Integrated Communication and Binary State Detection from Hoeffding's Perspective

Daewon Seo
EECS Department
DGIST, South Korea
dwseo@dgist.ac.kr

Sung Hoon Lim
School of Information Sciences
Hallym University, South Korea
shlim@hallym.ac.kr

***Abstract*—This work considers a problem of integrated sensing and communication (ISAC) from an information-theoretic perspective, in which the goal of sensing is to detect a binary state. We assume a broadcast channel that consists of a transmitter, a communication receiver, and a binary detector where the detector's channel is affected by an unknown binary state. Unlike most approaches where the detection error probability is considered, in our work, we analyze and disaggregate the error probability into more detailed components, false alarm and missed detection error exponents. We fully characterize the optimal three-way tradeoff between the coding rate for communication and the two possibly nonidentical error exponents for sensing in the asymptotic regime. The achievability and converse proofs rely on the analysis of the cumulant-generating function of the log-likelihood ratio. We also provide numerical evaluations of our results for binary and Gaussian channels.**

***Index Terms*—Integrated sensing and communication, binary hypothesis testing, Hoeffding's problem, constant composition codes, cumulant-generating functions**

I. Introduction

The rapid evolution of wireless communication technologies has created an unprecedented demand for higher data throughput. To meet this need, future wireless communication systems are expected to utilize higher frequencies, such as millimeter waves, which also allow the transmitted signal to be used for radar due to their high resolution. In this context, integrated sensing and communication (ISAC) technology is predicted to be one of the key features of the sixth generation (6G) communication networks, seamlessly merging communication and sensing functionalities into a unified system [1], [2]. Unlike previous generations that treated these aspects separately, ISAC enables networks to simultaneously transmit data and sense the environment, enhancing situational awareness and resource efficiency [3], [4]. This convergence supports advanced applications like autonomous driving, surveillance, and smart cities.

A key aspect of ISAC from a signal design perspective is that a single transmitted signal must function both as an active sensing signal and a data-bearing signal. These possibly conflicting roles create a new tradeoff between data rate and sensing performance. For example, reference [5] characterized a tradeoff between data rate and the error exponent of sensing errors for a binary multiplicative state, which was later extended to a general binary state [6] and to multiple states [7]. This body of literature indeed focuses on detection errors for the most indistinguishable state, i.e., the worst-case error, leading to a minimax or equalizer design of ISAC. However, detection errors often need to be considered in more detail, as they encompass false alarms and missed detections, which should be treated unequally. For instance, consider a motivating scenario where an autonomous vehicle communicates with a roadside infrastructure while also detects pedestrians in its path by shared transmitted signals to ensure safe navigation. In this case, there are two types of detection errors [8], [9]: false alarm (also called type I or false positive error), which occur when the system alerts the presence of a pedestrian when there is none, and missed detection (also called type II or false negative error), which occur when the system fails to alert even though pedestrians are present. Clearly, the cost of a missed detection is significantly higher than that of a false alarm. A missed detection could lead to catastrophic consequences whereas a false alarm may only result in a slight reduction in efficiency or unnecessary caution. Despite the critical importance of this issue, the problem of unequal error protection has been only partially studied in the ISAC literature [6]. Addressing this gap is essential for the reliable deployment of ISAC systems in safety-critical applications.

This work explores the problem of the above unequal error protection in ISAC from an information-theoretic perspective. In particular, our main goal of this work is to fully characterize the three-way tradeoff between communication rate, false alarm, and missed detection error exponents in the asymptotic regime.

Information-theoretic approaches on dual-functional signal design for both data transmission and sensing were initially explored in [10]–[12], in the context of "state amplification." Here, the channel state varies in an independent and identically distributed (i.i.d.) manner and is assumed to be known to the encoder. In this framework, the encoder embeds the state information into the transmitted codewords, enabling the receiver to perform simultaneous data decoding and state estimation. Subsequently, [13] considered a scenario where the state information is unavailable at the encoder and showed

D. Seo was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2024-00398948). S. H. Lim was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology under Grant RS-2023-00278337.

ISBN 978-3-903176-65-2 © 2024 IFIP

the tradeoff between communication and sensing. In recent monostatic ISAC systems in which transmitters aim to detect the state from backscattered signals, adaptive signaling can be utilized, e.g., [14]. In the above studies, two features of channel capacity for data transmission and rate distortion for state detection arise together due to the i.i.d. variation of states, and hence, the performance is characterized by the capacity-distortion function.

Unlike the above, our work is motivated by physical changes in the sensing environment that occur much slower than the time scale of communication. We assume that an unknown state remains fixed throughout the communication, such as an obstacle disrupting the sensing path. In this context, a relevant work to our setting is [5], which focuses on a binary multiplicative state; for example, if there is a full blockage of the sensing channel, the detector receives only noise regardless of the transmitted signal. Another relevant work [7] extends it to multiple states. This later work is more general as the sensing channel is not necessarily multiplicative and includes adaptive transmission for monostatic ISAC systems. For these two works, the sensing performance is measured by the error exponent of the total error probability, which is in turn dominated by the worst state-conditional error exponent. Formally, let x^n denote the transmitted codeword of length n and define the set of states by $\mathcal{S}$. Then, the probability of sensing error is indeed dominated by

$$P_{\mathsf{sensing}}^{(n)} = \max_{s\in\mathcal{S}} \max_{x^n} \mathsf{P}[\hat{S} \neq s | X^n = x^n, S = s].$$

Hence, the goal in [5], [7] is to identify a minimax sensing rule for ISAC and evaluate the rate-exponent region when $n \to \infty$. Lastly, reference [6, Sec. V] investigates the rate-exponent region using the Neyman-Pearson formulation, where the exponent E_α is for the probability of missed detection while keeping the probability of false alarm below $\alpha \in (0, 1)$. In other words,

$$E_\alpha = \lim_{n\to\infty} -\frac{1}{n} \log \max_{x^n} \min_{\hat{S}} \mathsf{P}[\hat{S} = 0 | X^n = x^n, S = 1],$$

where the minimum is over all detection rules such that the probability of false alarm is less than α,

$$\mathsf{P}[\hat{S} = 1 | X^n = x^n, S = 0] \leq \alpha.$$

To the best of our knowledge, this is the first study to consider unequal error protection in the context of ISAC. However, since the false alarm constraint is always met by any nontrivial detection rule as long as $n \to \infty$, the resulting E_α is indeed independent of α; there is no nontrivial tradeoff between α and E_α. In contrast, our work focuses on the nontrivial tradeoff between the two exponents for the probabilities of false alarm and missed detection, respectively, as well as the data rate.

In stand-alone classic binary hypothesis testing problems with i.i.d. sensing observations, the optimal tradeoff between the two exponents for the probabilities of false alarm and missed detection is characterized by Hoeffding [15]. It can be also thought of as a generalization of Neyman-Pearson setting. Our work extends this result to the ISAC setting.

