

Low-SNR Asymptotics of Cross-Receiver Mutual Information in Gaussian Channels

Adam R. Williamson Applied Research Laboratories, The University of Texas at Austin

Abstract

In this paper, we analyze the low-SNR behavior of the *cross-receiver mutual information* (CMI) between two received signals corrupted by uncorrelated, additive Gaussian noise. This framework has use in distributed, passive sensor applications, such as passive radar and collaborative opportunistic navigation. For Gaussian and BPSK signaling, the CMI can be expressed in terms of the effective input SNR between the receivers. On-off keying (OOK), while not optimal in terms of spectral efficiency for a single-receiver channel, is shown to have greater CMI than Gaussian or BPSK signaling. This is in spite of the fact that, given the same received SNRs, all three source distributions have the same linear correlation coefficient. This indicates that for OOK sources, effective SNR and correlation coefficient are not meaningful descriptors for passive receivers.

A full version of this paper is accessible at: <https://doi.org/10.36227/techrxiv.13710142>

I. INTRODUCTION

The input-output mutual information $I(X; Y)$ between transmitted signal X and received signal Y has been well-studied since Shannon [1] introduced the notion of channel capacity C , the maximum rate of reliable communication subject to an average source power constraint, \mathcal{P} . In particular, in the wideband, power-limited regime, the *capacity-cost function* $C(\mathcal{P})$ can be approximated as a second-order polynomial [2, (141)]:

$$C(\mathcal{P}) = \dot{C}(0)\mathcal{P} + \frac{1}{2}\ddot{C}(0)\mathcal{P}^2 + o(\mathcal{P}^2). \quad (1)$$

The first-order coefficient $\dot{C}(0)$ is referred to as the *channel capacity per unit-energy* [3] and represents the slope of the capacity vs. source power curve at zero power. Together, the polynomial coefficients $\dot{C}(0)$ and $\ddot{C}(0)$ define the *wideband slope* S_0 of the spectral efficiency versus $\frac{E_b}{N_0}$ curve, given by [2, Thm. 9]:

$$S_0 = \frac{2 \left[\dot{C}(0) \right]^2}{-\ddot{C}(0)}. \quad (2)$$

These parameters of the communications channel $X-Y$ have been used to analyze signaling distributions that are optimal in terms of $\dot{C}(0)$ (first-order optimality) and in terms of $\ddot{C}(0)$ (second-order optimality). For example, on the real additive white Gaussian noise (AWGN) channel, it has been shown that Gaussian signaling, binary phase-shift keying (BPSK), and on-off keying (OOK) all achieve the channel capacity per unit-energy $\dot{C}(0)$. However, of these source distributions, only Gaussian signaling and BPSK are second-order optimal [2].

In contrast, less attention has been paid to the *cross-receiver mutual information* (CMI)¹ $I(Y_1; Y_2)$, between noisy received signals Y_1 and Y_2 corresponding to the same transmitted signal, as shown in Fig. 1. While previous work (e.g., [9]–[11]) has used CMI for differential time-delay estimation, there has not been any research into the low-SNR behavior of CMI. This paper derives the first- and second-order Taylor series coefficients for the AWGN channel CMI at low SNR, for several source distributions.

Mutual information captures arbitrary statistical relationships, making it a more general measure of dependence than Pearson's correlation coefficient, which measures the linear dependence between two random variables. Further, while signal processing operations such as quantization may increase the cross-correlation between two received signals (e.g., see [12], [13]), no secondary processing of Y_1 and Y_2 can improve the CMI $I(Y_1; Y_2)$, due to the Data Processing Inequality. This suggests that CMI is more fundamental than the cross-correlation. Intuitively, we expect that given the same levels of additive noise, source distributions with greater CMI would perform better in passive receiver processing tasks, such as TDOA or FDOA estimation.

This paper assumes a particular source distribution (e.g., OOK) subject to average power constraint \mathcal{P} and characterizes the CMI $I(Y_1; Y_2)$ as a function of \mathcal{P} . The low-SNR asymptotics are obtained by letting \mathcal{P} tend to zero, with receiver noise powers fixed. This lens is useful for passive receiver applications, in which the receiver designer often has no choice in the

¹The term cross-receiver mutual information is motivated by the *cross-correlation* between two noisy received signals, an operation commonly performed in passive receiver applications such as radio interferometry [4, Ch. 8.1-8.4], radar, and collaborative opportunistic navigation [5]–[7]. Similarly, joint estimation of time-difference-of-arrival (TDOA) and frequency-difference-of-arrival (FDOA) is commonly performed by computing the cross-ambiguity (or complex ambiguity) function [8].

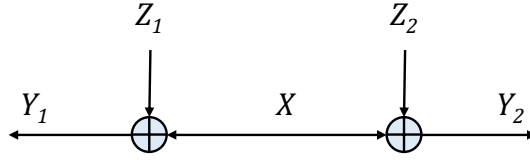


Fig. 1. Signal model showing transmitted signal X , uncorrelated, additive noise Z_i , and received signals Y_i ($i=1, 2$).

design of the transmitted signal.² We evaluate the low-SNR CMI of three source distributions that are known to be first-order optimal for the respective single-receiver AWGN channels $X-Y_i$ ($i=1, 2$): Gaussian signaling, BPSK, and OOK. Notably, the CMI per unit-energy $\dot{I}_{1,2}(0)$, defined precisely in Sec. IV, is shown to be zero for all three distributions. As a result, CMI in the low-SNR regime is quadratic in source power, governed by the second-order CMI coefficient $\ddot{I}_{1,2}(0)$. We show that while OOK is sub-optimal in terms of the single-receiver second-order coefficient, the second-order CMI coefficient of fixed-amplitude OOK surpasses those of Gaussian signaling and BPSK. This is notable, given that for the same received SNRs, all source distributions have the same linear correlation coefficient [13]. This indicates that the dependence between Y_1 and Y_2 is non-linear for OOK.

The remainder of this paper is organized as follows: Sec. I-A discusses additional related literature. Sec. I-B introduces relevant notation and Sec. II describes the additive noise model at each receiver. Sec. III reviews the low-SNR asymptotics of single-receiver channel capacity. Sec. IV introduces the low-SNR asymptotic expansion of CMI, while Sec. V provides the first- and second-order coefficients for specific sources on the AWGN channel: Gaussian signaling (Sec. V-A), BPSK signaling (Sec. V-B), and OOK signaling (V-C, V-D, and V-E). Sec. VI concludes the paper.

A. Related Literature

Cross-correlator-based TDOA estimation was originally introduced in the passive sonar literature, in which the acoustic signal of interest is commonly modeled as a Gaussian random process (e.g., [15]–[17]). While electromagnetic signals of interest tend to be highly structured (e.g., radar or coded communications waveforms), the literature for TDOA estimation of electromagnetic signals often assumes the signal is Gaussian as well. Fowler and Hu [18] showed, however, that differences in the acoustic and electromagnetic signal structures lead to differences in Fisher information, Cramer-Rao bounds, and maximum-likelihood estimators.

Following the work of Bell [19], which optimized radar signal detection by maximizing the mutual information between a Gaussian target ensemble and the received signal, various authors have applied information-theoretic concepts to the analysis and design of radar signals. This work includes the design of matched illumination waveforms [20], [21], analysis of the radar information channel [22]–[24], the constant information radar [25], [26], and co-design of spectrally efficient joint radar-communications waveforms [27]. However, the previous works cited above have analyzed the active sensor case, in which the receiver knows the transmitted signal, rather than the passive sensor case considered here. Stein et al. [28], in contrast, considered the universal delay estimation problem in discrete channels, for both the single-channel (active) case and the two-sensor (passive) case. They showed that for large enough signal duration, the error exponent of the universal delay estimator is equal to that of the maximum-likelihood estimator.

In distributed, passive receiver problems, sensor data is commonly communicated from the points of observation to a centralized estimator. Recognizing that the sensor data is often compressed, Zhang and Berger [29] introduced the notion of estimation from compressed information. In contrast to traditional rate-distortion theory that studies the distortion of a reconstructed source subject to rate-constrained compression, [29] analyzed rate-distortion trade-offs in terms of compression rate vs. estimator variance. A survey of statistical inference under rate-constrained multiterminal data compression is given by Han and Amari [30]. The inference problems surveyed in [30] are categorized as hypothesis testing, parameter estimation (e.g., [31]), or classification.

Vasudevan et al. [9]–[11] introduced the MaxMI method of discrete-time TDOA estimation, which computes the empirical mutual information between received signals for each delay hypothesis and selects the delay with the maximum (cross-receiver) mutual information. Further, to compress the received signals, [9]–[11] selected quantization thresholds that would maximize CMI at the centralized estimator. In numerical examples with Gaussian sources, the MaxMI quantizer-detector in [9]–[11] was shown to have lower error probability than a traditional approach using a minimum mean-squared error (MMSE) quantizer and a correlation detector. In contrast to [9], which gave numerical examples with received SNRs as low as -1 dB, this paper focuses on the low-SNR asymptotics of CMI. We do not, however, evaluate estimator performance for any particular application. Rather, the low-SNR CMI analysis in this paper indicates that when designing multiterminal compression algorithms, particular

²Although the signal model shown in Fig. 1 bears similarity to that of the broadcast channel [14], the analysis in this paper is markedly different from the broadcast channel literature. It is true that the physical signal X is assumed to be transmitted in a broadcast channel, in the sense that multiple geographically separated sensors may receive it. However, this paper does not evaluate achievable rate regions for the two broadcast channel users, nor does it analyze channel coding schemes, since the source distribution is assumed to be fixed from the receiver's perspective.

attention should be paid to the source distribution, as different sources can have extremely different CMI values at the distributed receivers.

While [3] introduced the term *capacity per unit-cost* for general cost functions, the specific energy cost function $b(x)=x^2$ was analyzed earlier by Prelov in [32] and [33], establishing an expression for the capacity per unit-energy two decades prior to [3]. Asymptotic expressions for the capacity of channels with weak input signals were derived for scalar input signals in [34] and for vector-valued input signals in [35], under both average and peak-power constraints. A formula for the second-order expansion of input-output mutual information in Gaussian noise channels as SNR tends to zero is provided in [36]. Building upon the connection between channel capacity per unit-cost and Fisher information observed in [3], a more general result in [37], called the I-MMSE relationship, relates the derivative of mutual information as a function of SNR to the MMSE.

Various extensions to capacity per unit-energy have appeared in the literature. El-Halabi et al. [38], for example, extended the concept of capacity per unit-energy [3] to secrecy capacity per unit-energy. Koch and Lapidot [39] showed that while 1-bit symmetric-threshold quantization degrades the capacity per unit-energy, at low SNR some of the capacity loss can be recovered by doubling the sampling rate. Specifically, symmetric quantization of Gaussian sources and of equiprobable, antipodal binary sources (e.g., BPSK) degrades the capacity per unit-energy by $\frac{2}{\pi}$ (−1.96 dB), whereas [40], [41] showed that an asymmetric source distribution and an asymmetric-threshold quantizer can recover the capacity per unit-energy of the unquantized channel [41, Thm. 2]. Further, they showed that in order for the quantized channel to achieve the capacity per unit-energy of the unquantized channel, the source distribution must be *flash signaling* [41, Thm. 3], a class of binary distributions introduced in [2, Def. 2]. Zhang et al. [42] studied the channel capacity per unit-energy for the single-receiver noncoherent OOK channel, showing that noncoherent OOK is first-order optimal for both soft- and hard-decision decoding.

