pulse symbol with amplitude \( A \) sent at one of \( 2^k \) possible positions of \( 2^k \) OOK channel uses. As mentioned in the introduction, in the coherent channel setting, the Shannon limit can be achieved by PPM with hard decoding as was demonstrated by Golay [7] and Wozencraft and Jacobs [8]. The same result was also provided by [9]. In the following, we focus on hard decoding of noncoherent PPM in phase fading and show it can also achieve the Shannon limit.
We assume that there is a threshold $\gamma$. The receiver decodes that there is a pulse sent at $l$th position if it corresponds to the unique received sample that exceeds the threshold at position $l$. Therefore, an upper-bound for error probability $P_e$ can be expressed as

\[
P_e \leq \int_0^\gamma \frac{y}{\sigma^2} \exp\left(-\frac{y^2 + A^2}{2\sigma^2}\right) \frac{A}{\sigma} I_0 \left(\frac{A y}{\sigma^2}\right) dy
\]

\[+ (2^k - 1) \int_\gamma^\infty \frac{y}{\sigma^2} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy, \quad (64)
\]

\[= 1 - Q_1 \left(\frac{A}{\sigma} \frac{y}{\sigma}\right) + (2^k - 1) \exp\left(-\frac{y^2}{2\sigma^2}\right). \]

Because the channel gain is unity, by considering $\sigma^2 = N_0/2$, we have that $A^2/\sigma^2 = 2kE_b/N_0$, which gives

\[
P_e \leq 1 - Q_1 \left(\sqrt{\frac{2kE_b}{N_0}} \sqrt{\frac{y^2}{N_0}}\right) + (2^k - 1) \exp\left(-\frac{y^2}{N_0}\right). \]

To show that for hard decision noncoherent PPM achieves the Shannon limit, we make the assignments

\[
E_b/N_0 = (1 + 2\epsilon)^2 \ln 2, \quad (65)
\]

\[
y^2/N_0 = k(1 + \epsilon)^2 \ln 2, \quad (66)
\]

where $0 < \epsilon < 1$, that is, we are willing to spend slightly more than $N_0 \ln 2$ Joule per bit and make the threshold $\gamma$ less than the amplitude $A$ which is given by $\sqrt{2kE_b}$ for the $2^k$-position PPM. Now, the upper bound on the symbol error probability becomes

\[
P_e \leq 1 - Q_1 (1 + 2\epsilon) \sqrt{2k \ln 2}, (1 + \epsilon) \sqrt{2k \ln 2})
\]

\[+ (2^k - 1) \exp(-k(1 + \epsilon)^2 \ln 2). \quad (67)
\]

According to [27], a lower bound on the Marcum-Q function of the Chernoff-type is

\[
Q_1(s, t) \geq 1 - \frac{s}{s - t} \exp\left[-\frac{(t - s)^2}{2}\right], \quad \text{for} \quad t < s. \quad (68)
\]

By using (68), the combination of the first two terms in (67) is bounded as

\[
1 \leq Q_1 ((1 + 2\epsilon) \sqrt{2k \ln 2}, (1 + \epsilon) \sqrt{2k \ln 2})
\]

\[\leq 1 + 2\epsilon \exp(-ke^2 \ln 2), \quad (69)
\]

and the third term in (67) is bounded using $2^k - 1 \leq 2^k = e^{k \ln 2}$ as

\[
(2^k - 1) \exp(-k(1 + \epsilon)^2 \ln 2) \leq \exp(-k(2\epsilon + \epsilon^2) \ln 2). \quad (70)
\]

Combining (69) and (70), we come to the conclusion that for all small but positive $\epsilon$, the symbol-error probability $P_e$ can be made arbitrary small by increasing $k$. Therefore, it is demonstrated that the Shannon limit can be achieved by using noncoherent PPM with hard decisions in the phase fading channel.

