Rate and Detection Error-Exponent Tradeoffs of Joint Communication and Sensing

Meng-Che Chang, Tuna Erdoğan, Shi-Yuan Wang, Matthieu R. Bloch
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332
Email: {mchang301,etuna3,shi-yuan.wang,matthieu}@gatech.edu

***Abstract*—We consider a communication model in which a transmitter attempts to communicate with a receiver over a state-dependent channel and simultaneously estimate the state using strictly causal noisy state observations. Motivated by joint communication and sensing scenarios in which the physical phenomenon of interest for sensing evolves at a much slower rate than the rate of communication, the state is assumed to remain constant over the duration of the transmission. We derive a complete characterization of the optimal asymptotic trade-off between communication rate and detection-error exponent when coding strategies are open loop. We also show that closed-loop strategies result in strict improvements of the trade-offs.**

I. Introduction

The ability to jointly communicate and sense the wireless environment is one of the key features currently explored for 6G communication systems [1], [2]. While communication and sensing technologies have traditionally been independently designed and deployed, the push towards mmWave systems whose bandwidth can support high-resolution sensing is enabling a convergence that promises multiple benefits including: efficient systems with a single hardware for both communication and sensing; increased spectral efficiency; new applications that leverage accurate localization and sensing. The challenges posed by joint communication and sensing systems are therefore multi-fold, from designing efficient hardware with multi-functional antenna arrays, to designing dedicated signal processing algorithms, and has attracted much interest in the context of joint communication and radar [3]–[5].

Of particular relevance to the present work, several works have analyzed the information-theoretic limits of joint communication and sensing [3], [6]–[9] by drawing parallels with joint communication and state estimation [10], [11]. These results formulate the information-theoretic limits of joint communication and sensing as the characterization of a capacity-distortion region for state-dependent channel models in which an independent and identically distributed (i.i.d.) state, which represents the quantity to sense, governs the behavior channel [6], [9]. Because models are set up to preclude any prediction, detection strategies are open-loop and the interplay between communication and sensing reduces to a resource allocation problem, in which the choice of a channel input distribution dictates the trade-off.

This work was supported in part by NSF grants 1955401 and NSF 1910859.

978-1-6654-0579-9/22/$31.00 ©2022 IEEE

Motivated by the observation that the physical phenomena sensed, e.g., the presence or absence of an obstacle, might change on a time scale much slower than the rate of communication, we study a different model in which the state of the channel remains constant over the block-length used for communication. In this setup, estimation is generally possible so that a trade-off exists between the rate of communication and the detection-error exponent. Accordingly, our results draw on the extensive literature around controlled sensing [12], [13] and channel estimation with pilot sequences [14], [15] rather than the rate-distortion literature.

The main contribution of the present work is an exact characterization of the optimal asymptotic trade-off between communication rate and detection-error exponent when coding strategies are open loop, i.e., the transmitter does not exploit the state observation feedback for adaptation. Perhaps unsurprisingly, we also show that closed-loop strategies result in strict improvements of the trade-offs. The remainder of the paper is organized as follows. After a brief review of notation in Section II and a formal introduction of the model studied in Section III, we present and derive our main result in Section IV. We illustrate our main result numerically in Section V.

II. Notation

For any discrete set $\mathcal{X}$, $\mathcal{P}_{\mathcal{X}}$ is the set of all probability distributions on $\mathcal{X}$. For $n \in \mathbb{N}^*$, a sequence of length n is implicitly denoted $\mathbf{x} \triangleq (x_1, \cdots, x_n) \in \mathcal{X}^n$, while $x^i \triangleq (x_1, \cdots, x_i) \in \mathcal{X}^i$ denotes a sequence of length i. For $\mathbf{x} \in \mathcal{X}^n$, $\hat{p}_{\mathbf{x}}$ denotes the type of $\mathbf{x}$, i.e., $\hat{p}_{\mathbf{x}}(x) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{x_i = x\}$. For any type P, $\mathcal{T}_P^n$ is the corresponding type class, i.e., the set of all sequence $\mathbf{x} \in \mathcal{X}^n$ such $\hat{p}_{\mathbf{x}} = P$. Finally, $\mathcal{P}_{\mathcal{X},n}$ is the set of all possible types of length n sequence on $\mathcal{X}^n$. If $W_{Y|X}$ is a conditional distribution on $Y \in \mathcal{Y}$ given $X \in \mathcal{X}$, we denote $\mathbb{I}(P_X, W_{Y|X})$ the mutual information between X and Y when $X \sim P_X$ and $Y \sim \sum_x P_X(x) W_{Y|x}$.