In [13], Williamson studied the impacts of additive Gaussian noise and 2-level quantization on the correlation coefficient of correlated sources, considering both Gaussian sources and binary sources. In contrast to a classic analysis by Van Vleck [43], [13] considered the effects of noise and quantization separately, showing that as source power tends to zero, symmetric-threshold quantization of Gaussian sources and of equiprobable, antipodal binary sources degrades the correlation by $\frac{2}{\pi}$. However, flash OOK sources and receivers with asymmetric quantization can avoid the $\frac{2}{\pi}$ correlation penalty: in the limit of zero SNR, the quantized signals' correlation approaches the sources' correlation. This has applications to cross-correlator-based TDOA estimation for sparse OOK sources.

Low-duty OOK signaling can be accomplished by PPM, which along with overlapping PPM and multipulse PPM has a long history within the optical communications literature [44]. Additional variants of PPM have been proposed in [45]–[48]. Recently, there has also been renewed interest in OOK for low-power RF communications, due to the reduced receiver hardware requirements for noncoherent OOK, as compared to coherent modulations. For example, OOK is used as the Wake-Up Packet modulation in IEEE 802.11 Wake-Up Radio.

B. Notation

We use capital letters to denote random variables and lowercase letters to denote their realizations; $P_X(x)=P[X=x]$ is the probability that random variable X takes on value $x \in \mathcal{X}$ for alphabet \mathcal{X} . The circumflex notation \hat{Y} denotes the output of a quantizer with input Y . We write $\log x$ to denote the logarithm in an arbitrary base.

II. SIGNAL MODEL

As shown in Fig. 1, we consider the following signal model: source $X \sim P_X$ is corrupted by additive noise at each of two receivers ($i=1, 2$). The received signals are

$$Y_i = X + Z_i, \quad (3)$$

where $Z_i \sim \mathcal{N}(0, \sigma_i^2)$ is zero-mean Gaussian noise with variance $\sigma_i^2 > 0$. The noise terms Z_1 and Z_2 are mutually uncorrelated and uncorrelated with X . The source X is subject to the power constraint

$$E_{P_X}[X^2] \leq \mathcal{P}, \quad (4)$$

corresponding to a maximum SNR of $\gamma_i = \frac{\mathcal{P}}{\sigma_i^2}$ at receiver i .

A common metric in TDOA applications is the *effective input SNR* γ_{eff} , defined as [49, (11)]^{3 4}:

$$\gamma_{\text{eff}} = \frac{\gamma_1 \gamma_2}{1 + \gamma_1 + \gamma_2}. \quad (5)$$

When each of the received SNRs is small ($\gamma_i \ll 1$), the effective input SNR can be approximated as $\gamma_{\text{eff}} \approx \gamma_1 \gamma_2$. When additionally the received SNRs are equal ($\gamma_1 = \gamma_2 = \gamma \ll 1$), the effective input SNR can be approximated as $\gamma_{\text{eff}} \approx \gamma^2$. For example, individual input SNRs of −20 dB correspond to an effective input SNR of approx. −40 dB.

³Note that the definition of effective input SNR in [49] differs by a factor of two compared to the definition in [8].

⁴Estimation-theoretic bounds such as the Cramer-Rao lower bound on TDOA estimator variance are commonly expressed in terms of the effective output SNR, which is equal to the effective input SNR multiplied by the processing gain (i.e., the time-bandwidth product) [49].

Two related quantities are the *equivalent noise* σ_{eq}^2 and the *equivalent SNR* γ_{eq} of the two receivers, defined as

$$\sigma_{\text{eq}}^2 = (\sigma_1^{-2} + \sigma_2^{-2})^{-1}, \quad (6)$$

$$\gamma_{\text{eq}} = \mathcal{P}/\sigma_{\text{eq}}^2 = \gamma_1 + \gamma_2. \quad (7)$$

Note that $\gamma_{\text{eff}} \neq \gamma_{\text{eq}}$. When $\sigma_1^2 = \sigma_2^2 = \sigma$, we have $\sigma_{\text{eq}}^2 = \frac{1}{2}\sigma^2$.

III. SINGLE-RECEIVER ASYMPTOTICS

This section reviews the low-SNR asymptotics of channel capacity for a single receiver. Expressed in terms of the cost constraint \mathcal{P} , the capacity-cost function $C_i(\mathcal{P})$ of the single-receiver channel $X-Y_i$ is [3, Thm. 1]

$$C_i(\mathcal{P}) = \sup_{P_X: E[b(X)] \leq \mathcal{P}} I(X; Y_i), \quad (8)$$

where $b: \mathcal{X} \rightarrow [0, \infty)$ is a non-negative cost function.

A. First-Order Asymptotics: $\dot{C}_i(0)$

As shown in [3, Thm. 2], the first-order coefficient is

$$\dot{C}_i(0) = \lim_{\mathcal{P} \downarrow 0} \frac{C_i(\mathcal{P})}{\mathcal{P}} \quad (9)$$

$$= \sup_{\mathcal{P} > 0} \frac{C_i(\mathcal{P})}{\mathcal{P}} = \sup_{P_X} \frac{I(X; Y_i)}{E[b(X)]}. \quad (10)$$

When the cost is $b(x) = x^2$, $\dot{C}_i(0)$ is referred to as capacity per unit-energy. The main result of [3] is the following:

Theorem 1. [3, Thm. 3] *If there is a free input symbol, i.e., $b(0) = 0$, then the capacity per unit-cost of a memoryless channel is equal to*

$$\dot{C}_i(0) = \sup_{x \in \mathcal{X} \setminus 0} \frac{D(P_{Y_i|X=x} || P_{Y_i|X=0})}{b(x)}, \quad (11)$$

where $P_{Y_i|X=x}$ denotes the distribution of Y_i given that the input is x , and $D(P||Q)$ is the Kullback-Leibler divergence (i.e., relative entropy) between probability distributions P and Q :

$$D(P||Q) = \begin{cases} \int \log \left(\frac{dP}{dQ} \right) dP, & P \ll Q \\ +\infty, & \text{otherwise.} \end{cases} \quad (12)$$

Whereas (10) requires optimization over input distributions P_X in order to compute $\dot{C}_i(0)$, (11) requires optimization only over the non-zero input letters $x \in \mathcal{X} \setminus 0$. Further, (11) requires evaluation of the divergence between two conditional output distributions, which is often simpler to compute than the mutual information. An input distribution parameterized by \mathcal{P} is said to be *first-order optimal* if the ratio $I(X; Y_i)/\mathcal{P}$ tends to $\dot{C}_i(0)$ as \mathcal{P} tends to zero [2, Def. 1]. Capacity-achieving distributions are first-order optimal by definition. For the real AWGN channel with quadratic cost function $b(x) = x^2$, we have

$$D(P_{Y_i|X=x} || P_{Y_i|X=0}) = (x^2 \log e) / 2\sigma_i^2, \quad (13)$$

$$\dot{C}_i(0) = (\log e) / 2\sigma_i^2. \quad (14)$$

It is well known that Gaussian signaling, i.e., $P_X \sim \mathcal{N}(0, \mathcal{P})$, achieves the AWGN channel capacity $C_i(\mathcal{P})$ and is therefore first-order optimal. Additionally, the OOK distribution P_X defined by

$$P_X(x) = \begin{cases} \alpha = \frac{\mathcal{P}}{a^2}, & x = a \\ 1 - \alpha, & x = 0 \end{cases} \quad (15)$$

for any fixed $a > 0$ is also first-order optimal [3]. That is, as power tends to zero by letting the on-probability α tend to zero, binary inputs can closely approach channel capacity.

B. Second-order Asymptotics: $\ddot{C}_i(0)$

The second-order term $\ddot{C}_i(0)$ is defined as

$$\ddot{C}_i(0) = 2 \lim_{\mathcal{P} \downarrow 0} \frac{C_i(\mathcal{P}) - \dot{C}_i(0)\mathcal{P}}{\mathcal{P}^2}, \quad (16)$$

which leads to the Taylor series representation (1) of the capacity-cost function for small \mathcal{P} , which we re-write here with the receiver index i made explicit:

$$C_i(\mathcal{P}) = \dot{C}_i(0)\mathcal{P} + \frac{1}{2}\ddot{C}_i(0)\mathcal{P}^2 + o(\mathcal{P}^2). \quad (17)$$

IV. CROSS-RECEIVER MUTUAL INFORMATION

As defined earlier, $I(Y_1; Y_2)$ is the CMI between received signals Y_1 and Y_2 , and can be expressed as:

$$I(Y_1; Y_2) = I(X; Y_1) + I(X; Y_2) - I(X; Y_1, Y_2) \quad (18)$$

$$\begin{aligned} &= D(P_{Y_1 Y_2} \| P_{Y_1 Y_2 | X=0}) \\ &\quad - D(P_{Y_1} \| P_{Y_1 | X=0}) - D(P_{Y_2} \| P_{Y_2 | X=0}), \end{aligned} \quad (19)$$

where $P_{Y_1 Y_2}$ is the joint distribution of Y_1 and Y_2 , and $P_{Y_1 Y_2 | X=0} = P_{Y_1 | X=0} P_{Y_2 | X=0}$. The goal of this paper is characterize $I(Y_1; Y_2)$ at low SNR for different source distributions, which are assumed to be fixed from the receiver's perspective. This is a different objective than other multiuser information theory settings, such as the two-user broadcast channel, in which the goal may be to determine the achievable rate region in two dimensions. For a given source P_X parameterized by \mathcal{P} , we define the mutual-information-cost functions:

$$I_i(\mathcal{P}) = I(X; Y_i) \quad i = 1, 2 \quad (20)$$

$$I_{1,2}(\mathcal{P}) = I(Y_1; Y_2). \quad (21)$$

Note that $I_i(\mathcal{P})$ depends on the distribution P_X and that $I_i(\mathcal{P}) \leq C_i(\mathcal{P})$, with equality iff P_X is a capacity-achieving input distribution at power \mathcal{P} . Further, there is no analogous term $C_{1,2}(\mathcal{P})$, since the cross-receiver "capacity" is not well-defined. We define the following CMI series coefficients in the low-power regime:

$$\dot{I}_i(0) = \lim_{\mathcal{P} \downarrow 0} \frac{I_i(\mathcal{P})}{\mathcal{P}} \quad (22)$$

$$\ddot{I}_i(0) = 2 \lim_{\mathcal{P} \downarrow 0} \frac{I_i(\mathcal{P}) - \dot{I}_i(0)\mathcal{P}}{\mathcal{P}^2} \quad (23)$$

$$\dot{I}_{1,2}(0) = \lim_{\mathcal{P} \downarrow 0} \frac{I_{1,2}(\mathcal{P})}{\mathcal{P}} \quad (24)$$

$$\ddot{I}_{1,2}(0) = 2 \lim_{\mathcal{P} \downarrow 0} \frac{I_{1,2}(\mathcal{P}) - \dot{I}_{1,2}(0)\mathcal{P}}{\mathcal{P}^2}. \quad (25)$$

That is, the low-SNR CMI-cost function is approximated as

$$I_{1,2}(\mathcal{P}) = \dot{I}_{1,2}(0)\mathcal{P} + \frac{1}{2}\ddot{I}_{1,2}(0)\mathcal{P}^2 + o(\mathcal{P}^2). \quad (26)$$

