Sequential Joint Communication and Sensing of Fixed Channel States

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

2023 IEEE Information Theory Workshop (ITW) | 979-8-3503-0149-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ITW55543.2023.10161688

***Abstract*—We consider a communication model in which a transmitter attempts to communicate with a receiver over a state-dependent channel and simultaneously estimates the state using strictly causal noisy state observations. The state is assumed to remain constant over the duration of the transmission. We analyze the trade-off between the state-error exponent and the communication rate in the sequential setting, in which the transmitter determines what and how many symbols to transmit in an online manner.**

I. Introduction

Joint communication and sensing, also referred to as integrated sensing and communication, is envisioned as a key component of next-generation networks, motivated in large part by the convergence of frequency bands used for communication and radar [1]–[3]. The dual use of radio waveforms for sensing and communication naturally leads to unavoidable tradeoffs in sensing and communication performance and a host of related resource allocation optimization problems [4]–[7].

In particular, there has been interest in studying the information-theoretic limits of joint communication and sensing, in the sense of characterizing the *optimal* tradeoffs incurred by the joint operation irrespective of the computational power available. Two main approaches have been adopted thus far. In high-mobility scenarios, joint communication and sensing can be abstracted as a joint communication and channel estimation problem using information-theoretic models developed for wireless channels [4]. In particular, [6] characterizes the rate-distortion capacity region for a state-dependent Discrete Memoryless Channel (DMC) when the transmitter is tasked with simultaneously estimating the independent and identically distributed (i.i.d.) state and communicating reliably with the receiver. Extensions to multi-user models [6] and secure joint communication and sensing [7]–[9] have also been considered. In low-mobility scenarios, in which the communication rate largely exceeds the rate of change of the sensed phenomenon, joint communication and sensing can be abstracted as a joint communication and parameter estimation problem using ideas from controlled sensing and active hypothesis testing [10], [11]. For parameters taking finitely many values, tradeoffs take the form of a rate-detection exponent capacity region that can be fully characterized in open loop [12]–[14]. In the case of parameters taking a continuum of values, the tradeoff takes the form of a Cramér-Rao vs. rate region, see, e.g., [15]. While the benefits of closed-loop operation are readily established [12], an exact characterization of the capacity region remains elusive.

This work has been supported by the National Science Foundation (NSF) under grant 2148400.

We are interested in furthering the information-theoretic study of low-mobility scenarios by considering a *sequential* estimation setting [16], in which the communication and the sensing are allowed to operate with variable length. This is motivated by the better performance of sequential tests compared to fixed length-tests [11], [17] and by the observation that transmissions often occur over multiple consecutive packets and not just a single packet. The main insight resulting from our study is that the benefits of sequential adaptation to estimate the channel state can be achieved by adapting the transmission parameters at the packet level, resulting in an improved tradeoff between sensing and communication performance.

II. Notation

We let $\mathbb{F}_2 \triangleq \{0,1\}$ and define $\mathcal{M} = \bigcup_{k=1}^{\infty} \mathbb{F}_2^k$; with a slight abuse of notation, we also let $\mathbb{F}_2^{\infty}$ denote a countably infinite sequence consisting of $\{0,1\}$ without additional algebraic structure. For any $U \in \mathcal{M}$, we denote $U[1;k]$ the length k subsequence of U. We also define $\mathcal{B} = \{\text{stop}, \text{continue}\}$. Let $\mathcal{X}$ be any discrete set, and $\mathcal{P}_{\mathcal{X}}$ be the set of all probability distributions on $\mathcal{X}$. We denote by X a random variable in $\mathcal{X}$ and X^n a random sequence with length $n \in \mathbb{N}_+$, where $\mathbb{N}_+$ is a set of all positive integers. $\sigma(X)$ is the sigma algebra generated by the random variable X. Let $\mathcal{Y}$ be another discrete set. Let $P_{Y|X}$ and $W_{Y|X}$ be two conditional distributions on $Y \in \mathcal{Y}$ given $X \in \mathcal{X}$, we define $\mathbb{H}(W_{Y|X}|P_X) \triangleq \mathbb{E}_{P_X}\left[\mathbb{H}(W_{Y|X}(\cdot|X))\right]$ the conditional entropy of $W_{Y|X}$ given input distribution P_X and $\mathbb{I}(P_X, W_{Y|X}) \triangleq \mathbb{H}(W_{Y|X} \circ P_X) - \mathbb{H}(W_{Y|X}|P_X)$ the mutual information between X and Y, where $X \sim P_X$ and $Y \sim P_X \circ W_{Y|X} \triangleq \sum_x P_X(x) W_{Y|X}(\cdot|x)$. We also define the relative entropy given an input distribution P_X as

$$\mathbb{D}(P_{Y|X} \,\|\, W_{Y|X}|P_X) \triangleq \mathbb{E}_{P_X}\left[\mathbb{D}(P_{Y|X}(\cdot|X) \,\|\, W_{Y|X}(\cdot|X))\right].$$

III. Channel Models and Main Results

We consider a sequential mono-static joint communication and sensing problem, in which a transmitter communicates with a receiver over a state-dependent DMC, also known as a compound channel, while simultaneously probing the channel state in a strictly causal manner through a sensing channel. For each time $t \in \mathbb{N}$, we denote by $X_t \in \mathcal{X}$ the input of the state-dependent DMC channel $W_{YZ|XS}$, and by