III. Joint communication and sensing model

We consider the communication model illustrated in Fig. 1, in which a transmitter attempts to communicate with a receiver over a state-dependent Discrete Memoryless Channel (DMC), often referred to as a compound channel in the literature, while simultaneously probing the channel state in a strictly causal manner through a sensing channel. Specifically, the transmitter encodes a uniformly distributed message $W \in [1; M]$ into a

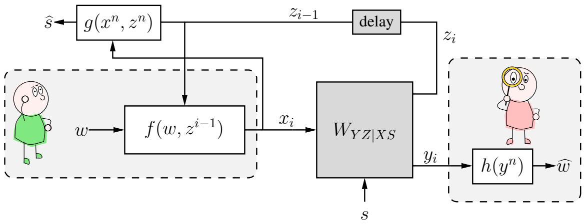

Fig. 1. Joint communication and sensing model.

length n codeword X^n, which symbols are transmitted over a DMC with transition probability $W_{YZ|XS}$. The a priori unknown state S is assumed to be *fixed* during the whole duration of the transmission and takes value in a finite set $\mathcal{S}$. The transmitter has the ability to estimate the channel state by using past observations obtained from the output Z of the DMC, allowing it to adapt its transmission in an online fashion. Formally, the encoder is a set of functions

$$f_i : [1; M] \times \mathcal{Z}^{i-1} \to \mathcal{X} : (w, z^{i-1}) \mapsto x_i \triangleq f_i(w, z^{i-1})$$

define for every $i \in [1; n]$ while the estimator is a function

$$g : \mathcal{X}^n \times \mathcal{Z}^n \to \mathcal{S} : (x^n, z^n) \to \hat{s}.$$

The decoder is a function

$$h : \mathcal{Y}^n \to [1; M] : y^n \mapsto \hat{w}.$$

A code $\mathcal{C}$ then consists of the tuple $(\{f_i\}_{i\in[1;n]}, g, h)$, as well as the implicitly defined associated message set $[1; M]$.

The performance of the system is measured in terms of the asymptotic rate of reliable communication and asymptotic detection-error exponent. Formally, we define the communication-error probability and the detection-error probability as follows

$$P_{\text{c}}^{(n)} \triangleq \max_{s\in\mathcal{S}} \max_{w\in[1;M]} \mathbb{P}(h(Y^n) \neq w | W = w, S = s), \quad (1)$$

$$P_{\text{d}}^{(n)} \triangleq \max_{s\in\mathcal{S}} \frac{1}{M} \sum_{w=1}^{M} \mathbb{P}(g(Z^n) \neq s | S = s, W = w). \quad (2)$$

The rate and detection-error exponent are

$$R \triangleq \frac{1}{n} \log M \text{ and } E_{\text{d}}^{(n)} \triangleq -\frac{1}{n} \log P_{\text{d}}^{(n)},$$

respectively.

Definition 1 (Achievability). *A rate/detection-error exponent (R, E) is achievable if for any $\epsilon > 0$, there exist a sufficiently large n and a code $\mathcal{C}$ of length n such that*

$$P_{\text{c}}^{(n)} \leqslant \epsilon, \quad (3)$$

$$E_{\text{d}}^{(n)} \geqslant E - \epsilon, \quad (4)$$

$$\frac{1}{n} \log |\mathcal{C}| \geqslant R - \epsilon. \quad (5)$$

Our objective is to characterize the set of all achievable rate/detection-error exponent pairs, which we call with a slight abuse of terminology the joint communication and sensing capacity region.

Since our model allows for adaptive coding schemes that use the feedback, we also define state-dependent error probabilities

$$P_{\text{c},s}^{(n)} \triangleq \max_{w} \mathbb{P}(h(Y^n) \neq w | W = w, S = s), \quad (6)$$

$$P_{\text{d},s}^{(n)} \triangleq \frac{1}{M} \sum_{w=1}^{M} \mathbb{P}(g(Z^n) \neq s | S = s, W = w). \quad (7)$$

The state-dependent detection-error exponent is defined as

$$E_{\text{d},s}^{(n)} \triangleq -\frac{1}{n} \log P_{\text{d},s}^{(n)}. \quad (8)$$

Definition 2 (s-Achievability). *Given the state $S = s$, a rate/detection-error exponent (R, E) is achievable if for any $\epsilon > 0$, there exists a sufficiently large n and a code $\mathcal{C}$ of length n such that*

$$P_{\text{c},s}^{(n)} \leqslant \epsilon, \quad (9)$$

$$E_{\text{d},s}^{(n)} \geqslant E - \epsilon, \quad (10)$$

$$\frac{1}{n} \log |\mathcal{C}| \geqslant R - \epsilon. \quad (11)$$

The notion of s-achievability is required to characterize coding schemes that learn the state s online. The non-ergodic nature of the state would otherwise prevent us from properly defining a notion of achievability without considering again a worst case scenario.