As we will see in Sec. V, there are a number of source distributions for which $\dot{I}_{1,2}(0) = 0$. For these sources, the CMI-cost function in (26) is

$$I_{1,2}(\mathcal{P}) = \frac{1}{2}\ddot{I}_{1,2}(0)\mathcal{P}^2 + o(\mathcal{P}^2). \quad (27)$$

V. CMI FOR THE GAUSSIAN NOISE CHANNEL

A. Gaussian Sources

This section considers Gaussian sources: $X \sim \mathcal{N}(0, \mathcal{P})$. While the capacity per unit-energy is $\dot{C}_i(0) = (\log e)/2\sigma_i^2$ for each receiver ($i = 1, 2$), the CMI per unit-energy is zero:

Theorem 2. *For the signaling model in Fig. 1 with Gaussian source X , the CMI per unit-energy between received signals Y_1 and Y_2 is zero:*

$$\dot{I}_{1,2}(0) = \lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}} = 0. \quad (28)$$

Proof. For Gaussian sources, we have

$$I(X; Y_1, Y_2) = \frac{1}{2} \log(1 + \gamma_{\text{eq}}). \quad (29)$$

As \mathcal{P} tends to zero, we have

$$\lim_{\mathcal{P} \downarrow 0} \frac{I(X; Y_1, Y_2)}{\mathcal{P}} = \lim_{\mathcal{P} \downarrow 0} \frac{\frac{1}{2} \log(1 + \mathcal{P}/\sigma_{\text{eq}}^2)}{\mathcal{P}} \quad (30)$$

$$= \lim_{\mathcal{P} \downarrow 0} \frac{1/\sigma_{\text{eq}}^2}{2(1 + \mathcal{P}/\sigma_{\text{eq}}^2)} \log e \quad (31)$$

$$= \frac{\log e}{2\sigma_{\text{eq}}^2}, \quad (32)$$

where (31) is by L'Hopital's rule. Thus, we have

$$\lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}} = \lim_{\mathcal{P} \downarrow 0} \frac{I(X; Y_1)}{\mathcal{P}} + \frac{I(X; Y_2)}{\mathcal{P}} - \frac{I(X; Y_1, Y_2)}{\mathcal{P}} \quad (33)$$

$$= \frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} - \left[\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right] \quad (34)$$

$$= 0. \quad (35)$$

□

Thm. 2 can be understood by expressing the Gaussian CMI in terms of the effective SNR defined in (5):

$$I(Y_1; Y_2) = \frac{1}{2} \log(1 + \gamma_{\text{eff}}). \quad (36)$$

Because the effective SNR γ_{eff} can be approximated as $\gamma_1\gamma_2 = \mathcal{P}^2/\sigma_1^2\sigma_2^2$ at low SNRs, it follows that the \mathcal{P}^2 term inside the log dominates the cost \mathcal{P} in the denominator of (28). The non-zero second-order term $\ddot{I}_{1,2}(0)$ is:

Theorem 3. For the signaling model in Fig. 1 with Gaussian source X , the second-order term $\ddot{I}_{1,2}(0)$ is given by:

$$\ddot{I}_{1,2}(0) = 2 \lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}^2} = \log e \prod_{i=1,2} \frac{1}{\sigma_i^2}. \quad (37)$$

Proof. Consider that

$$\frac{d}{d\mathcal{P}} \gamma_{\text{eff}} = \frac{d}{d\mathcal{P}} \left(\frac{\gamma_1\gamma_2}{1 + \gamma_1 + \gamma_2} \right) \quad (38)$$

$$= \frac{d}{d\mathcal{P}} \left(\frac{\mathcal{P}^2}{\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)} \right) \quad (39)$$

$$= \frac{2\mathcal{P}(\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)) - \mathcal{P}^2(\sigma_1^2 + \sigma_2^2)}{[\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)]^2} \quad (40)$$

$$= \frac{2\mathcal{P}\sigma_1^2\sigma_2^2 + \mathcal{P}^2(\sigma_1^2 + \sigma_2^2)}{[\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)]^2} \quad (41)$$

$$(42)$$

From (36), we have

$$\lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1, Y_2)}{\mathcal{P}^2} = \lim_{\mathcal{P} \downarrow 0} \frac{\frac{1}{2} \log(1 + \gamma_{\text{eff}})}{\mathcal{P}^2} \quad (43)$$

$$= \frac{\log e}{2} \lim_{\mathcal{P} \downarrow 0} \frac{(1 + \gamma_1 + \gamma_2) \frac{d}{d\mathcal{P}}(\gamma_{\text{eff}})}{2\mathcal{P}(1 + \gamma_1 + \gamma_2 + \gamma_1\gamma_2)} \quad (44)$$

$$= \frac{\log e}{2} \lim_{\mathcal{P} \downarrow 0} \frac{[\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)] \frac{d}{d\mathcal{P}}(\gamma_{\text{eff}})}{2\mathcal{P}[\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2) + \mathcal{P}^2]} \quad (45)$$

$$= \frac{\log e}{2} \lim_{\mathcal{P} \downarrow 0} \frac{2\mathcal{P}\sigma_1^2\sigma_2^2 + \mathcal{P}^2(\sigma_1^2 + \sigma_2^2)}{2\mathcal{P}[\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2) + \mathcal{P}^2][\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)]} \quad (46)$$

$$= \frac{\log e}{2} \lim_{\mathcal{P} \downarrow 0} \frac{2\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)}{2[\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2) + \mathcal{P}^2][\sigma_1^2\sigma_2^2 + \mathcal{P}(\sigma_1^2 + \sigma_2^2)]} \quad (47)$$

$$= \frac{\log e}{2\sigma_1^2\sigma_2^2}, \quad (48)$$

where (44) is by L'Hopital's rule. □

B. BPSK Sources

This section considers equiprobable, antipodal sources, such as BPSK. The output distribution is given by

$$P_{Y_i} = \frac{1}{2}P_{Y_i|X=+a} + \frac{1}{2}P_{Y_i|X=-a}, \quad (49)$$

where $a^2 = \mathcal{P}$. BPSK sources have the same first- and second-order expansion of CMI as Gaussian sources:

Theorem 4. For the signaling model in Fig. 1 with equiprobable, antipodal source X (e.g., BPSK signaling), the first- and second-order terms $\dot{I}_{1,2}(0)$ and $\ddot{I}_{1,2}(0)$ are given by:

$$\dot{I}_{1,2}(0) = \lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}} = 0. \quad (50)$$

$$\ddot{I}_{1,2}(0) = 2 \lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}^2} = \log e \prod_{i=1,2} \frac{1}{\sigma_i^2}. \quad (51)$$

Proof. Thm. 4 is proved in Appendix A. □

C. OOK Sources, Fixed Amplitude

This section considers OOK sources with P_X defined in (15). In particular, we fix $a > 0$ and take $\mathcal{P} \downarrow 0$ by letting the on-probability $\alpha \downarrow 0$. Again, the CMI per unit-energy $\dot{I}_{1,2}(0)$ is zero:

Theorem 5. For the signaling model in Fig. 1 with OOK source P_X defined in (15), with fixed amplitude $a > 0$, the CMI per unit-energy between received signals Y_1 and Y_2 is zero:

$$\dot{I}_{1,2}(0) = \lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}} = 0. \quad (52)$$

Proof. The proof of Thm. 5 uses the following lemma from [3]:

Lemma 1. [3, p. 1023] For any pair of probability measures P and Q such that $P \ll Q$,

$$\lim_{\delta \downarrow 0} \frac{D(\delta P + (1 - \delta)Q \| Q)}{\delta} = 0. \quad (53)$$

Consider that

$$P_{Y_1} = \alpha P_{Y_1|X=a} + (1 - \alpha) P_{Y_1|X=0} \quad (54)$$

$$P_{Y_2} = \alpha P_{Y_2|X=a} + (1 - \alpha) P_{Y_2|X=0} \quad (55)$$

$$P_{Y_1 Y_2} = \alpha P_{Y_1|X=a} P_{Y_2|X=a} + (1 - \alpha) P_{Y_1|X=0} P_{Y_2|X=0}. \quad (56)$$

Application of Lemma 1 gives the following limits:

$$\lim_{\alpha \downarrow 0} \frac{D(P_{Y_1} \| P_{Y_1|X=0})}{\alpha} = 0 \quad (57)$$

$$\lim_{\alpha \downarrow 0} \frac{D(P_{Y_2} \| P_{Y_2|X=0})}{\alpha} = 0 \quad (58)$$

$$\lim_{\alpha \downarrow 0} \frac{D(P_{Y_1 Y_2} \| P_{Y_1 Y_2|X=0})}{\alpha} = 0. \quad (59)$$

Putting these together, and observing that the on-probability α and source power \mathcal{P} are related as $\mathcal{P} = \alpha a^2$ with $a > 0$ fixed, we have

$$\dot{I}_{1,2}(0) = \lim_{\alpha \downarrow 0} \frac{I(Y_1; Y_2)}{\alpha a^2} \quad (60)$$

$$= \frac{1}{a^2} \lim_{\alpha \downarrow 0} \frac{1}{\alpha} \left\{ D(P_{Y_1 Y_2} \| P_{Y_1 Y_2|X=0}) - D(P_{Y_1} \| P_{Y_1|X=0}) - D(P_{Y_2} \| P_{Y_2|X=0}) \right\} \quad (61)$$

$$= 0, \quad (62)$$

where (61) is from (19). □

The main result of this section is Thm. 6, which together with Thm. 3 shows that, at low SNR, the CMI of fixed-amplitude OOK sources exceeds the CMI of Gaussian sources.

Theorem 6. For the signaling model in Fig. 1 with OOK source P_X defined in (15), with fixed amplitude $a > 0$, the second-order term $\ddot{I}_{1,2}(0)$ is given by:

$$\ddot{I}_{1,2}(0) = 2 \lim_{\alpha \downarrow 0} \frac{I(Y_1; Y_2)}{\alpha^2} = \log e \prod_{i=1,2} \frac{e^{a^2/\sigma_i^2} - 1}{a^2}. \quad (63)$$

Proof. To prove Thm. 6, we will use the following lemma from [2]:

Lemma 2. [2, (162)] For any pair of probability measures P and Q such that $P \ll Q$,

$$\lim_{\delta \downarrow 0} \frac{D(\delta P + (1-\delta)Q||Q)}{\delta^2} = \frac{\log e}{2} \chi_2(P||Q), \quad (64)$$

where $\chi_2(P||Q)$ is Pearson's χ^2 -divergence⁵:

$$\chi_2(P||Q) = E_Q \left[\left(\frac{dP}{dQ} - 1 \right)^2 \right]. \quad (65)$$

Application of Lemma 2 gives the following limits:

$$\lim_{\alpha \downarrow 0} \frac{D(P_{Y_1}||P_{Y_1|X=0})}{\alpha^2} = \frac{\log e}{2} \chi_2(P_{Y_1|X=a}||P_{Y_1|X=0}) \quad (66)$$

$$\lim_{\alpha \downarrow 0} \frac{D(P_{Y_2}||P_{Y_2|X=0})}{\alpha^2} = \frac{\log e}{2} \chi_2(P_{Y_2|X=a}||P_{Y_2|X=0}) \quad (67)$$

$$\begin{aligned} \lim_{\alpha \downarrow 0} \frac{D(P_{Y_1 Y_2}||P_{Y_1 Y_2|X=0})}{\alpha^2} & \quad (68) \\ &= \frac{\log e}{2} \chi_2(P_{Y_1|X=a} P_{Y_2|X=a} || P_{Y_1|X=0} P_{Y_2|X=0}). \end{aligned}$$