A notable departure from standard notations in this paper is that we use $D(p(Y)\|q(Y)) := D(p_Y\|q_Y)$ to clearly state KL divergences with distributions conditioned on a specific realization.

The remainder of this paper is organized as follows. In the next section, we formally define our problem, followed by some preliminaries in Sec. III. The main results are presented in Sec. IV, and some numerical evaluations are provided in Sec. V.

II. Problem Statement

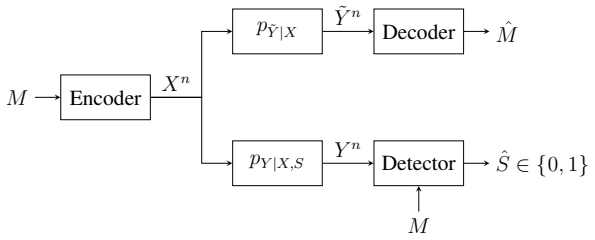

Fig. 1. Channel model for the joint communication and detection problem.

We consider a discrete memoryless state-dependent three-node network, depicted in Fig. 1,

$$p(\tilde{y}^n, y^n | x^n, s) = \prod_{i=1}^{n} p(\tilde{y}_i | x_i) p(y_i | x_i, s)$$

that consists of a transmitter, a receiver, and a detector, where $\tilde{y}^n \in \tilde{\mathcal{Y}}^n$ is the observation sequence at the receiver, $y^n \in \mathcal{Y}^n$ is the observation sequence at the detector, and $s \in \mathcal{S}$ is an unknown state that is fixed throughout the transmission. For simplicity, we will often denote $p_s(y|x) = p(y|x, s)$. This is the model for bistatic ISAC scenarios, but if the encoder and detector are equipped in a single device, it can be also thought of as a monostatic ISAC system that does not use feedback codes (e.g., open-loop coding [7]). In the latter case, $p_{Y|X,S}$ denotes the backscatter channel.

The transmitter wishes to send a message $m \in [1 : 2^{nR}]$ to the receiver while simultaneously enabling the detector to estimate the unknown state S. For the detector, we assume that the message m is given as side information. Formally, a $(2^{nR}, n)$ code for the joint communication and detection problem consists of

- a message set $m \in [1 : 2^{nR}]$,
- an encoder that assigns a sequence $x^n(m) \in \mathcal{X}^n$ to each message $m \in [1 : 2^{nR}]$,
- a decoder that assigns a message estimate $\hat{m}(\tilde{y}^n)$ to each observation sequence $\tilde{y}^n$, and
- a detector that assigns a state estimate $\hat{s}(y^n, x^n(m)) \in \mathcal{S}$ to each detector observation sequence y^n and message side information m.

We denote the codebook by $\mathcal{C}^{(n)} = \{x^n(m), m \in [1 : 2^{nR}]\}$ and suppose alphabet spaces are finite, i.e., $|\mathcal{X}|, |\mathcal{Y}|, |\tilde{\mathcal{Y}}| < \infty$.

The performance metrics are defined as follows. The probability of error for communication is defined as $P_\epsilon^{(n)} = \mathsf{P}(M \neq$

$\hat{M}$). We define two performance metrics for detection, the false alarm probability and the missed detection probability conditioned on a codeword $x^n(m)$ as

$$P_{\mathsf{FA}}^{(n)}(x^n) = \mathsf{P}[\hat{S}(Y^n, x^n) = 1 | X^n(M) = x^n, S = 0],$$
$$P_{\mathsf{MD}}^{(n)}(x^n) = \mathsf{P}[\hat{S}(Y^n, x^n) = 0 | X^n(M) = x^n, S = 1],$$

respectively.

We say that a rate and error exponent tuple $(R, E_{\mathsf{FA}}, E_{\mathsf{MD}})$ is achievable if there exists a sequence of $(2^{nR}, n)$ codes such that $\lim_{n\to\infty} P_\epsilon^{(n)} = 0$ and

$$\liminf_{n\to\infty} -\frac{1}{n} \log P_{\mathsf{FA}}^{(n)}(x^n) \geq E_{\mathsf{FA}},$$
$$\liminf_{n\to\infty} -\frac{1}{n} \log P_{\mathsf{MD}}^{(n)}(x^n) \geq E_{\mathsf{MD}},$$

for all codewords $x^n \in \mathcal{C}^{(n)}$ in the sequence of codebooks. The optimal region is the set of all achievable tuples. Then, the rate-exponent region $\mathcal{R}$ is the closure of the set of all achievable tuples $(R, E_{\mathsf{FA}}, E_{\mathsf{MD}})$, that is

$$\mathcal{R} := \mathrm{cl}\{(R, E_{\mathsf{FA}}, E_{\mathsf{MD}}) : (R, E_{\mathsf{FA}}, E_{\mathsf{MD}}) \text{ is achievable}\}.$$

Thus, our goal is to find the optimal tradeoff between the data rate R and error exponents $E_{\mathsf{FA}}, E_{\mathsf{MD}}$.

In this work, we are interested in the regime where both $E_{\mathsf{MD}} > 0$ and $E_{\mathsf{FA}} > 0$, since otherwise the problem becomes trivial. For instance, for the case $E_{\mathsf{FA}} = 0$, the detector can simply declare $\hat{S} = 1$ for all cases including $S = 0$. To further elaborate, we note that since $|\mathcal{X}|, |\mathcal{Y}| < \infty$, the log-likelihood ratio is bounded, i.e.,

$$\max_{x,y} \left| \log \frac{p_1(y|x)}{p_0(y|x)} \right| < \infty.$$

Hence, if the threshold of the log-likelihood ratio test is below a certain value (e.g., $\tau = -\infty$), it always declares $\hat{S} = 1$ independently of y^n. In this case, $P_{\mathsf{MD}} = 0$ and $E_{\mathsf{MD}} = \infty$; hence, $(E_{\mathsf{FA}}, E_{\mathsf{MD}}) = (0, \infty)$ is achievable. Therefore, we only consider the regime where $E_{\mathsf{FA}}, E_{\mathsf{MD}}$ are both strictly positive.

In the sequel, we assume the following necessary technical condition used in our converse proof. For any $x \in \mathcal{X}$, there exists a $C > 0$ such that

$$\mathrm{Var}\left(\log \frac{p_1(Y|x)}{p_0(Y|x)} \right) < C. \tag{1}$$

Note that the class of channels satisfying the condition includes discrete channels with finite alphabet spaces and an additive white Gaussian noise (AWGN) channel.