Remark 1. *There is an asymmetry in our definitions of the probability of errors in* (1) *and* (2)*. While* (1) *includes a maximum over a possible messages $w \in [1; M]$,* (2) *includes an average over all possible messages. This subtlety is only used in our converse proof for Theorem 3 in Section IV-B and is required to avoid having the detection performance dictated by a codeword whose type is not representative of the code.*

Remark 2. *The model of Fig. 1 differs from the ones in [6], [9], where the state is i.i.d. and changing from symbol to symbol. Our model captures a scenario in which the coherence time of the state is much longer than the duration of a transmission. As a result our model also captures the ability to adapt to the channel state in an online fashion, while the models in [6], [9] only allow for an offline adaptation based on a target rate/distortion pair. Neither model supersedes the other and both capture scenarios relevant to next generation communication networks.*

IV. Main result

We first restrict ourselves to the situation in which the encoder does not perform any online adaptation so that $f_i : [1; M] \mapsto \mathcal{X}$ is independent of the observation z^{i-1}. This corresponds to an *open-loop* strategy, which provides a baseline for assessing the usefulness of adaptation. For simplicity, we denote in this case the encoder that maps a message w to a codeword of n symbols by $f : [1; M] \mapsto \mathcal{X}^n$. We call such codes *open-loop schemes* and we denote the set of all achievable rate/detection-error exponent with open-loop schemes by $\mathcal{C}_{\text{open}}$. The following theorem provides an exact characterization of $\mathcal{C}_{\text{open}}$.

Theorem 3. *The joint communication and sensing capacity region of open-loop schemes is*

$$\mathcal{C}_{\text{open}} = \bigcup_{P_X \in \mathcal{P}_X} \left\{ \begin{array}{l} (R,E) \in \mathbb{R}_+^2 : \\ R \leqslant \min_{s\in\mathcal{S}} \mathbb{I}(P_X, W_{Y|X,s}) \\ E \leqslant \phi(P_X) \end{array} \right\} \quad (12)$$

where

$$\phi(P_X) = \min_{s\in\mathcal{S}} \min_{s'\neq s} \max_{\ell\in[0,1]} -\sum_x P_X(x) \times \log\left(\sum_z W_{Z|x,s}(z)^\ell W_{Z|x,s'}(z)^{1-\ell}\right). \quad (13)$$

Proof. See Section IV-A and Section IV-B. □

A couple of comments are in order. First, since open-loop schemes do not exploit the information about the state contained in past noisy observations of the state, achievable rates are necessarily upper bounded by the compound channel capacity $\max_{P_X} \min_{s\in\mathcal{S}} \mathbb{I}(P_X, W_{Y|X,s})$. This is a weakness of all open-loop schemes. Second, because of the open-loop nature of the coding schemes, the interplay between communication and sensing is captured by the choice of a distribution P_X that governs the empirical statistics of the codewords and is set *offline*. This is similar to what is obtained in other information-theoretic approaches based on rate-distortion [6], [9].

Corollary 4. *If there exists $x_0 \in \mathcal{X}$ such that for all $x \in \mathcal{X}$ there exists a permutation π_x on $\mathcal{Z}$ such that for every $s \in \mathcal{S}$*

$$W_{Z|X,s}(z|x) = W_{Z|X,s}(\pi_x(z)|x_0), \quad (14)$$

then

$$\mathcal{C}_{\text{open}} = \left\{ \begin{array}{l} (R,E) \in \mathbb{R}_+^2 : \\ R \leqslant \max_{P_X} \min_{s\in\mathcal{S}} \mathbb{I}(P_X, W_{Y|X,s}) \\ E \leqslant \max_{P_X} \phi(P_X) \end{array} \right\} \quad (15)$$

where $\phi(\cdot)$ is defined in (13).

In other words, there is no trade-off between rate and detection-error exponent in this case and one simultaneously achieves the optimal communication rate and the optimal detection performance.

Proof. For every $x \in \mathcal{X}\backslash\{x_0\}$,

$$\sum_{z\in\mathcal{Z}} W_{Z|X,s}(z|x)^\ell W_{Z|X,s'}(z|x)^{1-\ell} \quad (16)$$

$$= \sum_{z\in\mathcal{Z}} W_{Z|X,s}(\pi_x(z)|x_0)^\ell W_{Z|X,s'}(\pi_x(z)|x_0)^{1-\ell} \quad (17)$$

$$= \sum_{\pi_x^{-1}(z')\in\mathcal{Z}} W_{Z|X,s}(z'|x_0)^\ell W_{Z|X,s'}(z'|x_0)^{1-\ell} \quad (18)$$

$$= \sum_{z\in\mathcal{Z}} W_{Z|X,s}(z|x_0)^\ell W_{Z|X,s'}(z|x_0)^{1-\ell}. \quad (19)$$

Thus, we know that the detection-error exponent is invariant to the input type under this scenario. □

One of the compound channel families that falls into such a category is the set of Binary Symmetric Channels (BSCs). The maximal detection-error exponent and the compound capacity are then simultaneously achieved with a uniform input distribution.

The following theorem shows that *closed-loop schemes*, which exploit the feedback to adapt to the state, provide immediate improvements. Specifically, the theorem characterizes an inner bound of the set $\mathcal{C}^s_{\text{closed}}$ of s-achievable rate/detection-error exponent pairs.