Observing that $\chi_2(P||Q) = E_P \left(\frac{dP}{dQ} \right) - 1$, we have

$$\chi_2(P_1 P_2 || Q_1 Q_2) = E_{P_1 P_2} \left(\frac{dP_1}{dQ_1} \frac{dP_2}{dQ_2} \right) - 1 \quad (69)$$

$$= E_{P_1} \left(\frac{dP_1}{dQ_1} \right) E_{P_2} \left(\frac{dP_2}{dQ_2} \right) - 1, \quad (70)$$

where $P_i \ll Q_i$ and $P_1 P_2 \ll Q_1 Q_2$. Therefore,

$$\chi_2(P_1 P_2 || Q_1 Q_2) - \chi_2(P_1 || Q_1) - \chi_2(P_2 || Q_2) \quad (71)$$

$$\begin{aligned} &= \left[E_{P_1} \left(\frac{dP_1}{dQ_1} \right) - 1 \right] \left[E_{P_2} \left(\frac{dP_2}{dQ_2} \right) - 1 \right] \\ &= \chi_2(P_1 || Q_1) \chi_2(P_2 || Q_2). \end{aligned} \quad (72)$$

Letting α tend to zero, we have

$$\ddot{I}_{1,2}(0) = \frac{2}{a^4} \lim_{\alpha \downarrow 0} \frac{1}{\alpha^2} \{ D(P_{Y_1 Y_2} || P_{Y_1 Y_2|X=0}) \quad (73)$$

$$\begin{aligned} & - D(P_{Y_1} || P_{Y_1|X=0}) - D(P_{Y_2} || P_{Y_2|X=0}) \} \\ &= \frac{\log e}{a^4} \chi_2(P_{Y_1|X=a} || P_{Y_1|X=0}) \chi_2(P_{Y_2|X=a} || P_{Y_2|X=0}). \end{aligned} \quad (74)$$

For the real AWGN channel, the χ^2 -divergence and second-order term $\ddot{I}_{1,2}(0)$ are given by

$$\chi_2(P_{Y_i|X=x} || P_{Y_i|X=0}) = e^{x^2/\sigma_i^2} - 1, \quad (75)$$

$$\ddot{I}_{1,2}(0) = \log e \prod_{i=1,2} \frac{(e^{a^2/\sigma_i^2} - 1)}{a^2}. \quad (76)$$

□

⁵The χ^2 -divergence in (64) has the arguments switched relative to (162) in [2]; [2] appears to have an error in the statement of the lemma, but not in the proof. That (64) is correct can be verified by inspecting equations (164), (166), and (167) of [2].

Remarks: Thm. 6 requires fixed amplitude $a < \infty$. Although the ratio $\check{I}_{1,2}^{\text{OOK}}(0)/\check{I}_{1,2}^{\text{Gaussian}}(0)$ is unbounded as $a \rightarrow \infty$, letting $a \rightarrow \infty$ would violate a lemma used in the proof of Thm. 6.

Corollary 1. *For the signaling model in Fig. 1, the second-order term $\check{I}_{1,2}(0)$ of an OOK source P_X defined in (15), with fixed amplitude $a > 0$, is greater than that of a Gaussian source, i.e.,*

$$\check{I}_{1,2}^{\text{OOK}}(0) > \check{I}_{1,2}^{\text{Gaussian}}(0). \quad (77)$$

Proof. We start with the upper bound on the natural logarithm for $x > 0$:

$$\ln(1+x) < x \quad (78)$$

$$1+x < e^x \quad (79)$$

$$x < e^x - 1 \quad (80)$$

$$\frac{a^2}{\sigma_i^2} < e^{a^2/\sigma_i^2} - 1 \quad (81)$$

$$\frac{1}{\sigma_i^2} < \frac{e^{a^2/\sigma_i^2} - 1}{a^2}. \quad (82)$$

Equation (82) is true for $i=1, 2$, for all positive σ_i^2 , proving (77) by comparing (63) and (37). \square

We have shown that Gaussian sources, which are capacity-achieving for single-receiver channels $X-Y_i$, are not optimal in terms of the second-order CMI coefficient $\check{I}_{1,2}(0)$. Because the CMI per unit-energy is zero, we have:

Corollary 2. *For the signaling model in Fig. 1, as source power \mathcal{P} tends to zero, the CMI $I_{1,2}(\mathcal{P})$ of an OOK source P_X defined in (15), with fixed amplitude $a > 0$, is greater than that of a Gaussian source, i.e.,*

$$\lim_{\mathcal{P} \downarrow 0} \frac{I_{1,2}^{\text{OOK}}(\mathcal{P})}{I_{1,2}^{\text{Gaussian}}(\mathcal{P})} = \prod_{i=1,2} \frac{e^{a^2/\sigma_i^2} - 1}{a^2/\sigma_i^2} > 1. \quad (83)$$

Remarks: While we have shown that Gaussian signaling and BPSK signaling are not optimal in terms of the CMI in the low-SNR regime, we have not established that fixed-amplitude OOK is necessarily optimal. That is, there may be another source distribution with CMI exceeding that of OOK.

D. OOK Sources, Fixed Duty Cycle, Vanishing Amplitude

In Sec. V-C, we analyzed OOK sources with fixed source amplitude a and let the average duty-cycle α tend to zero. In this section, we analyze OOK sources and fix $\alpha \in (0, 1)$, letting power \mathcal{P} tend to zero by letting the amplitude a tend to zero, which is a better model for practical applications with a fixed transmit waveform (i.e., constant duty cycle). For example, received power may vanish due to propagation loss as the transmitter-receiver distance increases.

Theorem 7. *For the signaling model in Fig. 1 with OOK source P_X defined in (15), with fixed on-probability $\alpha \in (0, 1)$, the first- and second-order terms $\dot{I}_{1,2}(0)$ and $\ddot{I}_{1,2}(0)$ are given by:*

$$\dot{I}_{1,2}(0) = \lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}} = 0, \quad (84)$$

$$\ddot{I}_{1,2}(0) = 2 \lim_{\mathcal{P} \downarrow 0} \frac{I(Y_1; Y_2)}{\mathcal{P}^2} = \log e \prod_{i=1,2} \frac{(1-\alpha)}{\sigma_i^2}. \quad (85)$$

Proof. Thm. 7 is proved in Appendix B. \square

Comparing to Thm. 3 for Gaussian sources, we see that fixed-duty OOK sources are strictly worse than Gaussian sources in terms of the second-order coefficient, with the ratio $\check{I}_{1,2}^{\text{OOK}}(0)/\check{I}_{1,2}^{\text{Gaussian}}(0) = (1-\alpha)^2 < 1$. For sufficiently small α , however, the difference is negligible. Numerical examples are provided in the next section.



Fig. 2. (a) Signal model showing additive noise and quantization at receiver i and (b) Markov chain representation of transmitted and received signals.

E. OOK Sources, Fixed Duty Cycle, Finite Amplitude

Sec. V-C has established the *asymptotic* superiority of fixed-amplitude OOK versus Gaussian signaling in terms of CMI, while Sec. V-D has shown that fixed-on-probability OOK is strictly worse than Gaussian signaling. However, it's not clear from the asymptotic analysis how fixed-on-probability OOK and Gaussian signaling compare at small, but non-vanishing SNRs. This section considers OOK sources with P_X defined in (15), again with fixed on-probability $\alpha \in (0, \frac{1}{2})$, varying \mathcal{P} by varying the amplitude a . Numerical results in this section show that for a sufficiently small duty cycle ($\alpha \ll 1$), there exist a wide range of source powers $\mathcal{P} > 0$ such that $I_{1,2}^{\text{OOK}}(\mathcal{P}) > I_{1,2}^{\text{Gaussian}}(\mathcal{P})$, in stark contrast to the asymptotic conclusion of Sec. V-D.

To facilitate numerical evaluation of the CMI at low SNR, instead of evaluating $I(Y_1, Y_2)$, we evaluate the the CMI $I(\hat{Y}_1, \hat{Y}_2)$ of the 1-bit quantized signals \hat{Y}_1 and \hat{Y}_2 , defined as

$$\hat{Y}_i = \begin{cases} 1, & Y_i \geq \Delta_i \\ 0, & Y_i < \Delta_i, \end{cases} \quad (86)$$

where Δ_i is a receiver-specific quantization threshold. For simplicity, we will assume that the quantization thresholds are symmetric with respect to the source values 0 and a , i.e., $\Delta_i = \frac{a}{2}$. As a result, the conditional “error” probabilities ϵ_i of the quantizers are symmetric:

$$\epsilon_i = P_{Y_i|X=a}(0) = P_{Y_i|X=0}(1) = Q\left(\frac{a}{2\sigma_i}\right), \quad (87)$$

where $Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{t^2}{2}} dt$ is the tail of the standard normal distribution.⁶

Fig. 3 compares the CMI of Gaussian signaling with the CMI of OOK as a function of source power, for multiple values of the OOK on-probability α . All curves in this figure assume $\sigma_1^2 = 1/100$, $\sigma_2^2 = 1$. For Gaussian sources, Fig. 3 includes the single-receiver MI $I(X; Y_i)$ and the CMI $I(Y_1; Y_2)$. For each OOK curve in the figure, corresponding to a different value of α , the SNR $\gamma_i = \mathcal{P}/\sigma_i^2 = \alpha a^2/\sigma_i^2$ is varied by adjusting the amplitude a . Note that the CMI is displayed on the vertical axis using a logarithmic scale.

As the average power decays to zero, we observe that there is a wide range of powers \mathcal{P} such that the 1-bit quantized OOK CMI $I(\hat{Y}_1; \hat{Y}_2)$ is approximately constant before dropping off, particularly for small values of α . For each α , denote $\gamma_{\min}(\alpha)$ as the minimum SNR for which $I(\hat{Y}_1; \hat{Y}_2)$ is approximately constant. Markers in the OOK curves indicate the power levels for which SNR $\gamma_i = \gamma_{\min}(\alpha)$. For example, when $\alpha = 10^{-6}$, $I(\hat{Y}_1; \hat{Y}_2) \approx 2.1 \times 10^{-5}$ for all SNRs above $\gamma_{\min}(10^{-6}) \approx -39$ dB. Approximations for this constant CMI level and the minimum SNR $\gamma_{\min}(\alpha)$ are described by the following proposition:

Similar to Fig. 3, Fig. 4 compares the CMI of Gaussian signaling with the CMI of OOK as a function of source power, but with equal noise powers: $\sigma_1^2 = \sigma_2^2 = 1$. For Gaussian sources, Fig. 4 additionally includes the 1-bit quantized CMI $I(\hat{Y}_1; \hat{Y}_2)$ corresponding to symmetric quantization thresholds $\Delta_i = 0$.

Again, there is a wide range of powers \mathcal{P} such that the 1-bit quantized OOK CMI $I(\hat{Y}_1; \hat{Y}_2)$ is approximately constant. By Prop. 1, as long as $\gamma_i \geq \gamma_{\min}^{\text{OOK}}(\alpha)$, $I(\hat{Y}_1; \hat{Y}_2) \approx H_B(\alpha)$.