III. Preliminaries

In this section, we provide some preliminaries that is required throughout the paper. We denote the expectation over $p_s(Y|X)$ by $\mathsf{E}_s[\cdot]$ and hence, $\mathsf{E}_0, \mathsf{E}_1$ are the expectations over $p_0(y|x)$ and $p_0(y|x)$, respectively. Similarly, E_u is the expectation over p_u, which will be defined later. For simplicity, we define the likelihood ratios

$$\mathsf{LR}(y|x) := \frac{p_1(y|x)}{p_0(y|x)}, \quad \mathsf{LR}(y^n|x^n) := \frac{p_1(y^n|x^n)}{p_0(y^n|x^n)},$$

and the log-likelihood ratios by $\mathsf{LLR}(\cdot|\cdot) = \log \mathsf{LR}(\cdot|\cdot)$.

Define the normalized cumulant-generating function (CGF) of the n-length log-likelihood ratio as

$$\kappa_s(u|x^n) := \frac{1}{n} \log \mathsf{E}_s \left[\exp\left(u \cdot \mathsf{LLR}(Y^n|x^n) \right) \right],$$

for $u \in \mathbb{R}$ and $s \in \{0, 1\}$. Note that $\kappa_s(\cdot), s \in \{0, 1\}$ can be further simplified as

$$\begin{aligned}
\kappa_0(u|x^n) &:= \frac{1}{n} \log \mathsf{E}_0 \left[\exp\left(u \log \frac{p_1(Y^n|x^n)}{p_0(Y^n|x^n)} \right) \right] \\
&= \frac{1}{n} \log \mathsf{E}_0 \left[\frac{p_1^u(Y^n|x^n)}{p_0^u(Y^n|x^n)} \right] \\
&\overset{(a)}{=} \frac{1}{n} \log \prod_{i=1}^n \mathsf{E}_0 \left[\frac{p_1^u(Y_i|x_i)}{p_0^u(Y_i|x_i)} \right] \\
&\overset{(b)}{=} \frac{1}{n} \log \prod_{x\in\mathcal{X}} \left(\mathsf{E}_0 \left[\frac{p_1^u(Y|x)}{p_0^u(Y|x)} \right] \right)^{n_x} \\
&= \sum_{x\in\mathcal{X}} \frac{n_x}{n} \log \left(\mathsf{E}_0 \left[\frac{p_1^u(Y|x)}{p_0^u(Y|x)} \right] \right) \\
&\overset{(c)}{=} \sum_{x\in\mathcal{X}} p_X(x) \log \left(\mathsf{E}_0 \left[\frac{p_1^u(Y|x)}{p_0^u(Y|x)} \right] \right) \\
&= \sum_{x\in\mathcal{X}} p_X(x) \log \left(\sum_y p_0^{1-u}(y|x) p_1^u(y|x) \right),
\end{aligned}$$

where step (a) follows since the observations are conditionally independent, step (b) follows since n_x is the number of occurrence of x in x^n, and step (c) follows since p_X is the empirical pmf of x^n. Similarly,

$$\begin{aligned}
\kappa_1(v|x^n) &:= \frac{1}{n} \log \mathsf{E}_1 \left[\exp\left(v \log \frac{p_1(Y_1^n|x_1^n)}{p_0(Y_1^n|x_1^n)} \right) \right] \\
&= \sum_{x\in\mathcal{X}} p_X(x) \log \left(\mathsf{E}_1 \left[\frac{p_1^v(Y|x)}{p_0^v(Y|x)} \right] \right) \\
&= \sum_{x\in\mathcal{X}} p_X(x) \log \left(\sum_y p_0^{-v}(y|x) p_1^{v+1}(y|x) \right).
\end{aligned}$$

By observing the two definitions, we note that

$$\kappa_0(u|x^n) = \kappa_1(v|x^n), \quad \text{if } u = v + 1. \tag{2}$$

Property of κ_0: Consider a single-letter CGF

$$\kappa_0(u|x) := \log \mathsf{E}_0 \left[\exp\left(u \log \frac{p_1(Y|x)}{p_0(Y|x)} \right) \right].$$

Then, for every fixed x [9, Chap. 7],

1) $\kappa_0(u|x)$ is strictly convex in u if $p_0 \neq p_1$,
2) $\kappa_0(0|x) = \kappa_0(1|x) = 0$,
3) $\kappa_0'(0|x) = -D(p_0(Y|x) \| p_1(Y|x))$,
4) $\kappa_0'(1|x) = D(p_1(Y|x) \| p_0(Y|x))$,
5) $\kappa_0'(u|x)$ satisfies that for $0 \leq u \leq 1$,

$$-D(p_0(Y|x) \| p_1(Y|x)) \leq \kappa_0'(u|x) \leq D(p_1(Y|x) \| p_0(Y|x)).$$

That is, $\kappa_0(u|x)$ is a U-shape convex curve crossing zero at $u = 0, 1$. The normalized multi-letter CGF $\kappa_0(u|x^n)$ inherits all of the above properties, since it is a p_X-weighted linear combination of the single-letter CGF,

1) $\kappa_0(u|x^n)$ is strictly convex in u if $p_0 \neq p_1$,
2) $\kappa_0(0|x^n) = \kappa_0(1|x^n) = 0$,
3) $\kappa_0'(0|x^n) = -D(p_0(Y|x)\|p_1(Y|x)|p_X)$,
4) $\kappa_0'(1|x^n) = D(p_1(Y|x)\|p_0(Y|x)|p_X)$,
5) $\kappa_0'(u|x^n)$ satisfies that for $0 \leq u \leq 1$,

$$-D(p_0(Y|X)\|p_1(Y|X)|p_X) \leq \kappa_0'(u|x^n) \leq D(p_1(Y|X)\|p_0(Y|X)|p_X).$$

IV. Main Result

The following theorem characterizes the full tradeoff region between communication rate and error exponents.

Theorem 1. *The rate-exponent region $\mathcal{R}$ is the set of tuples $(R, E_{\mathsf{FA}}, E_{\mathsf{MD}})$ such that*

$$\begin{aligned} R &\leq I(X;\tilde{Y}), \\ E_{\mathsf{FA}} &\leq D(p_u(y|x)\|p_0(y|x)|p_X), \\ E_{\mathsf{MD}} &\leq D(p_u(y|x)\|p_1(y|x)|p_X), \end{aligned}$$

for some $p_X(x)$ and $u \in (0,1)$ where

$$p_u(y|x) = \frac{p_0^{1-u}(y|x)p_1^u(y|x)}{\sum_y p_0^{1-u}(y|x)p_1^u(y|x)}.$$

We note that $\mathcal{R}$ is parameterized by two parameters p_X and u. We provide some operational meaning on the role of u. The variable u can be considered as an intermediate parameter that controls the threshold of a log-likelihood ratio test which implies that u is controlled at the receiver (detector) side. Thus, for a fixed rate and code, the tradeoff point between the error exponents is determined at the receiver side. For example, assume that the detector receives a sequence of observations Y^n induced from a codeword. The detector can achieve multiple error exponent boundary points by varying the threshold of the log-likelihood test *on the single observation sequence*. This is somewhat different from typical rate regions in information theory where each operating point is obtained by a different codebook (determined at the encoder side).