Theorem 5. *Given each state $s \in \mathcal{S}$, the state-dependent joint communication and sensing capacity region satisfies*

$$\mathcal{C}^s_{closed} \supseteq \bigcup_{P_X \in \mathcal{P}_X} \left\{ \begin{array}{l} (R,E) \in \mathbb{R}_+^2 : \\ R \leqslant \mathbb{I}(P_X, W_{Y|X,s}) \\ E \leqslant \psi_s(P_X) \end{array} \right\}, \quad (20)$$

where

$$\psi_s(P_X) = \min_{s'\neq s} \max_{\ell\in[0,1]} -\sum_x P_X(x) \times \log\left(\sum_z W_{Z|X,s}(z|x)^\ell W_{Z|X,s'}(z|x)^{1-\ell}\right). \quad (21)$$

Theorem 5 is obtained by considering a simple strategy in which the transmitter learns the state, informs the receiver, and uses a code adapted to the learned channel state. The exact characterization of the optimal tradeoffs for closed-loop schemes is left of future work and presents non-trivial challenges, chief among them the absence of a known optimal detection error-exponent for multi-hypothesis controlled sensing [12].

A. Achievability Proof of Theorem 3

We show that all (R, E) pairs within the region C_{open} are achievable. Since we restrict ourselves to open-loop schemes, we may fix P_X as the type of all codewords. Fix any $\epsilon > 0$. By [16, Theorem 10.2], there exists a code with encoder f and decoder h such that $f(w) \in \mathcal{T}^n_{P_X}$, the rate is at least $\min_{s\in\mathcal{S}} \mathbb{I}(P_X, W_{Y|X,s}) - \epsilon$, and $\max_w \mathbb{P}(h(Y^n) \neq w|S = s) < \epsilon$ for all $s \in \mathcal{S}$. Then, the detection-error exponent $\phi(P_X)$ is given by the Lemma 6 modified from [12, Theorem 1].

Lemma 6. *Suppose that the the codeword corresponding to the message $w \in [1; M]$ has type $P_X \in \mathcal{P}_X$, the conditional detection-error exponent $E_{\text{d},w} \triangleq -\frac{1}{n}\log\max_{s\in\mathcal{S}} \mathbb{P}(g(Z^n) \neq s|S = s, W = w)$ in an open loop scheme is asymptotically upper bounded by*

$$\phi(P_X) \triangleq \min_{s} \min_{s'\neq s} \max_{\ell\in[0,1]} -\sum_x P_X(x) \times \log\left(\sum_z W_{Z|X,s}(z|x)^\ell W_{Z|X,s'}(z|x)^{1-\ell}\right).$$

Moreover, it is also asymptotically achievable by a maximum likelihood estimator.

Taking the union over all possible P_X leads to the region in (12).

B. Converse Proof of Theorem 3

Assume that the rate/detection-error exponent pair (R, E) is achievable. Then, for all $\epsilon > 0$, there exists n sufficiently large and a code $\mathcal{C}$ such that

$$\frac{\log|\mathcal{C}|}{n} \geqslant R - \epsilon$$

$$\max_{s\in\mathcal{S}} \max_{w\in[1;M]} \mathbb{P}(h(Y^n) \neq w | W = w, S = s) \leqslant \epsilon$$

$$-\frac{1}{n}\log\max_{s\in\mathcal{S}} \frac{1}{M}\sum_{w=1}^{M} \mathbb{P}(g(Z^n) \neq s | S = s, W = w) \geqslant E - \epsilon.$$

Since there is at most a polynomial number of types, there exists a set of types $\mathcal{T}$ such that, for all $P_X \in \mathcal{T}$, the subcode $\mathcal{C}_{P_X} \triangleq \{f(w) : \hat{p}_{f(w)} = P_X\} \subset \mathcal{C}$ satisfies

$$\max_{s\in\mathcal{S}} \max_{w\in f^{-1}(\mathcal{C}_{P_X})} \mathbb{P}(h(Y^n) \neq w | W = w, S = s) \leqslant \epsilon \quad (22)$$

and

$$\frac{\log|\mathcal{C}_{P_X}|}{n} > \frac{\log|\mathcal{C}|}{n} - \delta \geqslant R - \epsilon - \delta \quad (23)$$

for some δ vanishing with ϵ. By [16, Corollary 6.4],

$$\frac{\log|\mathcal{C}_{P_X}|}{n} < \min_s \mathbb{I}(P_X, W_{Y|X,s}) + \tau \quad (24)$$

for some τ vanishing with ϵ. Choose now $P_X^* \in \mathcal{T}$ such that

$$P_X^* \triangleq \operatorname*{argmin}_{P_X\in\mathcal{T}} \phi(P_X) \quad (25)$$

with $\phi(\cdot)$ defined in (13). Then,

$$\begin{aligned}
&E - \epsilon \\
&\leqslant -\frac{1}{n}\log\max_{s\in\mathcal{S}} \frac{1}{M}\sum_{w=1}^{M} \mathbb{P}(g(Z^n) \neq s | S = s, W = w) \\
&\overset{(a)}{\leqslant} -\frac{1}{n}\log\max_{s\in\mathcal{S}} \frac{1}{M}\sum_{w\in f^{-1}(\mathcal{C}_{P_X^*})} \mathbb{P}(g(Z^n) \neq s | S = s, W = w) \\
&\overset{(b)}{\leqslant} -\frac{1}{n}\log\max_{s\in\mathcal{S}} \mathbb{P}(g(Z^n) \neq s | S = s, W = w) + \delta \\
&\overset{(c)}{\leqslant} \phi(P_X^*) + \delta, \quad (26)
\end{aligned}$$