Proposition 1. *The CMI $I(\hat{Y}_1; \hat{Y}_2)$ of an OOK source P_X defined in (15), with fixed on-probability $\alpha \ll 1$, can be approximated as*

$$I(\hat{Y}_1; \hat{Y}_2) \approx H_B(\alpha), \quad (88)$$

where $H_B(\alpha)$ is the binary entropy:

$$H_B(\alpha) = -\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha), \quad (89)$$

⁶From the Markov chain $\hat{Y}_1 - Y_1 - X - Y_2 - \hat{Y}_2$, we observe that $I(Y_1; Y_2) \geq I(\hat{Y}_1; \hat{Y}_2)$ by the Data Processing Inequality. Therefore, results in this section using $I(\hat{Y}_1; \hat{Y}_2)$ may be pessimistic, in the sense that the OOK CMI $I(Y_1; Y_2)$ is not explicitly evaluated. Additionally, the symmetric-threshold restriction may further degrade $I(\hat{Y}_1; \hat{Y}_2)$ with respect to optimal quantization.

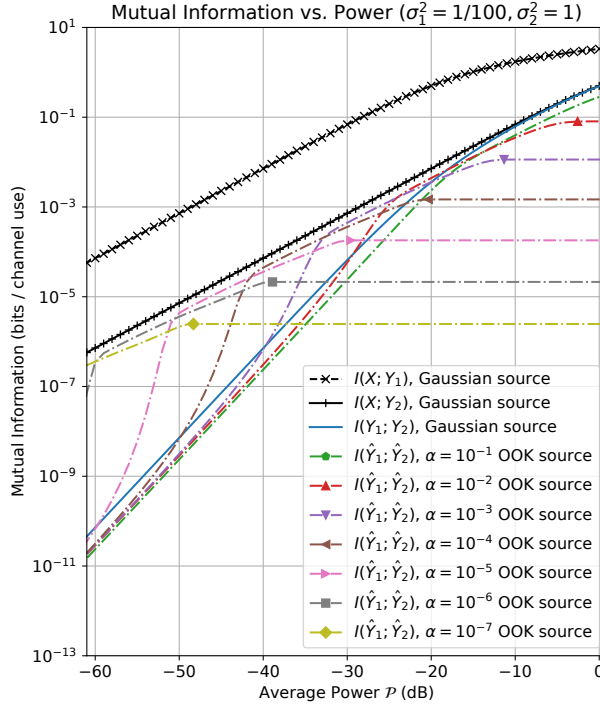


Fig. 3. Comparison of cross-receiver mutual information (CMI) for Gaussian sources and OOK sources, assuming *different* noise variances: $\sigma_1^2=1/100$, $\sigma_2^2=1$. OOK sources are parameterized by the on-probability α . Markers in the OOK curves indicate the minimum power $\mathcal{P}_{\min}^{\text{OOK}}(\alpha)$ for which $I(\hat{Y}_1; \hat{Y}_2) \approx H_B(\alpha)$.

whenever the SNRs γ_i at each receiver ($i=1,2$) are greater than or equal to $\gamma_{\min}^{\text{OOK}}(\alpha)$:

$$\gamma_{\min}^{\text{OOK}}(\alpha) = -8\alpha \ln \frac{\alpha}{r}, \quad (90)$$

for some constant $r \gg 1$.

Proof. Defining $t_i = \alpha(1 - 2\epsilon_i)$ and $t_{12} = \alpha(1 - \epsilon_1 - \epsilon_2)$, the CMI of the 1-bit quantized signals \hat{Y}_1 and \hat{Y}_2 is

$$I(\hat{Y}_1; \hat{Y}_2) = H(\hat{Y}_1) + H(\hat{Y}_2) - H(\hat{Y}_1, \hat{Y}_2) \quad (91)$$

$$= H_B(\epsilon_1 + t_1) + H_B(\epsilon_2 + t_2) \quad (92)$$

$$+ [\epsilon_1 \epsilon_2 + t_{12}] \log[\epsilon_1 \epsilon_2 + t_{12}]$$

$$+ [(1 - \epsilon_1)(1 - \epsilon_2) - t_{12}] \log[(1 - \epsilon_1)(1 - \epsilon_2) - t_{12}]$$

$$+ [\alpha(\epsilon_1 - \epsilon_2) + \epsilon_2(1 - \epsilon_1)] \log[\alpha(\epsilon_1 - \epsilon_2) + \epsilon_2(1 - \epsilon_1)]$$

$$+ [\alpha(\epsilon_2 - \epsilon_1) + \epsilon_1(1 - \epsilon_2)] \log[\alpha(\epsilon_2 - \epsilon_1) + \epsilon_1(1 - \epsilon_2)].$$

Assume now that $\epsilon_i \ll \alpha$ for both $i=1$ and $i=2$. In this region, $I(\hat{Y}_1; \hat{Y}_2)$ is dominated by the first four terms on the right-hand side of (92); the final two terms vanish. We can approximate the CMI as

$$I(\hat{Y}_1; \hat{Y}_2) \approx H_B(\epsilon_1 + t_1) + H_B(\epsilon_2 + t_2) \quad (93)$$

$$- H_B(\epsilon_1 \epsilon_2 + t_{12})$$

$$\approx H_B(\alpha). \quad (94)$$

Suppose we specifically have $\epsilon_i \leq \frac{\alpha}{r}$ for some $r \gg 1$, for both receivers ($i=1,2$). What is the minimum SNR $\gamma_{\min}^{\text{OOK}}(\alpha)$ required to achieve $\epsilon_i \leq \frac{\alpha}{r}$? Consider that

$$\epsilon_i = Q\left(\frac{a}{2\sigma_i}\right) \quad (95)$$

$$= Q\left(\frac{1}{2}\sqrt{\frac{\gamma_i}{\alpha}}\right) \quad (96)$$

$$\leq e^{-\gamma_i/8\alpha}, \quad (97)$$

where (97) is by the Chernoff bound for the Q-function. By setting γ_i greater than or equal to $\gamma_{\min}^{\text{OOK}}(\alpha) = -8\alpha \ln \frac{\alpha}{r}$, we have

$$\frac{\gamma_i}{8\alpha} \geq -\ln \frac{\alpha}{r} \quad (98)$$

$$e^{-\gamma_i/8\alpha} \leq \frac{\alpha}{r}. \quad (99)$$

By (97), we have guaranteed that $\epsilon_i \leq \frac{\alpha}{r} \ll \alpha$. As a result, $I(\hat{Y}_1; \hat{Y}_2) \approx H_B(\alpha)$. \square

Since we require that $\gamma_i \geq \gamma_{\min}^{\text{OOK}}(\alpha)$ for both receivers in order for the approximation of Prop. 1 to be valid, the relationship between the minimum SNR $\gamma_{\min}^{\text{OOK}}(\alpha)$ and the minimum power $\mathcal{P}_{\min}^{\text{OOK}}(\alpha)$ is given by

$$\mathcal{P}_{\min}^{\text{OOK}}(\alpha) = \max\{\sigma_1^2, \sigma_2^2\} \gamma_{\min}^{\text{OOK}}(\alpha). \quad (100)$$

The OOK markers in the figure indicate the minimum power $\mathcal{P}_{\min}^{\text{OOK}}(\alpha)$ at which OOK CMI can be approximated as $I(\hat{Y}_1; \hat{Y}_2) \approx H_B(\alpha)$. Consider the curve formed by all coordinates $(\mathcal{P}_{\min}^{\text{OOK}}(\alpha), H_B(\alpha))$ as α tends to zero, i.e., the envelope of the individual OOK(α) CMI curves. Using the approximation for $\gamma_{\min}^{\text{OOK}}(\alpha)$ in Prop. 1, it can be shown that the slope of this envelope is equal to the slope of the Gaussian source input-output mutual information curve. (Note the figure uses a log-log scale.) Gaussian sources are first-order optimal and in the low-power regime have input-output mutual information linear in source power: $I_i(\mathcal{P}) = O(\mathcal{P})$. As a result, the envelope of the OOK CMI curves is also linear in source power, as shown in the following proposition.

Proposition 2. *The family of fixed-on-probability OOK sources parameterized by α has CMI curves $I(\hat{Y}_1; \hat{Y}_2)$ vs. power \mathcal{P} whose envelope $(\mathcal{P}_{\min}^{\text{OOK}}(\alpha), I_{1,2}(\mathcal{P}_{\min}^{\text{OOK}}(\alpha)))$ as α tends to zero can be approximated as*

$$I_{1,2}(\mathcal{P}_{\min}^{\text{OOK}}(\alpha)) = O(\mathcal{P}_{\min}^{\text{OOK}}(\alpha)). \quad (101)$$

Before proving Prop. 2, we first state the following lemma:

Lemma 3. *Let $r > 0$ be a constant. Then*

$$\lim_{\alpha \downarrow 0} \frac{H_B(\alpha)}{-\alpha \ln \frac{\alpha}{r}} = \log e \quad (102)$$

Proof.

$$\lim_{\alpha \downarrow 0} \frac{H_B(\alpha)}{-\alpha \ln \frac{\alpha}{r}} = (\log e) \lim_{\alpha \downarrow 0} \frac{-\alpha \ln \alpha - (1 - \alpha) \ln(1 - \alpha)}{-\alpha \ln \alpha + \alpha \ln r} \quad (103)$$

$$= (\log e) \lim_{\alpha \downarrow 0} \frac{-1/\alpha - 1/(1 - \alpha)}{-1/\alpha} \quad (104)$$

$$= \log e \quad (105)$$

where (104) follows from applying L'Hopitals rule twice. \square

Proof of Prop. 2. To show that the OOK CMI curves' envelope is linear in source power, we need to show that

$$\lim_{\alpha \downarrow 0} \frac{I_{1,2}(\mathcal{P}_{\min}^{\text{OOK}}(\alpha))}{\mathcal{P}_{\min}^{\text{OOK}}(\alpha)} = c \quad (106)$$

for some constant $c > 0$. By the approximation in Prop. 1,

$$\lim_{\alpha \downarrow 0} \frac{I_{1,2}(\mathcal{P}_{\min}^{\text{OOK}}(\alpha))}{\mathcal{P}_{\min}^{\text{OOK}}(\alpha)} \approx \frac{1}{\max\{\sigma_1^2, \sigma_2^2\}} \lim_{\alpha \downarrow 0} \frac{H_B(\alpha)}{\gamma_{\min}^{\text{OOK}}(\alpha)} \quad (107)$$

$$= \frac{1}{\max\{\sigma_1^2, \sigma_2^2\}} \lim_{\alpha \downarrow 0} \frac{H_B(\alpha)}{-8\alpha \ln \frac{\alpha}{r}} \quad (108)$$

$$= \frac{\log e}{8 \max\{\sigma_1^2, \sigma_2^2\}}, \quad (109)$$

where (109) is by Lemma 3. \square

Thm. 7 shows that for a given on-probability α , OOK CMI in the low-power regime is quadratic in source power. Further, the $(1 - \alpha)^2$ constant in Thm. 7 makes fixed- α OOK source worse than Gaussian sources in terms of CMI. Remarkably, however, Prop. 2 shows that the *family* of fixed- α OOK CMI curves has an envelope that is linear in source power. Whereas

the CMI of Gaussian sources can be described by the effective SNR $\gamma_{\text{eff}} \approx \mathcal{P}^2 / \sigma_1^2 \sigma_2^2$ as in (36), effective SNR is only an accurate description of OOK CMI for a fixed α when the individual SNRs γ_1 and γ_2 are well below the minimum SNR $\gamma_{\min}^{\text{OOK}}(\alpha)$. Effective SNR is not meaningful when OOK SNRs are above $\gamma_{\min}^{\text{OOK}}(\alpha)$.