A. Achievability Proof

For the achievability proof, we use constant composition codes (CCCs) originally developed in [16]. Since we rely on the results from [16] for the communication performance analysis, we will only provide a brief summary here, and most of the details will be devoted to the analysis of CCC on detection error exponents.

Fix an input type p_X of length n, and let $\mathcal{T}_{p_X}^n$ be the set of sequences in $\mathcal{X}^n$ with composition p_X. Then, by [16, Thm. 10.2], rate R is achievable by some constant composition code $\mathcal{C}^{(n)}$ consisting of codewords $x^n(m) \in \mathcal{T}_{p_X}^n$ if

$$R < I(X;\tilde{Y})$$

as $n \to \infty$.

Next, we analyze the performance of the CCC for detection error exponents. Note that for binary detection problems, it is optimal to perform a log-likelihood ratio test with some proper threshold $\tau \in \mathbb{R}$. We first consider the case when

$$-D(p_0(Y|X)\|p_1(Y|X)|p_X) < \tau < D(p_1(Y|X)\|p_0(Y|X)|p_X),$$

which leads us to a nontrivial tradeoff; other trivial cases will be discussed at the end of this proof.

Consider an arbitrary codeword $x^n \in \mathcal{C}^{(n)}$ from the CCC, and define a random variable indicated by the log-likelihood ratio test as

$$\hat{S} = \begin{cases} 0 & \text{if } \frac{1}{n}\mathsf{LLR}(Y^n|x^n) < \tau, \\ 1 & \text{if } \frac{1}{n}\mathsf{LLR}(Y^n|x^n) \geq \tau. \end{cases}$$

Then, the false alarm probability can be bounded by the Chernoff bound for some $u > 0$ as

$$\begin{aligned} P_{\mathsf{FA}}(x^n) &= \mathsf{P}\left(\frac{1}{n}\mathsf{LLR}(Y^n|x^n) \geq \tau \middle| S = 0\right) \\ &= \mathsf{P}\left(\exp\left(u\log\frac{p_1(Y^n|x^n)}{p_0(Y^n|x^n)}\right) \geq \exp(nu\tau) \middle| S = 0\right) \\ &\leq \frac{\mathsf{E}_0\left[\exp\left(u\log\frac{p_1(Y^n|x^n)}{p_0(Y^n|x^n)}\right)\right]}{\exp(nu\tau)} \\ &= \exp(-n(u\tau - \kappa_0(u|x^n))), \end{aligned} \tag{3}$$

where the last equality follows from the definition of κ_0. Similarly, for any $v < 0$,

$$\begin{aligned} P_{\mathsf{MD}}(x^n) &= \mathsf{P}\left(\frac{1}{n}\mathsf{LLR}(Y^n|x^n) < \tau \middle| S = 1\right) \\ &= \mathsf{P}\left(\exp\left(v\log\frac{p_1(Y^n|x^n)}{p_0(Y^n|x^n)}\right) > \exp(nv\tau) \middle| S = 1\right) \\ &\leq \frac{\mathsf{E}_1\left[\exp\left(v\log\frac{p_1(Y^n|x^n)}{p_0(Y^n|x^n)}\right)\right]}{\exp(nv\tau)} \\ &= \exp(-n(v\tau - \kappa_1(v|x^n))) \\ &= \exp(-n((u-1)\tau - \kappa_0(u|x^n))), \end{aligned} \tag{4}$$

where the last equality holds by choosing $v = u - 1$ and the relation given in (2).

Note that for some fixed τ, (3) and (4) are both optimized by the same u^* such that $\kappa_0'(u^*|x^n) = \tau$. To further manipulate expressions and obtain the claimed bounds, we use the (multi-letter) mixture distribution p_u in the following form:

$$p_u(y^n|x^n) = \frac{p_0^{1-u}(y^n|x^n)p_1^u(y^n|x^n)}{\sum_{y^n} p_0^{1-u}(y^n|x^n)p_1^u(y^n|x^n)} \tag{5}$$

$$= \frac{p_0(y^n|x^n)\mathsf{LR}^u(y^n|x^n)}{\mathsf{E}_0[\mathsf{LR}^u(Y^n|x^n)]} \tag{6}$$

$$= \frac{p_0(y^n|x^n)\mathsf{LR}^u(y^n|x^n)}{\exp(n\kappa_0(u|x^n))}, \tag{7}$$

for some $u \in (0,1)$. Then, we have from (7) that

$$\begin{aligned} &D(p_u(Y^n|x^n)\|p_0(Y^n|x^n)) \\ &= \mathsf{E}_u\left[\log\frac{p_u(Y^n|x^n)}{p_0(Y^n|x^n)}\right] \end{aligned}$$

$$= u\,\mathsf{E}_u[\log \mathsf{LR}(Y^n|x^n)] - n\kappa_0(u|x^n). \quad (8)$$

For the $\mathsf{E}_u[\log \mathsf{LR}(Y^n|x^n)]$ term, we can change the underlying probability measure using (6) as follows:

$$\begin{aligned}
&\mathsf{E}_u[\log \mathsf{LR}(Y^n|x^n)] \\
&= \sum_{y^n} p_u(y^n|x^n) \log \mathsf{LR}(y^n|x^n) \\
&= \sum_{y^n} \frac{p_0(y^n|x^n)\mathsf{LR}^u(y^n|x^n)}{\mathsf{E}_0[\mathsf{LR}^u(Y^n|x^n)]} \log \mathsf{LR}(y^n|x^n) \\
&= \frac{\mathsf{E}_0[\mathsf{LR}^u(Y^n|x^n) \log \mathsf{LR}(Y^n|x^n)]}{\mathsf{E}_0[\mathsf{LR}^u(Y^n|x^n)]} \\
&= \frac{d}{du} \log \mathsf{E}_0[\mathsf{LR}^u(Y^n|x^n)] \\
&= n\kappa_0'(u|x^n),
\end{aligned}$$

which gives the relation

$$\frac{1}{n} D(p_u(Y^n|x^n)\|p_0(Y^n|x^n)) = u\kappa_0'(u|x^n) - \kappa_0(u|x^n).$$

The left side can also be written in a single-letter form by using the additive property of the KL divergence under product distributions by

$$\begin{aligned}
&D\left(p_u(Y^n|x^n)\|p_0(Y^n|x^n)\right) \\
&= D\left(\prod_{i=1}^n p_u(Y_i|x_i) \middle\| \prod_{i=1}^n p_0(Y_i|x_i)\right) \\
&= \sum_{i=1}^n D(p_u(Y_i|x_i)\|p_0(Y_i|x_i)) \\
&= nD(p_u(Y|X)\|p_0(Y|X)|p_X),
\end{aligned}$$

which results in

$$D(p_u(Y|X)\|p_0(Y|X)|p_X) = u\kappa_0'(u|x^n) - \kappa_0(u|x^n).$$

Therefore, at the optimal u^*, i.e., $\tau = \kappa_0'(u^*|x^n)$, we have

$$\begin{aligned}
&D(p_{u^*}(Y|X)\|p_0(Y|X)|p_X) \\
&= u^*\kappa_0'(u^*|x^n) - \kappa_0(u^*|x^n) \quad (9) \\
&= \max_u (u\tau - \kappa_0(u|x^n)).
\end{aligned}$$