where (a) follows since all terms in the sum are non-negative and keeping only the terms corresponding to messages in $\mathcal{C}_{P_X^*}$; (b) follows by lower-bounding $\frac{|\mathcal{C}_{P_X^*}|}{|\mathcal{C}|}$, noting that the detection error is the same for any message with the same type P_X^*, and using (23); (c) follows by Lemma 6. Combining (24) and (26), we conclude that for all $\epsilon > 0$, there exist $\tau, \delta > 0$ vanishing with ϵ and a type P_X^* such that

$$R \leqslant \min_s \mathbb{I}(P_X^*, W_{Y|X,s}) + \tau + \epsilon + \delta \quad (27)$$

$$E \leqslant \phi(P_X^*) + \delta + \epsilon. \quad (28)$$

Since ϵ, τ, δ can be chosen arbitrarily small as the block length n goes to infinity, E is upper bounded by $\phi(P_X)$ for some $P_X \in \mathcal{P}_X$ and the rate R is achieved by this P_X. Taking the union over all possible P_X completes the result of converse of Theorem 3.

C. Proof of Theorem 5

Fixing some $s \in \mathcal{S}$ and $P_X \in \mathcal{P}_X$, we show that the tuple (R, E) is achievable whenever $R \leqslant \mathbb{I}(P_X, W_{Y|X,s})$ and $E \leqslant \psi_s(P_X)$. We start with defining the code $\mathcal{C} = (\{f_i\}_{i\in[1;n]}, g, h)$. The state estimator is defined as an maximum likelihood estimator, i.e., $g(x^{i-1}, z^{i-1}) = \text{argmax}_{s\in\mathcal{S}} \prod_{\ell=1}^{i-1} W_{Z|x_\ell,s}(z_\ell)$. Fixing any $\Delta_1 > 0$, we define $P_X^{\#} = \text{argmax}_{P_X\in\mathcal{P}_{X,\Delta_1 n}} \phi(P_X)$ and pick a length $n\Delta_1$ sequence $\mathbf{v} = (v_1, ..., v_{\Delta_1 n})$ from the type class $\mathcal{T}_{P_X^{\#}}$. Then, for $1 \leqslant i \leqslant \Delta_1 n$, the encoder is defined as

$$f_i(w, z^{i-1}) = v_i \quad (29)$$

for all $w \in [1; M]$ and $z^{i-1} \in \mathcal{Z}^{i-1}$. At time $\Delta_1 n + 1$, the transmitter would estimate the state by using the maximum likelihood estimator $g(x^{\Delta_1 n}, z^{\Delta_1 n})$. Then, the transmitter would convey the information of the estimated state to the receiver by encoding the estimated state $\tilde{s}$ into a codeword. Since $|\mathcal{S}|$ does not grow with n, there exist a length $\Delta_2 n$ channel code $(\hat{f}, \hat{g})$ with arbitrarily small error probability, where $\hat{f} : \mathcal{S} \mapsto \mathcal{X}^{\Delta_2 n}$ is the encoder and $\hat{g} : \mathcal{Y}^{\Delta_2 n} \mapsto \mathcal{S}$ is the decoder. Denoting $\hat{\mathbf{x}}(\tilde{s}) = (\hat{x}_1(\tilde{s}), ..., \hat{x}_{\Delta_2 n}(\tilde{s})) = \hat{f}(\tilde{s})$ as the codeword corresponding to $\tilde{s}$. Then, for $\Delta_1 n < i \leqslant (\Delta_1 + \Delta_2)n$, the encoder is defined as

$$f_i(w, z^{i-1}) = \hat{x}_{i-\Delta_1 n}(g(x^{\Delta_1 n}, z^{\Delta_1 n})) \quad (30)$$

for all $w \in [1; M]$ and $z^{i-1} \in \mathcal{Z}^{i-1}$.

It is known that for every $\bar{P}_X \in \mathcal{P}_X$, there exists a channel code such that the rate is at least $\mathbb{I}(\bar{P}_X, W_{Y|X,s}) - 2\tau$ for any $\tau > 0$ for all sufficiently large n. Therefore, for the fixed $P_X \in \mathcal{P}_X$, there exist an (n, ϵ) channel code for the state s channel with rate $\mathbb{I}(P_X, W_{Y|X,s}) - 2\tau$ for any $\epsilon, \tau > 0$. Let the channel code for the state s channel be characterized by $(\tilde{f}_s, \tilde{h}_s)$, where $\tilde{f}_s : [1; M] \mapsto \mathcal{X}^{(1-\Delta_1-\Delta_2)n}$ is the encoder and $\tilde{h}_s : \mathcal{Y}^{(1-\Delta_1-\Delta_2)n} \mapsto [1 : M]$ is the decoder. Denoting $\tilde{\mathbf{x}}(w, s) = (\tilde{x}_1(w, s), ..., \tilde{x}_{(1-\Delta_1-\Delta_2)n}(w, s)) = \tilde{f}_s(w)$ as the codeword corresponding to the message $w \in [1; M]$. Then, for $(\Delta_1 + \Delta_2)n < i \leqslant n$, we define the encoder as