$Y_t \in \mathcal{Y}$ and $Z_t \in \mathcal{Z}$ the outputs of the communication channel and the sensing channel, respectively. The a priori unknown state $S \in \Theta$ is assumed to be *fixed* during the whole duration of the transmission and takes value in a finite set Θ. A sequential joint communication and sensing policy $\Lambda = (\{\psi_t\}_{t=1}^{\infty}, \{f_t\}_{t=1}^{\infty}, g, h, \{\phi_t\}_{t=1}^{\infty})$ consists of a stopping rule $\{\psi_t\}_{t=1}^{\infty}$, a set of encoding functions $\{f_t\}_{t=1}^{\infty}$, a state estimator g, a message decoder h, and a set of functions $\{\phi_t\}_{t=1}^{\infty}$ that determine the number of message bits being transmitted. The latter is required at the transmitter, since the number of channel uses is not fixed before transmission. Let M_t be the number of message bits determined at time t by the function ϕ_t. We require that the encoder only rely on the first M_t bits of the message sequence $\{U_t\}_{t-1}^{\infty}$, where $U_t \sim \text{Ber}(1/2)$. Detailed descriptions of each component of a policy Λ are as follows.

- $\phi_t : \mathbb{F}_2^{M_{t-1}} \times \mathcal{Z}^{t-1} \times \mathcal{X}^{t-1} \mapsto \mathbb{N} : (U^{M_{t-1}}, Z^{t-1}, X^{t-1}) \mapsto \phi_t(U^{M_{t-1}}, Z^{t-1}, X^{t-1}) = M_t$ is a function that determines the number of message bits M_t being transmitted by the time t. In addition, we require that each realization of $\{M_t\}_{t=1}^{\infty}$ be a non-decreasing sequence.
- $f_t : \mathbb{F}_2^{M_t} \times \mathcal{Z}^{t-1} \times \mathcal{X}^{t-1} \mapsto \mathcal{X} : (U^{M_t}, Z^{t-1}, X^{t-1}) \mapsto f_t(U^{M_t}, Z^{t-1}, X^{t-1}) = X_t$ is the function that determines the input of the channel at time t.
- $\psi_t : \mathbb{F}_2^{M_t} \times \mathcal{Z}^t \times \mathcal{X}^t \mapsto \mathcal{B}$: $(U^{M_t}, Z^t, X^t) \mapsto \psi_t(U^{M_t}, Z^t, X^t) = Q_t$, where $Q_t \in \mathcal{B}$ is the status indicating whether the transmission continues or not at time t. We also define $T \triangleq 1 + \sum_{t=1}^{\infty} \mathbf{1}\{Q_t = \text{continue}\}$ as the stopping time. Note that the transmission only happens at time t if $t \leqslant T$.
- $g : \mathbb{F}_2^{M_T} \times \mathcal{Z}^T \times \mathcal{X}^T \mapsto \Theta : (U^{M_T}, Z^T, X^T) \mapsto g(U^{M_T}, Z^T, X^T) = \widehat{S}$, where $\widehat{S} \in \Theta$ is the estimated state.
- $h : \mathcal{Y}^T \mapsto \mathcal{M} : Y^T \mapsto h(Y^T) = \widehat{U}$, where $\widehat{U} \in \mathcal{M}$ is the decoded message.

Under the above definition of a policy, the probability that $(M^t, X^t, Y^t, Z^t) = (m^t, x^t, y^t, z^t)$ when the state $S = \theta$ and the message sequence $U^{\infty} = w$ is given by

$$
\begin{aligned}
&\mathbb{P}(M^t = m^t, X^t = x^t, Y^t = y^t, Z^t = z^t | S = \theta, U^{\infty} = w) \\
&= \prod_{i=1}^{t} W_{YZ|XS}(y_i, z_i | x_i, \theta) \mathbf{1}\{f_i(w[1; m_i], z^{i-1}, x^{i-1}) = x_i\} \\
&\qquad \times \mathbf{1}\{\phi_i(w[1; m_{i-1}], z^{i-1}, x^{i-1}) = m_i\}.
\end{aligned}
$$

We shall denote $W_{Z|X,\theta}(\cdot|\cdot) \triangleq W_{Z|XS}(\cdot|\cdot, \theta)$ the conditional probability of Z given X when $S = \theta$ for any $\theta \in \Theta$. Moreover, we assume that $0 < \mathbb{I}(P_X, W_{Y|X,\theta}) < \infty$ for any $\theta \in \Theta$ and $P_X \in \mathcal{P}_X$, and $0 < \mathbb{D}\left(W_{Z|X,\theta}(\cdot|x) \| W_{Z|X,\theta'}(\cdot|x)\right) < \infty$ for any $x \in \mathcal{X}$, $\theta \in \Theta$ and $\theta' \neq \theta$. Let n be the constraint on the stopping time T. Specifically, we require that

$$
\min_{\theta \in \Theta} \min_{w \in \mathbb{F}_2^{\infty}} \mathbb{P}(T \leqslant n | S = \theta, U^{\infty} = w) \geqslant 1 - \delta_1 \quad \forall \delta_1 > 0.
$$

The detection-error probability and the communication-error probability when the expected stopping time satisfies the constraint n are defined as

$$
P_d^{(n)} \triangleq \max_{\theta \in \Theta} \max_{w \in \mathbb{F}_2^{\infty}} \mathbb{P}\left(g(U^{M_T}, Z^T, X^T) \neq \theta | S = \theta, U^{\infty} = w\right)
$$

$$
P_c^{(n)} \triangleq \max_{\theta \in \Theta} \max_{w \in \mathbb{F}_2^{\infty}} \mathbb{P}\left(h(Y^T) \neq w[1; M_T] | S = \theta, U^{\infty} = w\right)
$$