Since $\kappa_0(u|x^n)$ depends only on the empirical distribution of x^n, we can alternatively write κ_0 as a function of its type $\kappa_0(u|p_X)$ for any codeword in $\mathcal{C}^{(n)}$. In other words, we have a universal bound that holds for all $x^n \in \mathcal{C}^{(n)}$ as $\mathcal{C}^{(n)}$ is a CCC. Hence, (3) can be restated by, for any $x^n \in \mathcal{C}^{(n)}$

$$\begin{aligned}
P_{\mathsf{FA}}(x^n) &\le \exp(-n \cdot \max_{u\in(0,1)} (u\tau - \kappa_0(u|p_X))) \\
&= \exp(-nD(p_{u^*}(Y|X)\|p_0(Y|X)|p_X)).
\end{aligned}$$

Similarly, to bound the error exponent for $P_{\mathsf{MD}}(x^n)$, it also holds that

$$\begin{aligned}
&D(p_{u^*}(Y|X)\|p_1(Y|X)|p_X) \\
&= (u^*-1)\kappa_0'(u^*) - \kappa_0(u^*|x^n) \quad (10) \\
&= \max_u ((u-1)\tau - \kappa_0(u|x^n)).
\end{aligned}$$

Combined with (4), we have the bound

$$P_{\mathsf{MD}}(x^n) \le \exp(-nD(p_{u^*}(Y|X)\|p_1(Y|X)|p_X)),$$

for $x^n \in \mathcal{C}^{(n)}$.

From the strict convexity of κ_0, $\kappa_0'(u)$ is a monotonically increasing function which implies that τ and u^* are one-to-one. Therefore, for any τ such that

$$\tau \in (-D(p_0(Y|X)\|p_1(Y|X)|p_X), D(p_1(Y|X)\|p_0(Y|X)|p_X)),$$

there exists a u^* satisfying $\tau = \kappa_0'(u^*)$.

When τ does not belong to the above range, the problem becomes trivial. To see this, suppose that $\tau \le -D(p_0(Y|X)\|p_1(Y|X)|p_X)$. In this case, the false alarm probability converges to 1 since

$$\begin{aligned}
\mathsf{E}_0\left[\frac{1}{n}\mathsf{LLR}(Y^n|x^n)\right] &= \frac{1}{n}\mathsf{E}_0\left[\log\frac{p_1(Y^n|x^n)}{p_0(Y^n|x^n)}\right] \\
&= \frac{1}{n}\sum_{i=1}^n \mathsf{E}_0\left[\log\frac{p_1(Y_i|x_i)}{p_0(Y_i|x_i)}\right] \\
&= -\frac{1}{n}\sum_{i=1}^n D(p_0(Y_i|x_i)\|p_1(Y_i|x_i)) \\
&= -D(p_0(Y|X)\|p_1(Y|X)|p_X),
\end{aligned}$$

and thus, the log-likelihood ratio is greater than τ almost surely as $n \to \infty$, i.e., $P_{\mathsf{FA}} = 1$ and $E_{\mathsf{FA}} = 0$. Decreasing τ further decreases P_{MD} (i.e., increases E_{MD}) at no expense of E_{FA}; e.g., setting $\tau = -\infty$ achieves $(E_{\mathsf{FA}}, E_{\mathsf{MD}}) = (0, \infty)$. The case for $\tau \ge D(p_1(y|x)\|p_0(y|x)|p_X)$ can be addressed similarly.

Remark 1. *From* (9), (10)*, and* $\tau = \kappa_0'(u^*|x^n)$*, the threshold of the LRT* τ *can be obtained by*

$$\tau = D(p_{u^*}(Y|X)\|p_0(Y|X)|p_X) - D(p_{u^*}(Y|X)\|p_1(Y|X)|p_X).$$

Remark 2. *By setting* $\tau = 0$ *(i.e., both error events are equally penalized) specializing the exponent results for the binary multiplicative state channel recovers Thm. 1 in [5].*

B. Converse proof

Suppose that a rate-exponent tuple $(R, E_{\mathsf{FA}}, E_{\mathsf{MD}})$ is achievable by a codebook $\mathcal{C}^{(n)}$. That is, for any $\epsilon > 0$, there exists a length-n codebook $\mathcal{C}^{(n)}$ such that

$$\begin{aligned}
\frac{\log|\mathcal{C}^{(n)}|}{n} &\ge R - \epsilon \\
P_e^{(n)} &\le n \\
\max_{m\in[1:2^{nR}]} \mathsf{P}(\hat{S}(Y^n, m) = 1 | M = m, S = 0) &\le e^{-n(E_{\mathsf{FA}}-\epsilon)} \\
\max_{m\in[1:2^{nR}]} \mathsf{P}(\hat{S}(Y^n, m) = 0 | M = m, S = 1) &\le e^{-n(E_{\mathsf{MD}}-\epsilon)}
\end{aligned}$$

Note that there are exponentially many codewords in $\mathcal{C}^{(n)}$ while the number of types of length n is at most polynomial in n [16]. It implies that one can take a nonempty set of types $\mathcal{T}$ such that for all $p_X \in \mathcal{T}$, a subcodebook $\mathcal{C}_{p_X}^{(n)} := \{x^n(m) : \pi(x^n(m)) = p_X\} \subset \mathcal{C}^{(n)}$ such that

$$\frac{\log|\mathcal{C}_{p_X}^{(n)}|}{n} > R - \epsilon - \delta_n$$

for some $\delta_n > 0$ that tends to zero as $n \to \infty$ where $\pi(x^n)$ is the type of x^n.