$$f_i(w, z^{i-1}) = \tilde{x}_{i-(\Delta_1+\Delta_2)n}(w, g(x^{\Delta_1 n}, z^{\Delta_1 n})) \quad (31)$$

for all $w \in [1; M]$ and $z^{i-1} \in \mathcal{Z}^{i-1}$. Finally, the message decoder is

$$h(y^n) = \tilde{h}_{\hat{g}\left(y_{\Delta_1 n+1}^{\Delta_2 n}\right)}\left(y_{(\Delta_1+\Delta_2)n+1}^{n}\right) \quad (32)$$

for all $y^n \in \mathcal{Y}^n$. The state dependent error probability of communication is

$$\begin{aligned}
P_{\text{c,s}}^{(n)} &= \max_w \mathbb{P}(h(Y^n) \neq W | W = w, S = s) \\
&\leqslant \mathbb{P}(g(Y^{\Delta_1 n}) \neq s | S = s) + \mathbb{P}(\hat{h}(Y_{\Delta_1 n+1}^{\Delta_2 n}) \neq s | S = s) \\
&\quad + \max_w \mathbb{P}\left(\tilde{h}_s\left(y_{(\Delta_1+\Delta_2)n+1}^{n}\right) \neq w \middle| W = w, S = s\right),
\end{aligned}$$

which is arbitrarily small when n is sufficiently large by our construction of the code. The overall rate of communication is

$$R = \frac{1}{n}\log e^{n(1-\Delta_1-\Delta_2)\mathbb{I}(P_X, W_{Y|X,s})} \quad (33)$$

$$= (1 - \Delta_1 - \Delta_2)\mathbb{I}(P_X, W_{Y|X,s}) \tag{34}$$

By making Δ_1, Δ_2 arbitrarily small, we conclude that $\mathbb{I}(P_X, W_{Y|X,s})$ is achievable. Moreover, the error probability of detection is

$$\begin{aligned}
&\mathbb{P}(g(X^n, Z^n) \neq s | S = s) \\
&= \frac{1}{M} \sum_{w=1}^{M} \mathbb{P}(g(X^n, Z^n) \neq s | S = s, M = w) \\
&= \frac{1}{M} \sum_{w=1}^{M} \sum_{s' \neq s} \sum_{s'' \in \mathcal{S}} \\
&\qquad \mathbb{P}(g(X^n, Z^n) = s', g(X^{\Delta_1 n}, Z^{\Delta_1 n}) = s'' | S = s, M = w).
\end{aligned}$$

For each $s' \neq s$ and $s'' \in \mathcal{S}$, defining the set

$$\mathcal{A}_{s',s''} \triangleq \left\{ z^n : g(f(w, z^n), z^n) = s', g(f(w, z^{\Delta_1 n}), z^{\Delta_1 n}) = s'' \right\},$$

then for any $\ell \in [0, 1]$,

$$\begin{aligned}
&\mathbb{P}(g(X^n, Z^n) = s', g(X^{\Delta_1 n}, Z^{\Delta_1 n}) = s'' | S = s, M = w) \\
&= \sum_{z^n \in \mathcal{A}_{s',s''}} \prod_{i=1}^{n} \left(\sum_{x \in \mathcal{X}} W_{Z|x,s}(z_i) \mathbb{P}(X_i = x | Z^{i-1}, w) \right) \\
&= \sum_{z^n \in \mathcal{A}_{s',s''}} \left(\prod_{i=1}^{\Delta_1 n} W_{Z|v_i,s}(z_i) \right) \left(\prod_{i=\Delta_1 n+1}^{(\Delta_1+\Delta_2)n} W_{Z|\hat{x}_{i-\Delta_1 n}(s''),s}(z_i) \right) \\
&\qquad \times \left(\prod_{i=(\Delta_1+\Delta_2)n+1}^{n} W_{Z|\tilde{x}_{i-\Delta_1 n-\Delta_2 n}(w,s''),s}(z_i) \right) \\
&\leqslant \sum_{z^n \in \mathcal{Z}^n} \prod_{i=1}^{\Delta_1 n} \left(W_{Z|v_i,s'}(z_i)\right)^{\ell} \left(W_{Z|v_i,s}(z_i)\right)^{1-\ell} \\
&\times \prod_{i=\Delta_1 n+1}^{(\Delta_1+\Delta_2)n} \left(W_{Z|\hat{x}_{i-\Delta_1 n}(s''),s'}(z_i)\right)^{\ell} \left(W_{Z|\hat{x}_{i-\Delta_1 n}(s''),s}(z_i)\right)^{1-\ell} \\
&\times \prod_{i=(\Delta_1+\Delta_2)n+1}^{n} \left(W_{Z|\tilde{x}_{i-(\Delta_1+\Delta_2)n}(w,s''),s'}(z_i)\right)^{\ell} \\
&\qquad \times \left(W_{Z|\tilde{x}_{i-(\Delta_1+\Delta_2)n}(w,s''),s}(z_i)\right)^{1-\ell} \\
&= \exp\left(n \sum_{x} q_{s''}(x) \log \left(\sum_{z} W_{Z|x,s'}(z)^{\ell} W_{Z|x,s}(z)^{1-\ell} \right) \right),
\end{aligned}$$