We are interested in analyzing the tradeoff between the detection-error exponent and the rate of the policy Λ defined as follows:

$$
R^{(n)} \triangleq \frac{M_T}{n} \quad \text{and} \quad E_d^{(n)} \triangleq \frac{-1}{n} \log P_d^{(n)}.
$$

Note that the rate $R^{(n)}$ defined above is random. To simplify the notation, we denote by $\mathbb{P}_{\theta,w}$ the probability measure when $S = \theta$ and $U^{\infty} = w$. This leads to the following definition of achievability and sequential joint communication and sensing capacity.

Definition 1. *A policy Λ is (R, E) achievable if for any $\delta_1, \delta_2, \delta_3, \epsilon_1 > 0$, there exists some $n(\delta_1, \delta_2, \delta_3, \epsilon_1)$ such that*

$$
\max_{\theta \in \Theta} \max_{w \in \mathbb{F}_2^{\infty}} \mathbb{P}_{\theta,w}(T > n) \leqslant \delta_1 \tag{1}
$$

$$
\min_{\theta \in \Theta} \min_{w \in \mathbb{F}_2^{\infty}} \mathbb{P}_{\theta,w}\left(R^{(n)} \geqslant R\right) \geqslant 1 - \delta_2 \tag{2}
$$

$$
P_c^{(n)} \leqslant \delta_3 \tag{3}
$$

$$
E_d^{(n)} \geqslant E - \epsilon_1 \tag{4}
$$

for all $n \geqslant n(\delta_1, \delta_2, \delta_3, \epsilon_1)$.

Definition 2. *The closure region of sequential joint communication and sensing $\mathbf{C}$ is the union of all achievable tuples (R, E).*

Theorem 3. *The closure region of sequential joint communication and sensing $\mathbf{C}$ is larger than*

$$
\bigcup_{\{P_{X,\theta}\}_{\theta \in \Theta}} \bigcap_{\theta \in \Theta} \left\{ \begin{array}{l} (R, E) \in \mathbb{R}_+^2 : \\ R \leqslant \mathbb{I}(P_{X,\theta}, W_{Y|X,\theta}) \\ E \leqslant \min_{\theta' \neq \theta} \mathbb{D}\left(W_{Z|X,\theta'} \| W_{Z|X,\theta} | P_{X,\theta'}\right) \end{array} \right\}, \tag{5}
$$

where the notation $\bigcup_{\{P_{X,\theta}\}_{\theta \in \Theta}}$ means that we take the union over all possible $|\Theta|$-tuples of distributions on $\mathcal{X}$.

At first glance, the region (5) looks similar to the following open-loop capacity region [12]:

$$
\mathbf{C}_{\text{open}}^{(m)} = \bigcup_{P_X \in \mathcal{P}_{\mathcal{X}}} \left\{ \begin{array}{l} (R, E) \in \mathbb{R}_+^2 : \\ R \leqslant \min_{s \in \mathcal{S}} \mathbb{I}(P_X, W_{Y|X,s}) \\ E \leqslant \rho(P_X) \end{array} \right\}, \tag{6}
$$

where $\rho(P_X)$ is the minimum Chernoff information between states when the type of inputs is P_X [12]. However, the subtlety lies in the fact that we consider the union of all possible $|\Theta|$-tuples of distributions on $\mathcal{X}$, and this allows us to choose types independently for each channel state. Therefore, the region is no longer restricted to the compound capacity. In addition, the sensing exponent is boosted from the Chernoff information $\rho(P_X)$ to the relative entropy as in (5) because of the nature of the sequential setting. We provide a numerical example in Fig. 1 for the compound channel given in Table I to illustrate the benefit of operating in the sequential setting.

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

θ \ X	0	1
0	0.95	0.45
1	0.9	0.2
2	0.55	0.05

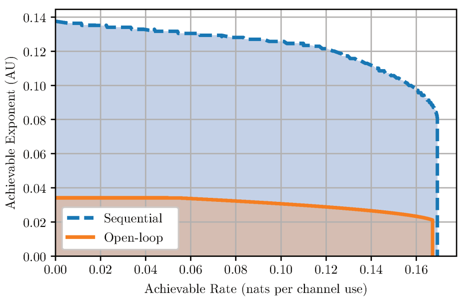

Fig. 1. Numerical illustration of the achievable region with sequential policy and open-loop capacity regions corresponding to the channel given in Table I.