In the sequel, we restrict our attention to $\mathcal{C}_{p_X}^{(n)}$. Let M' be a message taking values uniformly random from the codewords in $\mathcal{C}_{p_X}^{(n)}$ and define the chosen codeword by $\tilde{X}^n(m)$. Then, $\tilde{X}^n(m) \sim \tilde{p}^n(\tilde{x}^n)$ and $\tilde{X}_i \sim \tilde{p}(\tilde{x})$ where

$$\tilde{p}^n(x^n) = \frac{\mathbf{1}\{x^n \in \mathcal{C}_{p_X}^{(n)}\}}{|\mathcal{C}_{p_X}^{(n)}|},$$
$$\tilde{p}_i(x) = \frac{1}{|\mathcal{C}_{p_X}^{(n)}|} \sum_{x^n \in \mathcal{C}_{p_X}^{(n)}} \mathbf{1}\{x_i = x\},$$

that is, $\tilde{p}^n(\tilde{x}^n)$ is the uniform distribution over $\mathcal{C}_{p_X}^{(n)}$ and $\tilde{p}_i(\tilde{x}_i)$ is the i-th marginal distribution of symbols on $\mathcal{C}_{p_X}$. Also, let $\bar{p}(x)$ be the empirical distribution over the entire codebook $\mathcal{C}_{p_X}^{(n)}$ given by

$$\bar{p}(x) := \frac{1}{n} \sum_{i=1}^{n} \tilde{p}_i(x) = p_X(x),$$

where the last equality holds since all codewords in $\mathcal{C}_{p_X}^{(n)}$ are of type p_X.

By applying Fano's inequality and standard converse steps [17] to the sub-codebook $\mathcal{C}_{p_X}^{(n)}$,

$$\begin{aligned} n(R - \epsilon - \delta_n) &\le \log |\mathcal{C}_{p_X}^{(n)}| \\ &\le nI(X;\tilde{Y}) + n\epsilon_n, \end{aligned}$$

which gives the condition $R \le I(X;\tilde{Y}) + \epsilon + \delta_n + \epsilon_n$.

Next, we prove the bounds on the error exponents. Again, we first restrict our attention on the subcodebook $\mathcal{C}_{p_X}^{(n)}$. For binary hypothesis testing problems, it is well known that a log-likelihood test achieves the optimal tradeoff points of the error probabilities P_{FA} and P_{MD}. Moreover, by controlling the threshold parameter τ, we can choose the operating point on the optimal tradeoff curve. Thus, in the following, we bound the performance of the log-likelihood test on $\mathcal{C}_{p_X}^{(n)}$-codeword induced observations.

Define a binary test random variables T_s, $s \in \{0,1\}$,

$$T_s = \begin{cases} 0 & \text{if } \frac{1}{n}\mathsf{LLR}(Y^n|x^n) < \tau, \\ 1 & \text{if } \frac{1}{n}\mathsf{LLR}(Y^n|x^n) \ge \tau, \end{cases}$$

where τ is the threshold that achieves the error exponents E_{FA}, E_{MD}. The subscript index s indicates that the observation $Y^n|x^n$ is distributed as

$$Y^n|x^n \sim \begin{cases} p_0(y^n|x^n) \text{ for } T_0, \\ p_1(y^n|x^n) \text{ for } T_1. \end{cases}$$

Also, let u be such that

$$\mu_u := \mathsf{E}_u\left[\frac{1}{n}\mathsf{LLR}(Y^n|x^n)\right] = \tau, \tag{11}$$

and similarly for some $\delta > 0$, let v be such that

$$\mu_v = \mathsf{E}_v\left[\frac{1}{n}\mathsf{LLR}(Y^n|x^n)\right] = \tau + \delta,$$

where the underlying distribution p_v is the mixture distribution similarly defined in (5). Then, T_v is also similarly defined with respect to observations following $p_v(y^n|x^n)$.

With the above definitions,

$$\begin{aligned} &D(p_v(Y^n|x^n)\|p_0(Y^n|x^n)) \\ &\overset{(a)}{\ge} D(p(T_v|x^n)\|p(T_0|x^n)) \\ &= \mathsf{P}(T_v=0|x^n)\log\frac{\mathsf{P}(T_v=0|x^n)}{\mathsf{P}(T_0=0|x^n)} \\ &\qquad + \mathsf{P}(T_v=1|x^n)\log\frac{\mathsf{P}(T_v=1|x^n)}{\mathsf{P}(T_0=1|x^n)} \\ &= -H_2(T_v|x^n) - \mathsf{P}(T_v=0|x^n)\log\mathsf{P}(T_0=0|x^n) \\ &\qquad - \mathsf{P}(T_v=1|x^n)\log\mathsf{P}(T_0=1|x^n) \\ &\overset{(b)}{\ge} -1 - \mathsf{P}(T_v=1|x^n)\log\mathsf{P}(T_0=1|x^n), \end{aligned}$$

where (a) follows from the data processing inequality since the binary random variable T_s is a function of observations, and (b) follows since the binary entropy function is upper bounded by 1 and $\mathsf{P}(T_v=0|x^n)\log\mathsf{P}(T_0=0|x^n) < 0$. Rearranging terms and using the fact that $\mathsf{P}(T_v=1|x^n) \ge 1 - \frac{C}{n\delta^2}$ as shown in App. A, we have

$$\begin{aligned} &\log\mathsf{P}(T_0=1|x^n) \\ &\ge \frac{-1 - D(p_v(Y^n|x^n)\|p_0(Y^n|x^n))}{\mathsf{P}(T_v=1|x^n)} \\ &\ge \frac{-1 - D(p_v(Y^n|x^n)\|p_0(Y^n|x^n))}{1-\frac{C}{n\delta^2}} \\ &= (-1 - D(p_v(Y^n|x^n)\|p_0(Y^n|x^n)))(1+o(1)). \end{aligned}$$

Exponentiating both sides,

$$\begin{aligned} &P_{\mathsf{FA}}(x^n) = \mathsf{P}(T_0=1|x^n) \\ &\ge \exp\{-(1 + D(p_v(Y^n|x^n)\|p_0(Y^n|x^n)))(1+o(1))\}. \end{aligned}$$

Since the additivity of the KL divergence for product distributions gives us

$$D\left(p_v(Y^n|x^n)\|p_0(Y^n|x^n)\right) = nD(p_v(Y|X)\|p_0(Y|X)|p_X),$$

the above implies that for any $x^n \in \mathcal{C}_{p_X}^{(n)}$,

$$P_{\mathsf{FA}}(x^n) \ge \exp\left(-n(D(p_v(Y|X)\|p_0(Y|X)|p_X) + o(1))\right).$$

Finally, from two facts that $p_v \to p_u$ as $\delta \to 0$ and that $D(p_v(Y|X)\|p_0(Y|X)|p_X)$ is continuous in p_v,

$$P_{\mathsf{FA}}(x^n) \ge \exp\left(-n(D(p_u(Y|X)\|p_0(Y|X)|p_X) + o(1))\right)$$

holds as $\delta \to 0$.