The remainder of this section further compares the CMI of OOK and Gaussian sources. From Fig. 4, we see that although there is a range of SNRs for which the OOK source CMI is greater than the Gaussian source CMI, as SNR increases, eventually the Gaussian source CMI surpasses the OOK CMI. What is the minimum SNR $\gamma_{\min}^{\text{Gaussian}}(\alpha)$ for a Gaussian source to outperform an OOK(α) source in terms of CMI? Since $I^{\text{OOK}}(\hat{Y}_1; \hat{Y}_2) \approx H_B(\alpha)$, we must figure out when $I^{\text{Gaussian}}(Y_1; Y_2) \geq H_B(\alpha)$. That is, we define the minimum SNR $\gamma_{\min}^{\text{Gaussian}}(\alpha)$ as

$$\gamma_{\min}^{\text{Gaussian}}(\alpha) = \min\{\gamma > 0 : I_{1,2}^{\text{Gaussian}}(\mathcal{P}) \geq H_B(\alpha)\}. \quad (110)$$

where we've assumed equal receiver SNRs: $\gamma_1 = \gamma_2 = \gamma$. Note that while the statistics of the Gaussian source do not depend on α , the term $\gamma_{\min}^{\text{Gaussian}}(\alpha)$ is parameterized by α because it is comparing the Gaussian source relative to the OOK(α) source.

From Thms. 2 and 3 in Sec. V-A, the Gaussian-source CMI is approximated in the low-SNR regime as

$$I_{1,2}^{\text{Gaussian}}(\mathcal{P}) \approx \frac{\log e}{2\sigma_1^2 \sigma_2^2} \mathcal{P}^2 \quad (111)$$

$$= \frac{\log e}{2} \gamma^2. \quad (112)$$

when $\gamma_1 = \gamma_2 = \gamma$. We set the minimum SNR according to

$$\frac{\log e}{2} (\gamma_{\min}^{\text{Gaussian}}(\alpha))^2 = H_B(\alpha) \quad (113)$$

$$\gamma_{\min}^{\text{Gaussian}}(\alpha) = \sqrt{2 \frac{H_B(\alpha)}{\log e}}. \quad (114)$$

Therefore, as long as $\gamma_i \geq \gamma_{\min}^{\text{Gaussian}}(\alpha)$, we have $I_{1,2}^{\text{Gaussian}}(\mathcal{P}) \geq H_B(\alpha) \approx I_{1,2}^{\text{OOK}}(\mathcal{P})$.

The following proposition describes the relationship between $\gamma_{\min}^{\text{Gaussian}}(\alpha)$ and $\gamma_{\min}^{\text{OOK}}(\alpha)$ as α tends to zero:

Proposition 3. *We observe the following scaling of the minimum SNRs as the on-probability α tends to zero:*

$$\lim_{\alpha \downarrow 0} \frac{\gamma_{\min}^{\text{Gaussian}}(\alpha)}{\gamma_{\min}^{\text{OOK}}(\alpha)} = \infty \quad (115)$$

$$\lim_{\alpha \downarrow 0} \frac{(\gamma_{\min}^{\text{Gaussian}}(\alpha))^2}{\gamma_{\min}^{\text{OOK}}(\alpha)} = \frac{1}{4}. \quad (116)$$

Remarks: Prop. 3 indicates that as the on-probability vanishes, Gaussian signaling requires infinitely greater SNR to match or outperform OOK in terms of CMI. However, the ratio of the *squared* minimum Gaussian-source SNR to the minimum OOK SNR is constant.

Proof. Consider that as α tends to zero, we have

$$H_B(\alpha) = (\log e) [-\alpha \ln \alpha - (1 - \alpha) \ln(1 - \alpha)] \quad (117)$$

$$\approx -(\log e) \alpha \ln \alpha. \quad (118)$$

To prove (115), we have

$$\lim_{\alpha \downarrow 0} \frac{\gamma_{\min}^{\text{Gaussian}}(\alpha)}{\gamma_{\min}^{\text{OOK}}(\alpha)} = \lim_{\alpha \downarrow 0} \frac{\sqrt{2H_B(\alpha)/\log e}}{-8 \alpha \ln \frac{\alpha}{r}} \quad (119)$$

$$= \frac{\sqrt{2}}{8} \lim_{\alpha \downarrow 0} \frac{\sqrt{H_B(\alpha)/\log e}}{-\alpha \ln \frac{\alpha}{r}} \quad (120)$$

$$= \frac{\sqrt{2}}{8} \lim_{\alpha \downarrow 0} \frac{\sqrt{-\alpha \ln \alpha}}{-\alpha \ln \frac{\alpha}{r}} \quad (121)$$

$$= +\infty. \quad (122)$$

Thus, as α tends to zero, the ratio $\gamma_{\min}^{\text{Gaussian}}(\alpha)/\gamma_{\min}^{\text{OOK}}(\alpha)$ behaves as $\sqrt{2}/(8\sqrt{-\alpha \ln \alpha})$.

To prove (116):

$$\lim_{\alpha \downarrow 0} \frac{(\gamma_{\min}^{\text{Gaussian}}(\alpha))^2}{\gamma_{\min}^{\text{OOK}}(\alpha)} = \frac{1}{4} \lim_{\alpha \downarrow 0} \frac{H_B(\alpha)/\log e}{-\alpha \ln \frac{\alpha}{r}} \quad (123)$$

$$= \frac{1}{4}. \quad (124)$$

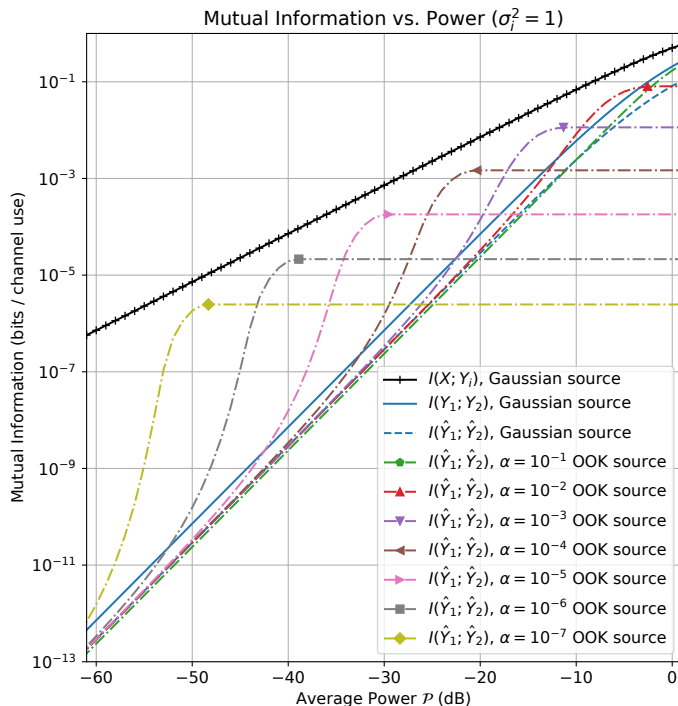


Fig. 4. Comparison of cross-receiver mutual information (CMI) for Gaussian sources and OOK sources, assuming *equal* noise variances: $\sigma_1^2 = \sigma_2^2 = 1$. OOK sources are parameterized by the on-probability α . Markers in the OOK curves indicate the minimum power $\mathcal{P}_{\min}^{\text{OOK}}(\alpha)$ for which $I(\hat{Y}_1; \hat{Y}_2) \approx H_B(\alpha)$.

where in (124) we have used Lemma 3. □

VI. CONCLUSION

Motivated by distributed, passive receiver applications with low-SNR, we defined the CMI-cost function $I_{1,2}(\mathcal{P})$ in terms of its first- and second-order series coefficients. Focusing on the AWGN channel, we evaluated these coefficients for three source distributions: Gaussian, BPSK, and OOK sources. For all three distributions, the first-order CMI coefficient is zero, indicating that CMI is quadratic in the source power \mathcal{P} . Gaussian and BPSK sources have the same second-order coefficient, which is less than the second-order coefficient for fixed-amplitude OOK sources, regardless of amplitude. For OOK sources with fixed on-probability, however, the asymptotic analysis indicates that the CMI of OOK sources is less than that of Gaussian sources. At small but non-zero SNRs, fixed-on-probability OOK sources have approximately constant CMI vs. SNR, which for sufficiently small on-probability can be much greater than the CMI of Gaussian sources. Further, in the limit of zero on-probability, the envelope of all fixed-on-probability OOK CMI curves is linear in source power, instead of quadratic. This indicates that the effective input SNR, while common in the TDOA estimation literature for Gaussian sources, may be misleading for OOK sources, depending on the individual receivers' SNRs. Due to the ability of CMI to capture arbitrary statistical dependences, CMI-based receiver processing techniques should be investigated for specific distributed inference problems.

ACKNOWLEDGEMENTS

A. Williamson would like to thank Dr. Johnathan York, Travis Cuvelier, and Dr. Boulat Bash for their helpful feedback on this work.

APPENDIX

A. Proof of Thm. 4

This section proves Thm. 4, which establishes the first- and second-order series coefficients $\dot{I}_{1,2}(0)$ and $\ddot{I}_{1,2}(0)$ for equiprobable, antipodal signaling (e.g., BPSK).

Proof. BPSK signaling on the Gaussian channel has the following KL divergences:

$$\begin{aligned} D(P_{Y_i}||P_{Y_i|X=0}) &= E_{P_{Y_i}} \left[\log \frac{\frac{1}{2}dP_{Y_i|X=+a} + \frac{1}{2}dP_{Y_i|X=-a}}{dP_{Y_i|X=0}} \right] \\ &= E_{P_{Y_i}} \left[\log \cosh \frac{aY_i}{\sigma_i^2} \right] - \frac{a^2}{2\sigma_i^2} \end{aligned} \quad (125)$$

$$\begin{aligned} D(P_{Y_1Y_2}||P_{Y_1Y_2|X=0}) &= E_{P_{Y_1Y_2}} \left[\log \cosh a \left(\frac{Y_1}{\sigma_1^2} + \frac{Y_2}{\sigma_2^2} \right) \right] \\ &\quad - \frac{a^2}{2\sigma_1^2} - \frac{a^2}{2\sigma_2^2}. \end{aligned} \quad (126)$$

To compute the KL divergences, consider the following Taylor series expansions:

$$\cosh ab = 1 + \frac{1}{2}a^2b^2 + \frac{1}{4!}a^4b^4 + o(a^4) \quad (127)$$

$$= 1 + f(a) + o(a^4) \quad (128)$$

$$\log_e \cosh ab = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} [\cosh ab - 1]^k \quad (129)$$

$$= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} [f(a) + o(a^4)]^k \quad (130)$$

$$= [f(a) + o(a)] - \frac{1}{2} [f(a) + o(a^4)]^2 + o(a^4) \quad (131)$$

$$= f(a) - \frac{1}{2}f(a)^2 + o(a^4) \quad (132)$$

$$= \frac{1}{2}a^2b^2 - \frac{1}{12}a^4b^4 + o(a^4), \quad (133)$$

where in (128) we have defined $f(a) = \frac{1}{2}a^2b^2 + \frac{1}{4!}a^4b^4$, and where in (131)-(133) we have included only the terms that are of order 4 or less in a . With some algebra, it can be shown that

$$D(P_{Y_i}||P_{Y_i|X=0}) = (\log e) \frac{a^4}{4\sigma_i^4} + o(a^4), \quad (134)$$

$$D(P_{Y_1Y_2}||P_{Y_1Y_2|X=0}) = (\log e) \frac{a^4}{4\sigma_{\text{eq}}^4} + o(a^4), \quad (135)$$

where

$$\frac{1}{\sigma_{\text{eq}}^4} = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^2 \quad (136)$$

$$= \frac{1}{\sigma_1^4} + \frac{1}{\sigma_2^4} + \frac{2}{\sigma_1^2\sigma_2^2} \quad (137)$$