IV. Construction of the Policy

We first construct the policy Λ. For any $\theta \in \Theta$ and $w \in \mathbb{F}_2^{\infty}$, we define the generalized likelihood ratio $A_{\theta,w}$ at any time $t \in \mathbb{N}$ as

$$A_{\theta,w}(t) \triangleq \log \frac{\mathbb{P}_{\theta,w}(M^t, Z^t, X^t)}{\max_{\theta' \neq \theta} \mathbb{P}_{\theta',w}(M^t, Z^t, X^t)}. \tag{7}$$

Similarly, for any θ, θ', w and any $t \in \mathbb{N}_+$, $A_{\theta,\theta',w}$ is the ordinary likelihood ratio defined as

$$A_{\theta,\theta',w}(t) \triangleq \log \frac{\mathbb{P}_{\theta,w}(M^t, Z^t, X^t)}{\mathbb{P}_{\theta',w}(M^t, Z^t, X^t)}. \tag{8}$$

For each fixed $n \in \mathbb{N}$, we define $N = \sqrt{n}$ as the block length. The core idea of of our policy is to determine whether the transmission stops or not at the end of each packet of blocklength N. The rate of each block is determined by the transmitter in a online manner and is unknown to the receiver of the communication channel. Therefore, a fraction $\Delta > 0$ of the block length N is utilized to convey the code structure. The detailed construction of our policy Λ is as follows.

a) Types of codes for each $\theta \in \Theta$:

The types that maximize the capacities for different channel states are different. Therefore, we start by choosing the set of types $\{P_{X,\theta}\}_{\theta\in\Theta}$. Fix any $\Delta > 0$, for each $\theta \in \Theta$, the type $P_{X,\theta} \in \mathcal{P}_X^{(1-\Delta)N}$ is any type of a length $(1-\Delta)N$ sequence. The choice of $\{P_{X,\theta}\}_{\theta\in\Theta}$ is done offline before the transmission, and we assume that the information of $\{P_{X,\theta}\}_{\theta\in\Theta}$ is known by both transceivers as part of the design of the policy. We also define

$$\theta_{\min} = \arg\min_{\theta\in\Theta} \mathbb{I}\left(P_{X,\theta}, W_{Z|X,\theta}\right)$$

and

$$P_{X,\min} = P_{X,\theta_{\min}}.$$

b) Stopping rule

When $t \bmod N = 0$, the stopping rule ψ_t is defined as

$$\psi_t(Z^t, X^t) = \begin{cases} \text{stop} & \text{if } \exists\, \theta \in \Theta \text{ such that } A_{\theta,\theta',w}(t) \geqslant \Gamma_{\theta,\theta'}\ \forall\, \theta' \neq \theta \\ \text{continue} & \text{otherwise,} \end{cases}$$

where

$$\Gamma_{\theta,\theta'} = n\left(\mathbb{D}\left(W_{Z|X,\theta} \middle\| W_{Z|X,\theta'} \middle| P_{X,\theta}\right) - \tau\right) \tag{9}$$

for some $\tau > 0$. On the other hand, if $t \bmod N \neq 0$,

$$\psi_t(Z^t, X^t) = \text{continue} \tag{10}$$

for all $Z^t \in \mathcal{Z}^t$ and $X^t \in \mathcal{X}^t$. That is, we only decide whether the sequential test should stop every N time steps. Under this definition of the stopping rule, the stopping time T can be written as

$$T = \inf_t \left\{ A_{\theta,\theta',w}\left(N\left\lfloor \frac{t}{N} \right\rfloor\right) \geqslant \Gamma_{\theta,\theta'}\ \forall\, \theta' \neq \theta \text{ for some } \theta \in \Theta \right\}.$$

c) Number of message bits being transmitted

For each $k \in \mathbb{N}_+$, the transmitted symbols during the k-th block $\mathbb{N}_k \triangleq [(k-1)N+1; kN]$ correspond to a codeword whose type simultaneously determines the performance of the communication and sensing. The number of message bits transmitted during the k-th block is defined as

$$I_k \triangleq \begin{cases} \left\lfloor N\left(\mathbb{I}\left(P_{X,\hat{\theta}((k-1)N)}, W_{Y|X,\hat{\theta}((k-1)N)}\right) - \epsilon\right)\right\rfloor & \text{if } k \geqslant n^{1/4}, \\ \left\lfloor N\left(\min_{\theta'} \mathbb{I}\left(P_{X,\theta'}, W_{Y|X,\theta'}\right) - \epsilon\right)\right\rfloor & \text{if } k < n^{1/4}, \end{cases}$$

where $\epsilon > 0$ is a positive value related to Δ and will be fully discussed in Section V-(d), and

$$\hat{\theta}(t) = \arg\max_{\theta\in\Theta} A_{\theta,w}(t) \tag{11}$$

is the estimate of θ at each time t. Then, the set of functions $\{\phi_t\}_{t=1}^{\infty}$ that determine the total number of message bits transmitted by time t are

$$\phi_t(Z^{t-1}, X^{t-1}) = \phi_{t-1}(Z^{t-2}, X^{t-2}) + \begin{cases} I_{\lceil \frac{t}{N} \rceil} & \text{if } t \equiv 1 \\ 0 & \text{otherwise} \end{cases},$$

where the notation $t \equiv 1$ means $t \equiv 1 (\bmod N)$. We emphasize that $\hat{\theta}(t)$ is a function of M^t, Z^t, X^t, which makes ϕ_t a function of Z^{t-1}, X^{t-1}. Moreover, from our definition of ϕ_t, the sequence of $\{M_t\}_{t=1}^{\infty}$ is non-decreasing.

d) Conveying the information of the code being used

We use a constant composition code with either type $P_{X,\hat{\theta}((k-1)N)}$ or $P_{X,\min}$, depending on whether $k \geqslant n^{1/4}$ or not, to transmit I_k message bits. When $k < n^{1/4}$, the value of I_k as well as the type $P_{X,\min}$ are known by the receiver. However, this information is unknown by the receiver when $k \geqslant n^{1/4}$. Since I_k only depends on $\hat{\theta}((k-1)N)$, at most $\log|\Theta|$ number of bits are required to convey the information about the structure of the code, i.e, $P_{X,\hat{\theta}((k-1)N)}$ and I_k. Note that $\lim_{n\to\infty} \frac{\log|\Theta|}{\Delta N} = 0$. for any $\Delta > 0$. Therefore, there exists some length ΔN code with type $\tilde{P}_X \in \mathcal{P}_X^{\Delta N}$ such that the error probability of conveying the information of $\hat{\theta}((k-1)N)$ is vanishing with N. Let $(\hat{f}, \hat{h})$ be the encoder/decoder pair of such code, where $\hat{f}(\theta') =$