To obtain the bound on P_{MD}, take v such that $\mu_v = \tau - \delta$. By repeating the same steps for $D(p_v(Y^n|x^n)\|p_1(Y^n|x^n))$ with $\delta \to 0$, we have

$$P_{\mathsf{MD}}(x^n) \ge \exp\left(-n(D(p_u(Y|X)\|p_1(Y|X)|p_X) + o(1))\right).$$

Since the above bounds hold for any $x^n \in \mathcal{C}_{p_X}^{(n)}$

$$P_{\mathsf{FA}} = \max_{m\in[1:M]} P_{\mathsf{FA}}(x^n(m))$$

$$
\begin{aligned}
&\geq \max_{m:x^n(m)\in\mathcal{C}_{p_X}^{(n)}} P_{\mathsf{FA}}(x^n(m)) \\
&\geq \exp\left(-n(D(p_u(Y|X)\|p_0(Y|X)|p_X)+o(1))\right), \\
P_{\mathsf{MD}} &= \max_{m\in[1:M]} P_{\mathsf{MD}}(x^n(m)) \\
&\geq \max_{m:x^n(m)\in\mathcal{C}_{p_X}^{(n)}} P_{\mathsf{MD}}(x^n(m)) \\
&\geq \exp\left(-n(D(p_u(Y|X)\|p_1(Y|X)|p_X)+o(1))\right).
\end{aligned}
$$

Finally, we have the corresponding bounds on the exponents

$$
\begin{aligned}
E_{\mathsf{FA}} &\leq D(p_u(Y|X)\|p_0(Y|X)|p_X), \\
E_{\mathsf{MD}} &\leq D(p_u(Y|X)\|p_1(Y|X)|p_X).
\end{aligned}
$$

V. Numerical Examples

A. Binary channel example

We begin with an example of a binary multiplicative state as described in [5]. The example can be considered as an abstract model for the cast that, if an obstacle is present, it completely disrupts the sensing path causing the detector to receive only noise. Specifically, we consider a state-dependent broadcast channel

$$
\begin{aligned}
\tilde{Y}_i &= x_i \oplus \tilde{Z}_i, \\
Y_i &= S\cdot x_i \oplus Z_i,
\end{aligned}
$$

where $\tilde{Z}_i \sim \text{Bern}(p)$ and $Z_i \sim \text{Bern}(q)$. In this setting, note that

$$
\begin{aligned}
p_u(1|0) &= \frac{p_0^{1-u}(1|0)p_1^u(1|0)}{p_0^{1-u}(0|0)p_1^u(0|0)+p_0^{1-u}(1|0)p_1^u(1|0)} \\
&= \frac{q^{1-u}q^u}{(1-q)^{1-u}(1-q)^u+q^{1-u}q^u} \\
&= q, \\
p_u(1|1) &= \frac{p_0^{1-u}(1|1)p_1^u(1|1)}{p_0^{1-u}(0|1)p_1^u(0|1)+p_0^{1-u}(1|1)p_1^u(1|1)} \\
&= \frac{q^{1-u}(1-q)^u}{(1-q)^{1-u}q^u+q^{1-u}(1-q)^u}.
\end{aligned}
$$

Then, letting $\alpha = p_X(1)$ for brevity and applying Thm. 1,

$$
\begin{aligned}
R &\leq I(X;\tilde{Y}) \\
&= H_2(\alpha * p) - H_2(p), \\
E_{\mathsf{FA}} &\leq D(p_u(y|x)\|p_1(y|x)|p_X) \\
&= (1-\alpha)D(p_u(y|0)\|p_1(y|0)) + \alpha D(p_u(y|1)\|p_1(y|1)) \\
&= \alpha D_2\left(\frac{q^{1-u}(1-q)^u}{(1-q)^{1-u}q^u+q^{1-u}(1-q)^u}\middle\|1-q\right), \\
E_{\mathsf{MD}} &\leq D(p_u(y|x)\|p_0(y|x)|p_X) \\
&= (1-\alpha)D(p_u(y|0)\|p_0(y|0)) + \alpha D(p_u(y|1)\|p_0(y|1)) \\
&= \alpha D_2\left(\frac{q^{1-u}(1-q)^u}{(1-q)^{1-u}q^u+q^{1-u}(1-q)^u}\middle\|q\right),
\end{aligned}
$$

where $\alpha * p = \alpha(1-p)+(1-\alpha)p$.

Note that for the special case with $u = 0.5$, it recovers the rate region given in [5] as follows.

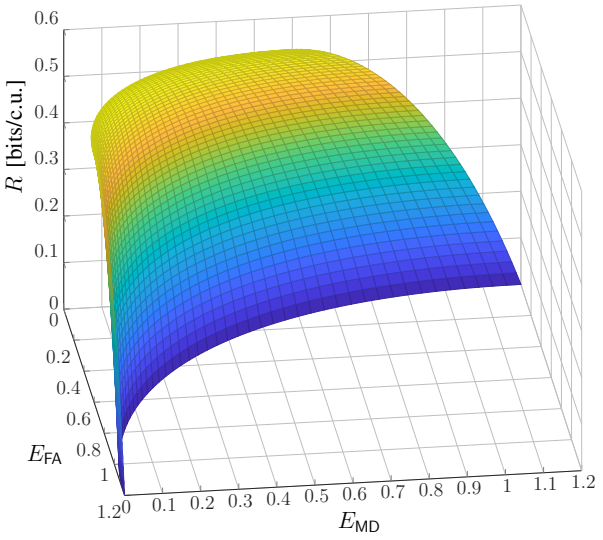

Fig. 2. Tradeoff boundary points of $\mathcal{R}$ for the binary multiplicative state channel with $p = 0.1, q = 0.2$. Units are bits.

Corollary 1 ([5, Thm. 1]). *For the binary multiplicative channel with $E = E_{\mathsf{FA}} = E_{\mathsf{MD}}$, the optimal trade-off region is given by the set of (R, E) pairs such that*

$$
\begin{aligned}
R &\leq H_2(\alpha * p) - H_2(p), \\
E &\leq \alpha D_2(0.5\|q).
\end{aligned}
$$

B. Gaussian example

Next, consider an AWGN communication channel

$$
\tilde{Y}_i = x_i + \tilde{Z}_i,
$$

and an AWGN sensing channel

$$
\begin{aligned}
Y_i &= x_i + Z_i, \text{ if } S = 0, \\
Y_i &= hx_i + Z_i, \text{ if } S = 1,
\end{aligned}
$$

where we assume an average input power constraint $\sum_{i=1}^n x_i^2 \leq nP$, $\tilde{Z}_i \sim \mathcal{N}(0,1)$, $Z_i \sim \mathcal{N}(0,1)$, and $h \neq 1$. This model can be considered as an abstract model for obstacle detection in which an obstacle alters the detector's wireless propagation path.