### 6. CAPACITY RESULTS FOR SLOW FADEING

Because ULP wireless systems normally operate in a low-mobility and low-transmission-rate regime, it is reasonable to model the fading as slow and flat, where the fading magnitude is random but remains constant, for example, over a block of transmissions longer than a codeword [28]. For this quasi-static scenario, we investigate the effective minimum bit energy required for reliable communication for noncoherent OOK. The fading magnitude is considered to be constant and unknown to the transceiver whilst on the other hand, the channel phase is still assumed to be time-varying over all on–off symbols. This corresponds to the practical scenario in which the receiver is not able to track and compensate channel imperfections, because typically the power budget is critically limited in ULP design.

Specifically, we consider a transmission model in which the fading magnitude $|h| = |\mathbf{h}|$ is randomly generated for each block and kept fixed for all on–off symbols $X_i$. The nonzero channel input symbol has an amplitude $|\mathbf{h}|A$ for each transmission such that the soft output of the envelope detector is still Rician, which is parameterized by $|\mathbf{h}|A$ instead of $A$, as before. Now, we can express the actual channel coefficient as $h_i = |\mathbf{h}|e^{j\theta_i}$, and the corresponding transmit power gain is equal to $G = h_ih_i^* = |\mathbf{h}|^2$ for a certain fading realization. Without loss of generality, we normalize the power gain by assuming $E[G] = 1$.

In the following, we will discuss the throughput per unit-cost behaviour of noncoherent OOK schemes in such a slow Rayleigh fading channel. The throughput describes the effective data rate of reliable transmissions in a slow fading channel [29]. The throughput per unit-cost thus gives the effective minimum required bit energy.

#### 6.1. Noncoherent on–off keying with soft decisions

Because it is assumed that transmitter has no access to the channel coefficient, no codes can be constructed that always achieve the channel capacity corresponding to $|\mathbf{h}|$, in a slow fading regime. Hence, the only thing we could do is to generate a capacity-achieving code by targeting at a specific fading magnitude $\mathbf{h}$. The following result first guarantees that we can achieve reliable communication using such code if $|\mathbf{h}| \geq \mathbf{h}$.

**Lemma 1.** The mutual information of the noncoherent BISO channel $I(\mathbf{X}; Y | \mathbf{h}A)$ is monotonically increasing in $|\mathbf{h}|$ for any fixed $A$ and input distribution $\{1 - P_A, P_A\}$. 

Proof. Set \( \sigma^2 = 1 \) without loss of generality. So, for a channel \( |h| \), the Rician distribution of the soft output \( Y \) of the envelope detector becomes

\[
p(y|X = |h|A) = y \exp\left(-\frac{y^2 + |h|^2 A^2}{2}\right) I_0(|h|Ay) \quad \text{for } y \geq 0
\]

(71)
corresponding to sending an ‘on’ symbol with amplitude \( A \). Meanwhile, note that the Rayleigh distribution \( p(y|X = 0) \) corresponding to sending an ‘off’ symbol is independent of \( |h| \).

Now, for any fixed \( A \) and input distribution \( P_A \in (0,1) \), the soft output \( Y \) is distributed as \( p(y) = P_A p(y|X = |h|A) + (1 - P_A) p(y|X = 0) \), and the mutual information \( I(X;Y) \) becomes a function in \( |h| \), which is denoted by \( I(|h|) \) for simplicity. By taking the first derivative of \( I(|h|) \) with respect to \( |h| \), we obtain

\[
\frac{dI(|h|)}{d|h|} = P_A \int_0^\infty \frac{\hat{p}(y|X = |h|A) \ln \frac{p(y|X = |h|A)}{p(y)}}{p(y)} dy
\]

\[
= |h|^2 A^2 P_A \int_0^\infty \frac{(y) - p(y|X = |h|A)) \ln \frac{p(y|X = |h|A)}{p(y)}}{p(y)} dy,
\]

(72)
where \( \hat{p}(y|X = |h|A) \) is the first derivative of (71) with respect to \( |h| \), and

\[
q(y) = \frac{y^2}{|h|A} \exp\left(-\frac{y^2 + |h|^2 A^2}{2}\right) I_1(|h|Ay), \quad \text{for } y \geq 0
\]

(73)
is the non-central chi density as seen in the proof of Corollary 1. In what follows, we will prove that for any fixed \( |h| \), we have \( \frac{dI(|h|)}{d|h|} > 0 \).