$\hat{\mathbf{x}}(\theta') = (\hat{x}_1(\theta'), \cdots, \hat{x}_{\Delta N}(\theta')) \in \mathcal{X}^{\Delta N}$ is the codeword corresponding to $\theta' \in \Theta$. Then, for each $k \in \mathbb{N}_+$, the encoder f_t for all $t \in [(k-1)N+1; (k-1+\Delta)N]$is defined as

$$f_t(U^{M_t}, Z^{t-1}, X^{t-1}) = \hat{x}_{t-(k-1)N}(\hat{\theta}((k-1)N)). \quad (12)$$

e) A constant composition encoding for each block

In part (d), we have defined the encoders for all $t \in [(k-1)N+1; (k-1+\Delta)N]$ for all $k \in \mathbb{N}_+$. We now proceed to define the encoders for the remaining time steps. Let $\{\mathcal{C}_{\theta'}\}_{\theta' \in \Theta}$ be the set of constant composition codes, where for each $\theta' \in \Theta$, $\mathcal{C}_{\theta'} = \{\tilde{\mathbf{x}}_{\theta'}(w)\}_w$ is the set of length $(1-\Delta)N$ codewords and $(\tilde{f}_{\theta'}, \tilde{h}_{\theta'})$ is the encoder/decoder pair of such code, where $\tilde{f}_{\theta'}(w) = \tilde{\mathbf{x}}_{\theta'}(w)$ is the type $P_{X,\theta'}$ codeword corresponding to the message $w \in [1; 2^{N(\mathbb{I}(P_{X,\theta'}, W_{Y|X,\theta'})-\epsilon)}]$ in the code $\mathcal{C}_{\theta'}$. We also denote $\tilde{\mathbf{x}}_{\theta'}(w) = (\tilde{x}_{\theta',1}(w), \cdots, \tilde{x}_{\theta',(1-\Delta)N}(w))$, where each $\tilde{x}_{\theta',i}(w)$ is the ith symbol of the codeword $\tilde{\mathbf{x}}_{\theta'}(w)$. Then, for each $k \in \mathbb{N}_+$, the encoder for each $t \in [(k-1+\Delta)N+1; kN]$ is defined as

$$f_t(U^{M_t}, Z^{t-1}, X^{t-1}) = \begin{cases} \tilde{x}_{\hat{\theta}((k-1)N), t-(k-1+\Delta)N}\left(U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]\right) \text{ if } k \geqslant n^{1/4}, \\ \tilde{x}_{\theta_{\min}, t-(k-1+\Delta)N}\left(U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]\right) \text{ if } k < n^{1/4}. \end{cases}$$

Note that for $1 \leqslant k < n^{1/4}$, we transmit a universal code according to the encoder $\tilde{f}_{\theta_{\min}}$, and the reason is to communicate reliably while obtaining an initial estimate of the state. This information is known by the receiver as part of the design of the policy. To simplify the notation, we sometimes write $\tilde{x}_{\theta',j}$ to denote the jth symbol of the codeword $\tilde{\mathbf{x}}_{\theta'}\left(U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]\right)$ when the message $U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]$ is clear in the context.

f) Message decoder

After the receiver obtains the sequence Y^T, the message decoder needs to estimate the transmitted message sequence $U[1; M_T]$. Since M_T is a priori unknown, the receiver first needs to decode from Y^T the value of M_T or, equivalently, I_k for each $k \in \mathbb{N}_+$, which can be done using the decoder $\hat{h}$. Recall that $\hat{h} : \mathcal{Y}^{\Delta N} \mapsto \Theta$ defined in part (d) is the decoder of the codes that convey the information of the code structure. Then, we define $h_k : \mathcal{Y}^N \mapsto \mathcal{M}$ as the message decoder for each block $k \in \mathbb{N}_+$ as follow.

$$h_k\left(Y_{(k-1)N+1}^{kN}\right) = \begin{cases} \tilde{h}_{\hat{h}(Y_{(k-1)N+1}^{(k-1+\Delta)N})}\left(Y_{(k-1+\Delta)N+1}^{kN}\right) \text{ if } k \geqslant n^{1/4} \\ \tilde{h}_{\theta_{\min}}\left(Y_{(k-1+\Delta)N+1}^{kN}\right) \text{ if } k < n^{1/4}. \end{cases}$$

Finally, the decoder $h : \mathcal{Y}^T \mapsto \mathcal{M}$ is defined as the concatenation of all message sub-sequences decoded from $\{h_k\}_{k=1}^{T/N}$, i.e.,

$$h(Y^T) = \left(h_k\left(Y_1^N\right), h_k\left(Y_{N+1}^{2N}\right), \cdots, h_k\left(Y_{T-N+1}^{T}\right)\right).$$

g) State detector

The state detector g is simply defined as

$$g(U^{M_T}, Z^T, X^T) = \hat{\theta}(T) \quad (13)$$

with $\hat{\theta}(t)$ defined in (11) for all $t \in \mathbb{N}_+$.