where for all $x \in \mathcal{X}$

$$\begin{aligned}
q_{s''}(x) = \frac{1}{n} \Bigg(&\sum_{i=1}^{\Delta_1 n} \mathbf{1}\{v_i = x\} + \sum_{i=\Delta_1 n+1}^{(\Delta_1+\Delta_2)n} \mathbf{1}\{\hat{x}_i(s'') = x\} \\
&+ \sum_{i=\Delta_1 n+\Delta_2 n+1}^{n} \mathbf{1}\{\tilde{x}_i(w, s'') = x\} \Bigg).
\end{aligned}$$

Because above inequality is true for all $\ell \in [0, 1]$, we have

$$\begin{aligned}
&\mathbb{P}(g(X^n, Z^n) = s', g(X^{\Delta_1 n}, Z^{\Delta_1 n}) = s'' | S = s, M = w) \\
&\leqslant \exp\Bigg(-n \Bigg(\max_{\ell \in [0,1]} - \sum_{x} q_{s''}(x) \\
&\qquad \times \log \left(\sum_{z} W_{Z|x,s'}(z)^{\ell} W_{Z|x,s}(z)^{1-\ell} \right) \Bigg) \Bigg)
\end{aligned} \tag{35}$$

Note that for all $w \in [1; M]$, the code $\tilde{\mathbf{x}}(w, s'')$ has the type P_X. By making Δ_1, Δ_2 sufficiently small, we have $|q_{s''}(x) - P_X(x)| \leqslant \eta$ for any $\eta > 0$ and any $s'' \in \mathcal{S}$. Since the smallest exponent will dominate the error probability in (35), we have the detection-error exponent

$$\begin{aligned}
E_{\text{d},s}^{(n)} &\geqslant \min_{s'' \in \mathcal{S}} \min_{s' \neq s} \max_{\ell \in [0,1]} - \sum_{x} q_{s''}(x) \\
&\qquad \times \log \left(\sum_{z} W_{Z|x,s'}(z)^{\ell} W_{Z|x,s}(z)^{1-\ell} \right) \\
&\geqslant \psi_s(P_X) - \xi(\eta),
\end{aligned}$$

where $\xi(\eta)$ is some continuous function in η and $\lim_{\eta \to 0} \xi(\eta) = 0$. By choosing Δ_1, Δ_2 and η arbitrarily small and n sufficiently large, we obtain $E_{\text{d},s}^{(n)} \geqslant \psi_s(P_X)$. By taking the union of all P_X, we conclude that the capacity region $\mathcal{C}_{\text{Joint}}^{s}$ is at least

$$\bigcup_{P_X \in \mathcal{P}_X} \left\{ (R, E) \in \mathbb{R}_+^2 : \begin{cases} R \leqslant \mathbb{I}(P_X, W_{Y|X,s}) \\ E \leqslant \psi_s(P_X) \end{cases} \right. . \tag{36}$$

V. Numerical illustration

We finally present two numerical examples to illustrate our main results. We fist define the channel $W_{YZ|X,S}$ as in Table I. In this example, we assume $\mathcal{Y} = \mathcal{Z} = \mathcal{X} = \{0, 1\}$ and $W_{Z|X,S} = W_{Y|X,S}$. Note that it is impossible to distinguish the state S by only transmitting $X = 0$ or $X = 1$ because for each $X \in \{0, 1\}$ there exists a pair (s, s') such that $W_{Z|X,s} = W_{Z|X,s'}$. Moreover, the input distribution of P_X is near uniform if one tries to maximize the capacity. However, the uniform distribution on X is not the best for estimating S because the distance between $W_{Z|X=0,S=1}$ and $W_{Z|X=0,S=2}$ is greater than the distance between $W_{Z|X=1,S=0}$ and $W_{Z|X=1,S=1}$, and hence, the optimal codeword for detection should have a higher weight on $X = 1$. The joint communication and sensing capacity region corresponding to Table I is given in Fig. 2. One can see that, beyond a certain rate, the optimum error exponent of detection is not achievable. In contrast, the channel given in Table II is BSC and, according to Corollary 4, the best error exponent of detection can always be achieved regardless of the type of codewords. The joint communication and sensing capacity region corresponding to Table II is given in Fig. 3.

TABLE I
TABLE FOR $W_{Z|X,S}(0) = W_{Y|X,S}(0)$ FOR ALL $X \in \{0, 1\}$ AND $S \in \{0, 1, 2\}$.