First, we evaluate the ratio of the two densities \( q(y) \) and \( p(y|X = |h|A) \), that is,

\[
r(y) \triangleq \frac{q(y)}{p(y|X = |h|A)} = \frac{y}{A|h|} I_1(|h|Ay) = \frac{B_1(x)}{B_2 I_0(x)},
\]

(74)
where we first denote the constant \( A|h| \) by \( B \) and next denote \( B_1(x) \) by \( x \), for simplicity. Observing that

\[
\frac{d}{dx} \left( \frac{x I_1(x)}{I_0(x)} \right) = \frac{x [I_0(x)]^2 - x [I_1(x)]^2}{[I_0(x)]^2} \geq 0,
\]

(75)
where according to \([24]\), \( d[x I_1(x)]/dx = x I_0(x) \) and \( I_0(x) \geq I_1(x) \) are applied, it can be concluded that \( x I_1(x)/I_0(x) \) is a monotonically increasing function in \( x \) and so does \( r(y) \) in \( y \).

Next, observe that \( q(y) \) and \( p(y|X = |h|A) \) must intersect because both are non-zero density functions defined for \( y > 0 \). Because of the monotonicity of \( r(y) \) in \( y \), there exists a \( y^* \) such that

\[
q(y) = p(y|X = |h|A) \quad \text{for} \quad y \in [y^*, \infty) \quad \text{(76)}
\]

and

\[
p(y|X = |h|A) > q(y) \quad \text{for} \quad y \in [0, y^*). \quad \text{(77)}
\]

Because \( q(y) \) and \( p(y|X = |h|A) \) are densities, relations (76) and (77) further give that

\[
\int_0^{y^*} [p(y|X = |h|A) - q(y)] dy = \int_0^{y^*} q(y) - p(y|X = |h|A) dy \quad \text{(78)}
\]

Now, denoting \( \ln(p(y|X = |h|A)/p(y)) \) by \( f(y) \), we can express the integral in (72) as

\[
\int_0^{y^*} (q(y) - p(y|X = |h|A)) f(y) dy = \int_0^{y^*} (q(y) - p(y|X = |h|A)) f(y) dy - \int_0^{y^*} (p(y|X = |h|A) - q(y)) f(y) dy \quad \text{(80)}
\]

By taking into account that \( f(y) \) is also monotonically increasing in \( y \), relation (78) implies that integral (79) is always larger than integral (80). Thus, the derivative (72) is positive and the lemma follows.

Lemma 1 indicates that the channel capacity of the non-coherent BISO channel parameterized by \( A \) increases with \( |h| \) where \( |h|A \) can be regarded as the effective ‘on’ symbol of the channel output. Therefore, we can always achieve reliable decoding if only \( |h| \geq h \). Note that a decoding procedure as maximum mutual information decoding \([30]\) could be used, because the receiver may not know the channel coefficient \( |h| \). On the other hand, when the chosen \( h \) is larger than the actual channel coefficient \( |h| \), the constructed code cannot be successfully decoded. In that case, we may say that the system is in outage, and the outage probability is given by

\[
P_{out} \triangleq \Pr(|h| < h). \quad \text{(81)}
\]

If furthermore we assumed that \( |h| \) is Rayleigh distributed with unit variance, \( P_{out} \) can be expressed as

\[
P_{out} = 1 - \exp(-h^2). \quad \text{(82)}
\]

Now, we can consider the throughput of our fading channel, and we have the following result.
Proposition 2. For the BISO noncoherent channel in slow fading, the throughput per unit-cost is
\[
C_{uc,\infty}^* = \frac{h^2 \exp(-h^2)}{N_0 \ln 2} \text{ (bits/Joule)},
\]
where \( h \in (0, \infty) \) and the maximum is achieved at \( h = 1 \).