V. Rate and Exponent Analysis

Before deriving the achievable region of the policy defined in Section IV, we first review a useful lemma that has been derived in the context of sequential hypothesis testing.

Lemma 4. *[17] [10, Lemma 19] The state estimator $\hat{\theta}(t)$ is incorrect for only finitely many $t \in \mathbb{N}_+$, i.e.,*

$$\sum_{t=1}^{\infty} \mathbb{P}_{\theta,w}\left(\hat{\theta}(t) \neq \theta\right) < \infty \quad (14)$$

for any $\theta \in \Theta$ and $w \in \mathbb{F}_2^\infty$.

a) Analysis of the stopping time

For any θ, $\theta' \neq \theta$, $w \in \mathbb{F}_2^\infty$, and $i \in \mathbb{N}_+$, we define

$$L_{\theta,\theta',w}(i) \triangleq \log \frac{W_{Z|X,S}(Z_i|X_i,\theta)}{W_{Z|X,S}(Z_i|X_i,\theta')}, \quad (15)$$

so that $A_{\theta,\theta',w}(t) = \sum_{i=1}^t L_{\theta,\theta',w}(i)$ for any $t \in \mathbb{N}_+$. Note that for each θ, $\theta' \neq \theta$, $w \in \mathbb{F}_2^\infty$, the discrete time stochastic process

$$V_{\theta,\theta',w}(t) \triangleq \sum_{i=1}^t L_{\theta,\theta',w}(i) - \sum_{i=1}^t \mathbb{E}_{\theta,w}[L_{\theta,\theta',w}(i)|\mathcal{F}_{i-1}]$$

is a martingale adapted to the filtration $\{\mathcal{F}_i\}_{i=1}^\infty \triangleq \{\sigma(X^i, Z^i, U^\infty)\}_{i=1}^\infty$. Then, for any $\theta \in \Theta$ and $w \in \mathbb{F}_2^\infty$,

$$\begin{aligned} &\mathbb{P}_{\theta,w}(T > n) \\ &\leqslant \mathbb{P}_{\theta,w}\left(\forall\, \theta' \in \Theta\, \exists\, \theta'' \neq \theta' \text{ s.t } A_{\theta',\theta'',w}(n-N) < \Gamma_{\theta',\theta''}\right) \\ &\leqslant \sum_{\theta'' \neq \theta} \mathbb{P}_{\theta,w}\Bigg(V_{\theta,\theta'',w}(n-N) \\ &\qquad < \Gamma_{\theta,\theta''} - \sum_{i=1}^{n-N} \mathbb{E}_{\theta,w}[L_{\theta,\theta'',w}(i)|\mathcal{F}_{i-1}]\Bigg). \quad (16) \end{aligned}$$

One can show that

$$\begin{aligned} &\sum_{i=1}^{n-N} \mathbb{E}_{\theta,w}[L_{\theta,\theta'',w}(i)|\mathcal{F}_{i-1}] \\ &\qquad \geqslant n(1-\Delta)\left(\mathbb{D}\left(W_{Z|X,\theta} \middle\| W_{Z|X,\theta''} \middle| P_{X,\theta}\right) - \epsilon'\right) \quad (17) \end{aligned}$$

for any $\epsilon' > 0$ when n is sufficiently large by leveraging Lemma 4 and the fact that $N = o(n)$; details are omitted for brevity. Therefore,

$$\begin{aligned} &\Gamma_{\theta,\theta''} - \sum_{i=1}^{n-N} \mathbb{E}_{\theta,w}[L_{\theta,\theta'',w}(i)|\mathcal{F}_{i-1}] \\ &\leqslant -n\left(\tau - \Delta\mathbb{D}\left(W_{Z|X,\theta} \middle\| W_{Z|X,\theta''} \middle| P_{X,\theta}\right) - (1-\Delta)\epsilon'\right). \quad (18) \end{aligned}$$

For any $\tau > 0$, we can choose $\Delta > 0$ and $\epsilon' > 0$ small enough such that

$$\tau - \Delta\mathbb{D}\left(W_{Z|X,\theta} \middle\| W_{Z|X,\theta''} \middle| P_{X,\theta}\right) - (1-\Delta)\epsilon'$$

is positive. Then, by (16), (18) and Azuma's inequality,

$$\mathbb{P}_{\theta,w}(T > n) \leqslant e^{-n\xi_1} \quad (19)$$

for some $\xi_1 > 0$ when n is sufficiently large.

b) Analysis of the rate

For any $\theta \in \Theta$, $\theta' \neq \theta$, $w \in \mathbb{F}_2^\infty$ and $t \in \mathbb{N}_+$, we have

$$A_{\theta,\theta',w}(t) = \sum_{i=1}^{t} \log \frac{W_{Z|X,S}(Z_i|X_i,\theta)}{W_{Z|X,S}(Z_i|X_i,\theta')} \leqslant tC_1, \tag{20}$$

for some $C_1 > 0$, where (20) follows from the fact that $\mathbb{D}\left(W_{Z|X,\theta}\middle\|W_{Z|X,\theta'}\middle|P_X\right)$ is bounded for all $\theta \in \Theta$, $\theta' \neq \theta$ and $P_X \in \mathcal{P}_\mathcal{X}$ by assumption. This implies $T = \Omega(n)$ by the definition of the stopping rule. Then, the number of blocks is $T/N = \Omega(\sqrt{n})$. When the true hypothesis is $\theta \in \Theta$, the number of transmitted message M_T is