S \ X	0	1
0	0.9	0.3
1	0.9	0.2
2	0.7	0.2

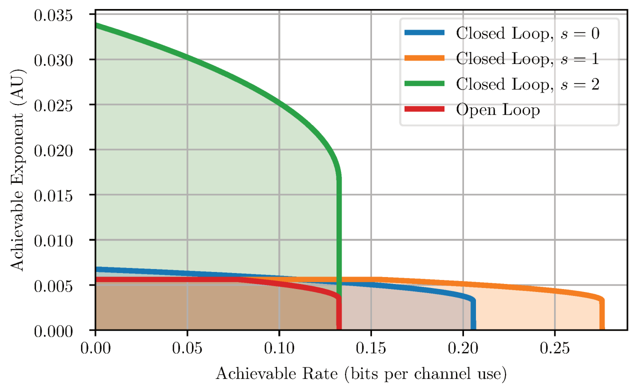

Fig. 2. Capacity region $\mathcal{C}_{\text{open}}$ and achievable region $\mathcal{C}^{s}_{\text{Joint}}$ corresponding to the channel of Table I.

TABLE II
TABLE FOR $W_{Z|X,S}(0) = W_{Y|X,S}(0)$ FOR ALL $X \in \{0, 1\}$ AND $S \in \{0, 1, 2\}$.

S \ X	0	1
0	0.9	0.1
1	0.8	0.2
2	0.7	0.3

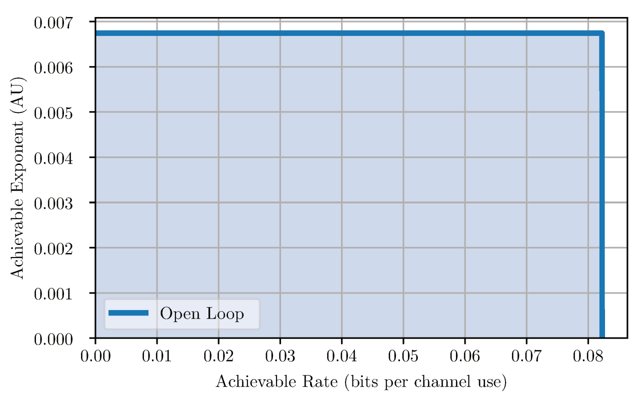

Fig. 3. Capacity region $\mathcal{C}_{\text{open}}$ corresponding to the channel of Table II.

REFERENCES

[1] C. de Lima, D. Belot, A. Berkvens, Rafael andBourdoux, D. Dardari, M. Guillaud, M. Isomursu, E.-S. Lohan, Y. Miao, A. N. Barreto, M. R. K. Aziz, J. Saloranta, T. Sanguanpuak, H. Sarieddeen, G. Seco-Granados, J. Suutala, T. Svensson, M. Valkama, H. Wymeersch, and B. van Liempd, "6g white paper on localization and sensing," 6G Flagship, Tech. Rep., 2020.

[2] T. Wild, V. Braun, and H. Viswanathan, "Joint design of communication and sensing for beyond 5g and 6g systems," *IEEE Access*, vol. 9, pp. 30 845–30 857, 2021.

[3] A. R. Chiriyath, B. Paul, G. M. Jacyna, and D. W. Bliss, "Inner bounds on performance of radar and communications co-existence," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 464–474, Jan. 2016.

[4] L. Zheng, M. Lops, Y. C. Eldar, and X. Wang, "Radar and communication coexistence: An overview: A review of recent methods," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 85–99, sep 2019.

[5] 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 Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, jun 2020.

[6] W. Zhang, S. Vedantam, and U. Mitra, "Joint transmission and state estimation: A constrained channel coding approach," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 7084–7095, oct 2011.

[7] M. Kobayashi, G. Caire, and G. Kramer, "Joint state sensing and communication: Optimal tradeoff for a memoryless case," in *Proc. of IEEE International Symposium on Information Theory*. IEEE, jun 2018.

[8] M. Kobayashi, H. Hamad, G. Kramer, and G. Caire, "Joint state sensing and communication over memoryless multiple access channels," in *Proc. of IEEE International Symposium on Information Theory*. IEEE, Jul. 2019.

[9] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An Information-Theoretic Approach to Joint Sensing and Communication," *arXiv preprint*, vol. 2107.14264, Jul. 2021.

[10] Y. H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.

[11] C. Choudhuri, Y.-H. Kim, and U. Mitra, "Causal State Communication," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3709–3719, Jun. 2013.

[12] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Transactions on Automatic and Control*, vol. 58, no. 10, pp. 2451–2464, Oct. 2013.

[13] M. Naghshvar and T. Javidi, "Sequentiality and adaptivity gains in active hypothesis testing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 768–782, Oct. 2013.

[14] A. Tchamkerten and I. E. Telatar, "On the use of training sequences for channel estimation," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1171–1176, Mar. 2006.

[15] A. Mahajan and S. Tatikonda, "Opportunistic capacity and error exponent region for the compound channel with feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4331–4341, Jul. 2012.

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