Assuming $X \sim \mathcal{N}(0, P)$, note that

$$
R \leq \frac{1}{2}\log(1+P).
$$

To evaluate the exponents, first note that

$$
\begin{aligned}
p_u(y|x) &= \frac{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(1-u)(y-x)^2}{2}\right)\exp\left(-\frac{u(y-hx)^2}{2}\right)}{\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(1-u)(y-x)^2}{2}\right)\exp\left(-\frac{u(y-hx)^2}{2}\right)dy} \\
&= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(y-(1-u+hu)x)^2}{2}\right) \\
&= \mathcal{N}((1-u+hu)x, 1).
\end{aligned}
$$

With the distribution at hand, we have

$$
D(p_u(y|X=x)\|p_1(y|X=x))
$$

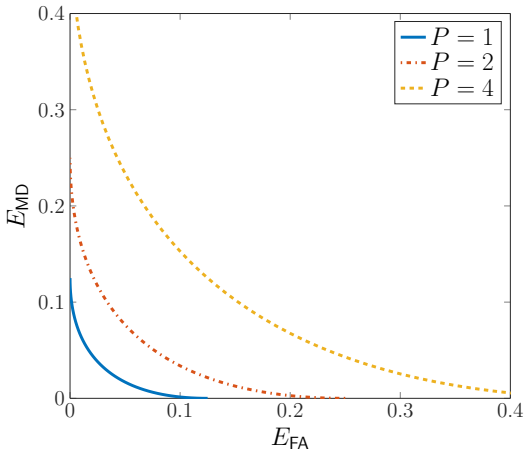

Fig. 3. Projection of $\mathcal{R}$ on the E_{MD}-E_{FA} plane for the Gaussian example with $P = 1, P = 2$ and $P = 4$. We set $h = 0.5$.

$$\begin{aligned} &= D(\mathcal{N}((1-u+hu)x, 1)\|\mathcal{N}(hx, 1)) \\ &= \frac{(1-u)^2(1-h)^2x^2}{2}, \end{aligned} \tag{12}$$

$$\begin{aligned} &D(p_u(y|X=x)\|p_0(y|X=x)) \\ &= D(\mathcal{N}((1-u+hu)x, 1)\|\mathcal{N}(x, 1)) \\ &= \frac{u^2(1-h)^2x^2}{2}. \end{aligned} \tag{13}$$

Then, we can evaluate Thm. 1 under the Gaussian input distribution $X \sim \mathcal{N}(0, P)$ which gives the following:

$$\begin{aligned} R &\le \frac{1}{2}\log(1+P), \\ E_{\mathsf{FA}} &\le D(p_u(Y|X)\|p_1(Y|X)|p_X) = \frac{(1-u)^2(1-h)^2P}{2}, \\ E_{\mathsf{MD}} &\le D(p_u(Y|X)\|p_0(Y|X)|p_X) = \frac{u^2(1-h)^2P}{2}. \end{aligned} \tag{14}$$

We note that the region (14) is in fact $\mathcal{R}$, i.e., the Gaussian distribution achieves the largest region over all possible input distributions. To see this, note that for each x, (12) and (13) hold. That means, for any distribution satisfying the constraint $\mathsf{E}(X^2) \le P$, the exponents are upper bounded by

$$\begin{aligned} D(p_u(Y|X)\|p_1(Y|X)|p_X) &\le \frac{(1-u)^2(1-h)^2P}{2}, \\ D(p_u(Y|X)\|p_0(Y|X)|p_X) &\le \frac{u^2(1-h)^2P}{2}. \end{aligned}$$

Since Gaussian distribution achieves the equalities and maximizes the mutual information, it attains $\mathcal{R}$ in Thm. 1.

VI. Conclusion

In this paper, we characterize the asymptotically optimal tradeoff between communication rate, false alarm, and missed detection error exponents for a fixed sensing state. By addressing the unequal error protection problem, our work introduces Hoeffding's hypothesis testing problem [15] to ISAC and extends existing literature on equal error protection [5], [7]. Our findings provide insights into how to separately control the reliability of the two error events simultaneously with an information-bearing signal.

Appendix A
Bound on $\mathsf{P}(T_v = 1|x^n)$

Recall that $\tau = \mathsf{E}_v\left[\frac{1}{n}\mathsf{LLR}(Y^n|x^n)\right] - \delta$ where $\delta > 0$. Then,

$$\begin{aligned} \mathsf{P}(T_v = 0|x^n) &= \sum_{y^n:\mathsf{LLR}(y^n|x^n)<n\tau} p_v(y^n|x^n) \\ &= \sum_{y^n:\mathsf{LLR}(y^n|x^n)-n\mu_v<-n\delta} p_v(y^n|x^n) \\ &\le \sum_{y^n:|\mathsf{LLR}(y^n|x^n)-n\mu_v|>n\delta} p_v(y^n|x^n) \\ &\overset{(a)}{\le} \frac{\mathsf{Var}(\mathsf{LLR}(Y^n|x^n))}{n^2\delta^2} \\ &\overset{(b)}{\le} \frac{C}{n\delta^2}, \end{aligned}$$

where (a) follows from the Chebyshev inequality, and (b) follows from our assumption (1). Since $\mathsf{P}(T_v = 0|x^n) = 1 - \mathsf{P}(T_v = 1|x^n)$, it holds that $\mathsf{P}(T_v = 1|x^n) \ge 1 - \frac{C}{n\delta^2}$.

References

[1] A. Bourdoux *et al.*, "6G white paper on localization and sensing," 2020, preprint available at https://arxiv.org/abs/2006.01779.

[2] A. Liu *et al.*, "A survey on fundamental limits of integrated sensing and communication," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 994–1034, 2022.

[3] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, Jun. 2020.

[4] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1236–1259, Jul. 2011.

[5] H. Joudeh and F. M. J. Willems, "Joint communication and binary state detection," *IEEE J. Sel. Areas Inf. Theory*, vol. 3, no. 1, pp. 113–124, Mar. 2022.

[6] H. Wu and H. Joudeh, "On joint communication and channel discrimination," in *Proc. 2022 IEEE Int. Symp. Inf. Theory*, June-July 2022, pp. 3321–3326.

[7] M.-C. Chang, S.-Y. Wang, T. Erdoğan, and M. R. Bloch, "Rate and detection-error exponent tradeoff for joint communication and sensing of fixed channel states," *IEEE J. Sel. Areas Inf. Theory*, vol. 4, pp. 245–259, May 2023.

[8] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, USA: Springer-Verlag, 1994.

[9] P. Moulin and V. Veeravalli, *Statistical Inference for Engineers and Data Scientists*. Cambridge, U.K.: Cambridge University Press, 2019.

[10] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1486–1495, Apr. 2005.

[11] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.

[12] C. Choudhuri, Y.-H. Kim, and U. Mitra, "Causal state amplification," in *Proc. 2011 IEEE Int. Symp. Inf. Theory*, July-Aug. 2011, pp. 2110–2114.

[13] W. Zhang, S. Vedantam, and U. Mitra, "Joint transmission and state estimation: A constrained channel coding approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 7084–7095, Oct. 2011.

[14] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An information-theoretic approach to joint sensing and communication," *IEEE Trans. Inf. Theory*, vol. 70, no. 2, pp. 1124–1146, Feb. 2024.

[15] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions," *Ann. Math. Stat.*, vol. 36, no. 2, pp. 369–401, Apr. 1965.

[16] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[17] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.