$$\begin{aligned} M_T = &\sum_{k=n^{1/4}}^{T/N} \left\lfloor N\left(\mathbb{I}\left(P_{X,\hat{\theta}((k-1)N)}, W_{Y|X,\hat{\theta}((k-1)N)}\right) - \epsilon\right)\right\rfloor \\ &+ (n^{1/4}-1) \times \left\lfloor N\left(\min_\theta \mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right) - \epsilon\right)\right\rfloor \\ \geqslant &(1-\epsilon'')\frac{T}{N}\left\lfloor N\left(\mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right) - \epsilon\right)\right\rfloor \end{aligned} \tag{21}$$

for any $\epsilon'' > 0$ when n is sufficiently large, where (21) follows from the fact that $(n^{1/4}-1) = o(\sqrt{n})$ and that $\sum_{k=1}^\infty \mathbf{1}\left(\hat{\theta}((k-1)N) \neq \theta\right) < \infty$ by Lemma 4. Then, the rate satisfies

$$R^{(n)} \geqslant (1-\kappa) \times \frac{T}{n} \times \left(\mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right) - \epsilon\right) \tag{22}$$

for any $\kappa > 0$ when n is sufficiently large and the true hypothesis is θ. It remains to show that T is greater than $(1-\kappa)n$ with high probability. By definition of the stopping rule, we have

$$\begin{aligned} &\mathbb{P}_{\theta,w}\left(T \leqslant (1-\kappa)n\right) \\ &\leqslant \sum_{\theta' \in \Theta} \sum_{1 \leqslant t \leqslant (1-\kappa)n} \mathbb{P}_{\theta,w}\left(A_{\theta',\theta'',w}(t) \geqslant \Gamma_{\theta',\theta''}\ \forall \theta'' \neq \theta'\right), \end{aligned} \tag{23}$$

where one can show that (details omitted) each term in the summation of (23) can be upper bounded by $e^{-n\xi_2}$ for some $\xi_2 > 0$ when n is sufficiently large. Therefore,

$$\begin{aligned} &\mathbb{P}_{\theta,w}\left(R^{(n)} \geqslant (1-\kappa)^2 \times \left(\mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right) - \epsilon\right)\right) \\ &\quad \geqslant 1 - |\Theta|(1-\kappa)ne^{-n\xi_2} \end{aligned} \tag{24}$$

when n is sufficiently large for all $\theta \in \Theta$ and $w \in \mathbb{F}_2^\infty$.

c) Analysis of the detection-error exponent

Following the standard change-of-measure techniques [18], we can show that for any $\theta \in \Theta$ and $w \in \mathbb{F}_2^\infty$,

$$\begin{aligned} &\mathbb{P}_{\theta,w}\left(g(U^{M_T}, X^T, Z^T) \neq \theta\right) \\ &\quad \leqslant |\Theta| e^{-n\left(\min_{\theta' \neq \theta} \mathbb{D}\left(W_{Z|X,\theta'}\middle\|W_{Z|X,\theta}\middle|P_X(\theta')\right) - \tau\right)}; \end{aligned} \tag{25}$$

details are omitted for brevity.

d) Analysis of the communication-error probability

$$\begin{aligned} &P_c^{(n)} = \max_{\theta \in \Theta} \max_{w \in \mathbb{F}_2^\infty} \mathbb{P}\left(h(Y^T) \neq w[1;M_T] | S=\theta, U^\infty = w\right) \\ &\leqslant \max_{\theta \in \Theta} \max_{w \in \mathbb{F}_2^\infty} \\ &\left(\sum_{k=1}^{n^{1/4}-1} \mathbb{P}_{\theta,w}\left(\tilde{h}_{\theta_{\min}}(Y^{kN}_{(k-1+\Delta)N+1}) \neq U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]\right)\right. \\ &+ \sum_{k=n^{1/4}}^{T/N} \left(\mathbb{P}_{\theta,w}\left(\hat{\theta}((k-1)N) \neq \theta\right) + \mathbb{P}_{\theta,w}\left(\hat{h}(Y^{(k-1+\Delta)N}_{(k-1)N+1}) \neq \theta\right)\right) \\ &\left.+ \sum_{k=n^{1/4}}^{T/N} \mathbb{P}_{\theta,w}\left(\tilde{h}_\theta(Y^{kN}_{(k-1+\Delta)N+1}) \neq U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]\right)\right), \end{aligned} \tag{26}$$

where in (26) we use the law of total probability so that I_k in the third summation term of (26) satisfies

$$I_k = \left\lfloor N\left(\mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right) - \epsilon\right)\right\rfloor.$$

By Lemma 4, for any $\delta > 0$, there exists some n sufficiently large so that

$$\sum_{k=n^{1/4}}^{T/N} \mathbb{P}_{\theta,w}\left(\hat{\theta}((k-1)N) \neq \theta\right) \leqslant \delta. \tag{27}$$