Proof. For the case where \( |h| \geq h \), that is, where the system is not in outage, the achievable \( C_{uc,\infty}^* \) is \( h^2/(N_0 \ln 2) \) according to Corollary 1 because the equivalent amplitude of the ‘on’ symbol is \( hA \). By observing that \( \Pr(|h| \geq h) = \exp(-h^2) \), the throughput per unit-cost can be seen to be (83). Moreover, optimizing \( C_{uc,\infty}^* \) over \( h \) gives a maximum at \( h = 1 \).

Because reliable communication cannot always be realized in slow fading, the capacity per unit-cost in throughput sense depends on target channel coefficient \( h \). The largest \( C_{uc,\infty}^* \) is 1/(\( \sigma^2 \ln 4 \)), with an outage probability equal to \( 1 - e^{-1} \), which is achieved if codes were designed based on \( h = 1 \). When \( \sigma^2 = 1 \), it is seen that 1/(\( \sigma^2 \ln 4 \)) \approx 0.265 bit/Joule can be realized, which is smaller than the result for only the phase-fading OOK channel with soft decisions, that is, 1/(\( \ln 4 \)) \approx 0.721 bit/Joule.

6.2. Noncoherent on–off keying with hard decisions

Similar to what we have carried out for the BISO channel, we can generate codes for transmission at target fading magnitude \( h \), when hard decisions are made by the receiver. In what follows, we first observe that the capacity of the noncoherent BIBO channel increases with respect to \( |h| \) when we fix the channel input \( A \) and the threshold \( \gamma \).

Lemma 2. The mutual information of the noncoherent BIBO channel \( I(X;Y|A,\gamma) \) is monotonically increasing in \( |h| \) for any fixed \( A \) and input distribution \( \{1-P_A, PA\} \) and threshold \( \gamma \) at the receiver.

Proof. Note that the noncoherent BIBO channel is a binary asymmetric channel. With increasing \( |h| \) and fixed \( \gamma \), the transition probability \( P_1 := P(Z = 1|X = A, \gamma) \) increases whilst the transition probability \( P_0 := P(Z = 1|X = 0, \gamma) \) remains the same. Thus, the channel approaches a Z-channel when \( |h| \) increases. We can now take the first derivative of \( I(X;Y|A,\gamma) \) with respect to \( P_1 \), which is
\[
\frac{\partial I(X;Y|A,\gamma)}{\partial P_1} = \frac{\partial}{\partial P_1} \left[ \mathcal{H}((1-P_A)P_0+PA P_1) \right. \\
- \left. (1-P_A)\mathcal{H}(P_0)-P_A\mathcal{H}(P_1) \right] \\
= PA \ln \left( \frac{1-(1-P_A)P_0-PA P_1}{(1-P_A)P_0+PA P_1} \cdot \frac{P_1}{1-P_1} \right).
\]

In this way, it can be seen that \( \partial I/\partial P_1 > 0 \) if \( P_1 > P_0 \), which is always true if \( |h|A > 0 \), so the lemma follows. □

Therefore, just like for the BISO channel, if \( |h| \geq h \), we can successfully decode a code designed for \( h \) for the BIBO channel. Now, again, we can consider the throughput for such BIBO channel and obtain the following result.

Proposition 3. For the BIBO noncoherent channel in slow amplitude fading, the throughput per unit-cost \( C_{uc,\infty}^* \) is the same as \( C_{uc,\infty}^* \) and also achieves its maximization for target coefficient \( h = 1 \).