Therefore, the CMI is given by

$$I(Y_1; Y_2) = (\log e) \left(\frac{a^4}{4\sigma_{\text{eq}}^4} - \frac{a^4}{4\sigma_1^4} - \frac{a^4}{4\sigma_2^4} + o(a^4) \right) \quad (138)$$

$$= (\log e) \left(\frac{a^4}{2\sigma_1^2\sigma_2^2} + o(a^4) \right) \quad (139)$$

Taking the limit of (139) as $a^2 = \mathcal{P}$ tends to zero, we obtain the CMI series coefficients:

$$\dot{I}_{1,2}(0) = \lim_{a \downarrow 0} \frac{I(Y_1; Y_2)}{a^2} \quad (140)$$

$$= 0 \quad (141)$$

$$\ddot{I}_{1,2}(0) = 2 \lim_{a \downarrow 0} \frac{I(Y_1; Y_2)}{a^4} \quad (142)$$

$$= \frac{\log e}{\sigma_1^2\sigma_2^2}. \quad (143)$$

□

B. Proof of Thm. 7

This section proves Thm. 7, which establishes the first- and second-order series coefficients $\dot{I}_{1,2}(0)$ and $\ddot{I}_{1,2}(0)$ for OOK with fixed on-probability $\alpha \in (0, 1)$ and vanishing amplitude a .

Proof. Consider a random vector $Y=(Y_1, \dots, Y_n)$ distributed according to

$$Q_Y = \alpha Q_a + (1 - \alpha)Q_0, \quad (144)$$

where we have denoted the conditional probability distribution $Q_{Y|X=x}=Q_x$. Using the Taylor series expansion for the natural logarithm:

$$\log_e(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1} x^k}{k}, \quad |x| < 1, \quad (145)$$

we can write the KL divergence $D(Q||Q_0)$ in nats as

$$D(Q||Q_0) = E_{Q_Y} \left[\log_e \left(1 + \alpha \left(\frac{Q_a(Y)}{Q_0(Y)} - 1 \right) \right) \right] \quad (146)$$

$$= E_{Q_Y} \left[\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \alpha^k \left(\frac{Q_a(Y)}{Q_0(Y)} - 1 \right)^k \right] \quad (147)$$

$$= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \alpha^k E_{P_Y} \left(\frac{Q_a(Y)}{Q_0(Y)} - 1 \right)^k \quad (148)$$

$$= \sum_{k=2}^{\infty} \frac{(-1)^{k+1}}{k} (\alpha^{k+1} \chi_{k+1}(Q_a||Q_0) + \alpha^k \chi_k(Q_a||Q_0)) \quad (149)$$

$$= \sum_{k=2}^{\infty} \frac{(-1)^k \alpha^k}{k(k-1)} \chi_k(Q_a||Q_0), \quad (150)$$

where (149) is from the following observation:

$$E_{P_Y} \left(\frac{Q_a(Y)}{Q_0(Y)} - 1 \right)^k \quad (151)$$

$$\begin{aligned} &= \alpha E_{Q_a} \left(\frac{Q_a(Y)}{Q_0(Y)} - 1 \right)^k + (1 - \alpha) E_{Q_0} \left(\frac{Q_a(Y)}{Q_0(Y)} - 1 \right)^k \\ &= \alpha \chi_{k+1}(Q_a||Q_0) + \chi_k(Q_a||Q_0). \end{aligned} \quad (152)$$

The chi-divergence $\chi_k(Q_a||Q_0)$ can be expressed in terms of the moment-generating function as follows:

$$\chi_k(Q_a||Q_0) = E_{Q_0} \left(\frac{Q_a(Y)}{Q_0(Y)} - 1 \right)^k \quad (153)$$

$$= \sum_{\ell=0}^k \binom{k}{\ell} (-1)^{k-\ell} E_{Q_0} \left(\frac{Q_a(Y)}{Q_0(Y)} \right)^\ell \quad (154)$$

$$= (-1)^{k+1} (k-1) + \sum_{\ell=2}^k \binom{k}{\ell} (-1)^{k-\ell} E_{Q_0} \left(\frac{Q_a(Y)}{Q_0(Y)} \right)^\ell \quad (155)$$

$$= (-1)^{k+1} (k-1) + \sum_{\ell=2}^k \binom{k}{\ell} (-1)^{k-\ell} E_{Q_a} \left(\frac{Q_a(Y)}{Q_0(Y)} \right)^{\ell-1}, \quad (156)$$

where (155) follows because $E_{Q_0} \left(\frac{Q_a(Y)}{Q_0(Y)} \right)^\ell = 1$ for $\ell=0, 1$.

Conditioned on the source value x , the joint conditional probability distribution Q_x can be written as a product of marginal conditional probability distributions:

$$Q_x = \prod_{i=1}^n Q_{Y_i|X=x} \quad (157)$$

$$\frac{Q_a(Y)}{Q_0(Y)} = \prod_{i=1}^n \frac{Q_{Y_i|X=a}(Y_i)}{Q_{Y_i|X=0}(Y_i)} \quad (158)$$

$$= \prod_{i=1}^n e^{-a^2/2\sigma_i^2} e^{ay_i/\sigma_i^2} \quad (159)$$

$$\left(\frac{Q_a(Y)}{Q_0(Y)}\right)^\ell = \prod_{i=1}^n e^{-\ell a^2/2\sigma_i^2} e^{\ell a Y_i/\sigma_i^2}, \quad (160)$$

where (159) is the ratio of conditional densities particularized to the Gaussian channel: $Q_{Y_i|X=x} \sim \mathcal{N}(x, \sigma_i^2)$.

Because the Y_i are independent when conditioned on $X=a$, the expectation with respect to Q_a can be written as the product of expectations with respect to $Q_{Y_i|X=a}$:

$$E_{Q_a} \left(\frac{Q_a(Y)}{Q_0(Y)}\right)^\ell = \prod_{i=1}^n E_{Q_{Y_i|X=a}} \left(e^{-\ell a^2/2\sigma_i^2} e^{\ell a Y_i/\sigma_i^2}\right) \quad (161)$$

$$= \prod_{i=1}^n e^{-\ell a^2/2\sigma_i^2} E_{Q_{Y_i|X=a}} \left(e^{\ell a Y_i/\sigma_i^2}\right) \quad (162)$$

$$= \prod_{i=1}^n e^{-\ell a^2/2\sigma_i^2} M_{Q_{Y_i|X=a}} \left(\frac{\ell a}{\sigma_i^2}\right) \quad (163)$$

$$= \prod_{i=1}^n \exp \left\{ -\frac{\ell a^2}{2\sigma_i^2} + \frac{\ell a^2}{\sigma_i^2} + \frac{\ell^2 a^2}{2\sigma_i^2} \right\} \quad (164)$$

$$= \prod_{i=1}^n \exp \left\{ \frac{a^2}{\sigma_i^2} \frac{\ell(\ell+1)}{2} \right\} \quad (165)$$

$$= \exp \left\{ \frac{\ell(\ell+1)}{2} \sum_{i=1}^n \frac{a^2}{\sigma_i^2} \right\}, \quad (166)$$

where the individual expectations have been expressed in (163) in terms of the moment-generating function $M(t)$:

$$M_{Q_{Y_i|X=a}}(t) = E_{Q_{Y_i|X=a}}(e^{tY_i}) \quad (167)$$

$$= e^{at} e^{\sigma_i^2 t^2/2}. \quad (168)$$

Putting this together with (156), we have

$$\chi_k(Q_a||Q_0) = (-1)^{k+1}(k-1) + \sum_{\ell=2}^k \binom{k}{\ell} (-1)^{k-\ell} \exp \left\{ \frac{\ell(\ell-1)}{2} \sum_{i=1}^n \frac{a^2}{\sigma_i^2} \right\}. \quad (169)$$

Application of (169) to the distributions P_{Y_1, Y_2} , P_{Y_1} , and P_{Y_2} yields the following result:

$$\chi_k(P_{Y_1, Y_2|X=a} || P_{Y_1, Y_2|X=0}) \quad (170)$$

$$= \chi_k(P_{Y_1|X=a} || P_{Y_1|X=0}) - \chi_k(P_{Y_2|X=a} || P_{Y_2|X=0})$$

$$= (-1)^k (k-1) + \sum_{\ell=2}^k \binom{k}{\ell} (-1)^{k-\ell} (r_\ell - 1), \quad (171)$$

$$= \sum_{\ell=2}^k \binom{k}{\ell} (-1)^{k-\ell} r_\ell, \quad (172)$$

where in (171) we have defined

$$r_\ell = 1 + \exp \left\{ \frac{\ell(\ell-1)}{2} \left(\frac{a^2}{\sigma_1^2} + \frac{a^2}{\sigma_2^2} \right) \right\} \quad (173)$$

$$- \exp \left\{ \frac{\ell(\ell-1)}{2} \frac{a^2}{\sigma_1^2} \right\} - \exp \left\{ \frac{\ell(\ell-1)}{2} \frac{a^2}{\sigma_2^2} \right\} \\ = \prod_{i=1,2} \left(\exp \left\{ \frac{\ell(\ell-1)}{2} \frac{a^2}{\sigma_i^2} \right\} - 1 \right) \quad (174)$$

$$= \prod_{i=1,2} \left(\exp \{ c_{\ell,i} a^2 \} - 1 \right), \quad (175)$$

where in (175) we have defined $c_{\ell,i} = \frac{\ell(\ell-1)}{2\sigma_i^2}$.