Proof. For any certain unknown \( |h| \geq h \), we can design codes for channel magnitude \( h \) to achieve
\[
C_{uc,\infty}^*(h) = Q_1 \left( \frac{hA}{\sigma}, \frac{h\gamma}{\sigma} \right) \frac{h^2 \gamma^2}{\sigma^2 A^2 \ln 4} \left|_{A \to \infty} \right. = \frac{h^2}{N_0 \ln 2},
\]
which follows the derivation of (56). Note that the threshold \( \gamma \) should satisfy property (61). By considering the outage probability \( P_{out} \), the result for \( C_{uc,\infty}^* \) follows. Moreover, optimizing over \( h \) shows that \( h = 1 \) achieves the highest throughput capacity per unit-cost. □

This result shows that using hard decisions for a slow fading channel will not destroy the throughput per unit-cost compared with the case of using soft decisions. The same capacity 0.265 bit/Joule can be achieved again.

6.3. Comparison with fast fading

We have seen that in the noncoherent channel setting with slow fading, both the soft decision and hard decision schemes have the same loss in capacity per unit-cost. It is also interesting to compare these results with the corresponding results in the case of fast fading. The capacity cannot be destroyed in fast fading at low SNR if soft decisions are used [6]. So the capacity per unit-cost 1/(\( N_0 \ln 2 \)) is achieved. In the fast Rayleigh fading case with hard decisions, it is shown that capacity per unit-cost is \( 1/e \) times less than 1/(\( N_0 \ln 2 \)) as was demonstrated by Koch and Lapidoth [9]. The capacity drop is caused by using hard decisions. So in fast fading, using 1-bit quantization should be avoided.

On the contrary, in slow fading, achievable rate loss is inevitable regardless whether soft or hard decisions are made. These results are based on the scenario where the receiver does not have access to the channel realization, which results in a throughput scenario. For this setting, it should be noticed that using hard decisions will not introduce a rate degradation at low SNR, which is remarkable for ULP design.
7. CONCLUSION AND DISCUSSION

In this paper, we investigated noncoherent OOK channels in the low-SNR regime to reveal the required minimum SNR per bit for reliable communication. By modelling the BISO channel and the BIBO channel corresponding to using soft and hard decisions, we explicitly studied the capacity per unit-cost for both channels in phase fading settings. One important result that we found is that noncoherent OOK with threshold detection (hard decisions) can also achieve the Shannon limit when using soft decisions, which is $\ln 2$. This implies that quantization does not affect the capacity in the low-SNR regime for noncoherent channels. This result complements a similar result by Koch and Lapidoth [9], which shows that binary quantization does not decrease the capacity for the coherent case.

Because we assumed that the channel phase is i.i.d and uniformly distributed such no information can be conveyed by the phase, the derived Shannon limit achieving result can be extended to other phase models, such as the Tikhonov distribution and Wiener process as pointed out by one of the reviewers, where the carrier phase partially conveys by the phase, the derived Shannon limit achieving result can be extended to other phase models, such as the Tikhonov distribution and Wiener process as pointed out by one of the reviewers, where the carrier phase partially conveys information about the transmitted signal.

We have also demonstrated that it is necessary to let the amplitude of the ‘on’ symbols go to infinity, that is, to use flash signalling [6], and to use an unbounded threshold. In addition, we have also studied the wideband slope for the two OOK channels, which demonstrated the poor spectral efficiency of noncoherent OOK.

Based on the results of the noncoherent BIBO channel, it was shown that noncoherent PPM with hard decisions achieves the Shannon limit if the threshold is chosen properly. This result complements the work of Golay [7] for coherent PPM schemes with a threshold.

In the end, we considered noncoherent OOK in a slow fading scenario where the channel magnitude is unknown to the transceiver but fixed across the transmission of a codeword. For Rayleigh slow fading, we studied the throughput per unit-cost. We calculated the throughput loss relative to the Shannon limit. The rate degradation was identical for both the BISO and BIBO settings.

ACKNOWLEDGEMENTS

The authors wish to thank T. Koch for discussions on an early version of this work and the editor-in-chief and anonymous reviewers for their valuable comments.

REFERENCES