Substituting (172) into (150) and using an arbitrary base for the logarithm, we have

$$I(Y_1; Y_2) = D(P_{Y_1, Y_2} \| P_{Y_1, Y_2 | X=0}) \quad (176)$$

$$- D(P_{Y_1} \| P_{Y_1 | X=0}) - D(P_{Y_2} \| P_{Y_2 | X=0}) \\ = (\log e) \sum_{k=2}^{\infty} \frac{(-1)^k \alpha^k}{k(k-1)} \sum_{\ell=2}^k \binom{k}{\ell} (-1)^{k-\ell} r_\ell, \quad (177)$$

$$= (\log e) \sum_{k=2}^{\infty} \sum_{\ell=2}^k \frac{(-1)^\ell \alpha^k}{k(k-1)} \binom{k}{\ell} r_\ell, \quad (178)$$

Fixing $\alpha \in (0, 1)$, we can take the limit as $\mathcal{P} \downarrow 0$ by letting $a \downarrow 0$. The first- and second-order terms can be expressed in terms of the following:

$$\dot{I}_{1,2}(0) = \lim_{a \downarrow 0} \frac{I(Y_1; Y_2)}{\alpha a^2} \quad (179)$$

$$\ddot{I}_{1,2}(0) = \lim_{a \downarrow 0} \frac{I(Y_1; Y_2) - \dot{I}_{1,2}(0) \alpha a^2}{\alpha^2 a^4}. \quad (180)$$

Introducing the index m to cover the cases that $m=1, 2$, we obtain the following limit:

$$\lim_{a \downarrow 0} \frac{I(Y_1; Y_2)}{\alpha^m a^{2m}} = \lim_{a \downarrow 0} \log e \sum_{k=2}^{\infty} \sum_{\ell=2}^k \frac{(-1)^\ell \alpha^{k-m}}{k(k-1) a^{2m}} \binom{k}{\ell} r_\ell \quad (181)$$

$$= \log e \lim_{a \downarrow 0} \sum_{k=2}^{\infty} \sum_{\ell=2}^k \frac{(-1)^\ell \alpha^{k-m}}{k(k-1)} \binom{k}{\ell} \prod_{i=1,2} \frac{e^{c_{\ell,i} a^2} - 1}{a^m} \quad (182)$$

$$= \log e \sum_{k=2}^{\infty} \sum_{\ell=2}^k \frac{(-1)^\ell \alpha^{k-m}}{k(k-1)} \binom{k}{\ell} \lim_{a \downarrow 0} \prod_{i=1,2} \frac{e^{c_{\ell,i} a^2} - 1}{a^m}, \quad (183)$$

$$= \log e \sum_{k=2}^{\infty} \sum_{\ell=2}^k \frac{(-1)^\ell \alpha^{k-m}}{k(k-1)} \binom{k}{\ell} \prod_{i=1,2} \lim_{a \downarrow 0} \frac{e^{c_{\ell,i} a^2} - 1}{a^m}, \quad (184)$$

where the individual limits for receiver i are

$$\lim_{a \downarrow 0} \frac{e^{c_{\ell,i} a^2} - 1}{a^m} = \begin{cases} 0, & m = 1 \\ c_{\ell,i}, & m = 2. \end{cases} \quad (185)$$

As a result, the first-order term (using $m=1$) is $\dot{I}_{1,2}(0)=0$, so we can ignore the $\dot{I}_{1,2}(0)$ term in (180). The second-order

term (using $m=2$) is

$$\frac{\ddot{I}_{1,2}(0)}{2 \log e} = \sum_{k=2}^{\infty} \sum_{\ell=2}^k \frac{(-1)^\ell \alpha^{k-2}}{k(k-1)} \binom{k}{\ell} \prod_{i=1,2} c_{\ell,i} \quad (186)$$

$$= \frac{1}{\sigma_1^2 \sigma_2^2} \sum_{k=2}^{\infty} \sum_{\ell=2}^k \frac{(-1)^\ell \alpha^{k-2}}{k(k-1)} \binom{k}{\ell} \binom{\ell}{2}^2 \quad (187)$$

$$= \frac{1}{\sigma_1^2 \sigma_2^2} \sum_{k=2}^4 \frac{(-1)^k \alpha^{k-2}}{k(k-1)} \sum_{\ell=2}^k (-1)^{k-\ell} \binom{k}{\ell} \binom{\ell}{2}^2 \quad (188)$$

$$= \frac{1}{\sigma_1^2 \sigma_2^2} \left(\frac{1}{2} - \frac{6\alpha}{6} + \frac{6\alpha^2}{12} \right) \quad (189)$$

$$= \frac{(1-\alpha)^2}{2\sigma_1^2 \sigma_2^2} \quad (190)$$

$$\ddot{I}_{1,2}(0) = \frac{(1-\alpha)^2}{\sigma_1^2 \sigma_2^2} \log e, \quad (191)$$

where (188) is due to the following fact, from [50, (3)]

$$\sum_{\ell=2}^k (-1)^{k-\ell} \binom{k}{\ell} \binom{\ell}{2}^2 = 0, \quad k > 4, \quad (192)$$

since $\binom{\ell}{2}^2$ is a polynomial in ℓ of degree 4.

It is also interesting to note that fixed- α OOK is not first-order optimal on the Gaussian noise channel. It can be shown using the above approach that $\dot{I}_i(0) = (1-\alpha) \frac{\log e}{2\sigma_i^2}$, which is strictly less than $\dot{C}_i(0) = \frac{\log e}{2\sigma_i^2}$, since $\alpha > 0$. Still, the first-order coefficient of the CMI is zero: $\dot{I}_{1,2}(0) = 0$. □

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Jul./Oct. 1948.
- [2] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.
- [3] —, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.
- [4] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*. Springer, 2017.
- [5] Z. M. Kassas, "Collaborative opportunistic navigation [student research highlight]," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 28, no. 6, pp. 38–41, June 2013.
- [6] Z. M. Kassas and T. E. Humphreys, "Observability analysis of collaborative opportunistic navigation with pseudorange measurements," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 260–273, Feb. 2014.
- [7] J. J. Morales and Z. M. Kassas, "Optimal collaborative mapping of terrestrial transmitters: Receiver placement and performance characterization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 2, pp. 992–1007, April 2018.
- [8] S. Stein, "Algorithms for ambiguity function processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pp. 588–599, June 1981.
- [9] L. Vasudevan, A. Ortega, and U. Mitra, "Application-specific compression for time delay estimation in sensor networks," in *Proc. 1st Int. Conf. Embedded Netw. Sensor Syst. (SenSys)*, Los Angeles, California, USA, Nov. 2003, pp. 243–254.
- [10] —, "Application-optimized quantizers for time delay estimation in sensor networks," in *Proc. 2003 IEEE Workshop Stat. Signal Process.*, 2003, pp. 9–12.
- [11] —, "Jointly optimized quantization and time delay estimation for sensor networks," in *1st Int. Symp. Control, Commun., Signal Proc.*, 2004, pp. 203–208.
- [12] T. Cole, "Finite sample correlations of quantized Gaussians," *Australian J. Physics*, vol. 21, no. 3, pp. 273–282, 1968.
- [13] A. R. Williamson, "The impacts of additive noise and 1-bit quantization on the correlation coefficient in the low-SNR regime," in *Proc. 57th Allerton Conf. on Commun., Control, and Computing*, Sep. 2019, pp. 631–638.
- [14] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [15] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327, Aug. 1976.
- [16] G. Carter, "Time delay estimation for passive sonar signal processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pp. 463–470, June 1981.
- [17] G. C. Carter, "Coherence and time delay estimation," *Proc. IEEE*, vol. 75, no. 2, pp. 236–255, Feb. 1987.
- [18] M. L. Fowler and X. Hu, "Signal models for TDOA/FDOA estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 4, pp. 1543–1550, Oct. 2008.
- [19] M. R. Bell, "Information theory and radar waveform design," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1578–1597, Sep. 1993.
- [20] R. Romero and N. A. Goodman, "Information-theoretic matched waveform in signal dependent interference," in *2008 IEEE Radar Conf.*, May 2008.
- [21] R. A. Romero, J. Bae, and N. A. Goodman, "Theory and application of SNR and mutual information matched illumination waveforms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 912–927, April 2011.
- [22] J. A. Malas and J. A. Cortese, "The radar information channel and system uncertainty," in *2010 IEEE Radar Conf.*, May 2010, pp. 144–149.
- [23] —, "The radar system and information flow," in *2nd Int. Workshop Cognitive Inf. Process.*, June 2010, pp. 17–22.

- [24] J. A. Malas, J. A. Cortese, and P. Ryan, "Uncertainty propagation and the fano based information theoretic method," in *2015 IEEE Radar Conf.*, May 2015, pp. 1638–1643.
- [25] B. Paul and D. W. Bliss, "Constant information radar for dynamic shared spectrum access," in *Proc. 49th Asilomar Conf. Signals, Syst., Comput.*, Nov 2015, pp. 1374–1378.
- [26] B. Paul and D. Bliss, "The constant information radar," *Entropy*, vol. 18, no. 9, p. 338, Sep. 2016.
- [27] A. R. Chiriyath, B. Paul, and D. W. Bliss, "Radar-communications convergence: Coexistence, cooperation, and co-design," *IEEE Trans. Cognitive Commun. Netw.*, vol. 3, no. 1, pp. 1–12, Mar. 2017.
- [28] J. Stein, J. Ziv, and N. Merhav, "Universal delay estimation for discrete channels," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 2085–2093, Nov 1996.
- [29] Z. Zhang and T. Berger, "Estimation via compressed information," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 198–211, Mar. 1988.
- [30] T. S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2300–2324, Oct. 1998.
- [31] —, "Parameter estimation with multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1802–1833, Nov. 1995.
- [32] V. V. Prelov, "Asymptotic behavior of the capacity of a continuous channel with a large amount of noise," *Probl. Pered. Inform.*, vol. 6, no. 2, pp. 40–57, 1970.
- [33] —, "Asymptotic behavior of the capacity of a continuous channel with large nonadditive noise," *Probl. Pered. Inform.*, vol. 8, no. 4, pp. 22–27, 1972.
- [34] I. A. Ibragimov and R. Z. Khas' minskii, "Weak signal transmission in a memoryless channel," *Probl. Pered. Inform.*, vol. 8, no. 4, pp. 28–39, 1972.
- [35] V. V. Prelov and E. C. van der Meulen, "An asymptotic expression for the information and capacity of a multidimensional channel with weak input signals," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1728–1735, Sep. 1993.
- [36] V. V. Prelov and S. Verdú, "Second-order asymptotics of mutual information," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1567–1580, Aug. 2004.
- [37] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.
- [38] M. El-Halabi, T. Liu, and C. N. Georghiades, "Secrecy capacity per unit cost," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1909–1920, Sep. 2013.
- [39] T. Koch and A. Lapidoth, "Increased capacity per unit-cost by oversampling," in *Proc. 26th IEEE Conv. Elect. Electron. Engineers Israel*, Eliat, Israel, Nov. 2010, pp. 684–688.
- [40] —, "Asymmetric quantizers are better at low SNR," in *Proc. 2011 IEEE Int. Symp. Inf. Theory (ISIT)*, St. Petersburg, Russia, July 2011, pp. 2592–2596.
- [41] —, "At low SNR, asymmetric quantizers are better," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5421–5445, Sep. 2013.
- [42] P. Zhang, F. M. J. Willems, and L. Huang, "Investigations of noncoherent OOK based schemes with soft and hard decisions for WSNs," in *49th Allerton Conf. on Commun., Control, and Comput.*, Sept. 2011, pp. 1702–1709.
- [43] J. H. Van Vleck and D. Middleton, "The spectrum of clipped noise," *Proc. IEEE*, vol. 54, no. 1, pp. 2–19, Jan. 1966.
- [44] C. N. Georghiades, "Modulation and coding for throughput-efficient optical systems," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1313–1326, Sept. 1994.
- [45] F. Xu, M. Khalighi, and S. Bourennane, "Coded PPM and multipulse PPM and iterative detection for free-space optical links," *IEEE/OSA J. Optical Commun. Netw.*, vol. 1, no. 5, pp. 404–415, Oct. 2009.
- [46] M. Noshad and M. Brandt-Pearce, "Expurgated PPM using symmetric balanced incomplete block designs," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 968–971, July 2012.
- [47] —, "Multilevel pulse-position modulation based on balanced incomplete block designs," in *Proc. 2012 IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2012, pp. 2930–2935.
- [48] Y. Fujiwara, "Self-synchronizing pulse position modulation with error tolerance," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5352–5362, Sept. 2013.
- [49] E. Weinstein and A. Weiss, "Fundamental limitations in passive time-delay estimation—Part II: Wide-band systems," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 5, pp. 1064–1078, Oct. 1984.
- [50] K. N. Boyadzhiev, "Close encounters with the Stirling numbers of the second kind," *Mathematics Magazine*, vol. 85, no. 4, pp. 252–266, 2012.