By well-known results for constant composition codes [19], for any $\theta \in \Theta$, $w \in \mathbb{F}_2^\infty$ and $k \in \mathbb{N}_+$,

$$\mathbb{P}_{\theta,w}\left(\hat{h}(Y^{(k-1+\Delta)N}_{(k-1)N+1}) \neq \theta\right) \leqslant e^{-\Delta N \xi_4} \tag{28}$$

for some $\xi_4 > 0$ when n is sufficiently large. Moreover, for any $\theta \in \Theta$, $w \in \mathbb{F}_2^\infty$ and $k \geqslant n^{1/4}$,

$$\mathbb{P}_{\theta,w}\left(\tilde{h}_\theta(Y^{kN}_{(k-1+\Delta)N+1}) \neq U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]\right) \leqslant e^{-(1-\Delta)N\xi_5}$$

for some $\xi_5 > 0$ when n is sufficiently large if

$$\frac{I_k}{(1-\Delta)N} < \mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right) \tag{29}$$

Similarly, for any $\theta \in \Theta$, $w \in \mathbb{F}_2^\infty$ and $k < n^{1/4}$,

$$\mathbb{P}_{\theta,w}\left(\tilde{h}_{\theta_{\min}}(Y^{kN}_{(k-1+\Delta)N+1}) \neq U\left[\sum_{\ell=1}^{k-1} I_\ell + 1; \sum_{\ell=1}^{k} I_\ell\right]\right) \leqslant e^{-(1-\Delta)N\xi_6}$$

for some $\xi_6 > 0$ when n is sufficiently large if

$$\frac{I_k}{(1-\Delta)N} < \min_{\theta' \in \Theta} \mathbb{I}\left(P_{X,\theta'}, W_{Y|X,\theta'}\right). \tag{30}$$

For any $\Delta > 0$, we can choose

$$\epsilon = 1.5\Delta \max_{\theta' \in \Theta} \mathbb{I}\left(P_{X,\theta'}, W_{Y|X,\theta'}\right) \tag{31}$$

so that (29) and (30) hold for all $k \in \mathbb{N}_+$. Then, when n is sufficiently large,

$$P_c^{(n)} \leqslant \sum_{k=1}^{T/N}\left(e^{-\xi_4 \Delta N} + e^{-\xi_5(1-\Delta)N} + e^{-\xi_6(1-\Delta)N}\right) + \delta. \tag{32}$$

e) Claim of achievable region

By considering all $\theta \in \Theta$ and $w \in \mathbb{F}_2^\infty$ in (19), (24), (25) and (32), we show that the rate $\left(\mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right) - \epsilon\right)$ and the exponent $\min_{\theta \in \Theta}\left(\min_{\theta' \neq \theta} \mathbb{D}\left(W_{Z|X,\theta'}\middle\|W_{Z|X,\theta}\middle|P_X(\theta')\right) - \tau\right)$ are achievable when n is sufficiently large. Note that κ and Δ depend on the choice of τ, and ϵ is a function of Δ as discussed in paragraphs (a), (b) and (d). Therefore, by choosing τ arbitrarily small, there exists some $n(\delta_1, \delta_2, \delta_3, \epsilon_1)$ so that (1)-(4) are satisfied with the rate and exponent tuple

$$\left(\min_{\theta \in \Theta} \mathbb{I}\left(P_{X,\theta}, W_{Y|X,\theta}\right), \min_{\theta' \neq \theta} \mathbb{D}\left(W_{Z|X,\theta'}\middle\|W_{Z|X,\theta}\middle|P_X(\theta')\right)\right).$$

We complete the proof by taking the union over all $\{P_{X,\theta}\}$.

REFERENCES

[1] B. Paul, A. R. Chiriyath, and D. W. Bliss, "Survey of RF communications and sensing convergence research," *IEEE Access*, vol. 5, pp. 252–270, 2017.

[2] 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, 2020.

[3] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1295–1315, 2021.

[4] 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.

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

[6] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An information-theoretic approach to joint sensing and communication," *IEEE Transactions on Information Theory*, pp. 1–1, May 2022.

[7] O. Günlü, M. R. Bloch, R. Schaefer, and A. Yener, "Secure joint communication and sensing," in *Proc. of IEEE International Symposium on Information Theory*, Helskinki, Finland, Jun. 2022, pp. 844–849.

[8] N. Su, F. Liu, and C. Masouros, "Secure radar-communication systems with malicious targets: Integrating radar, communications and jamming functionalities," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 83–95, Jan. 2021.

[9] Z. Ren, L. Qiu, J. Xu, and D. W. K. Ng, "Robust transmit beamforming for secure integrated sensing and communication," Jul. 2022.

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

[11] 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.

[12] M.-C. Chang, T. Erdogan, S.-Y. Wang, and M. R. Bloch, "Rate and detection error-exponent tradeoffs of joint communication and sensing," in *Proc. of IEEE International Symposium on Joint Communications & Sensing*, Vienna, Austria, Mar. 2022, pp. 1–6.

[13] H. Joudeh and F. M. J. Willems, "Joint communication and binary state detection," *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 1, pp. 113–124, Mar. 2022.

[14] H. Wu and H. Joudeh, "On joint communication and channel discrimination," in *Proc. of IEEE International Symposium on Information Theory*, Espoo, Finland, Jun. 2022.

[15] Y. Xiong, F. Liu, Y. Cui, W. Yuan, and T. X. Han, "Flowing the information from shannon to fisher: Towards the fundamental tradeoff in isac," arXiv: 2204.06938, Apr. 2022.

[16] H. Chernoff, "Sequential design of experiments," *The Annals of Mathematical Statistics*, vol. 30, no. 3, pp. 755–770, Sep. 1959.

[17] Y. Li, V. Y. F. Tan, and M. Tomamichel, "Optimal adaptive strategies for sequential quantum hypothesis testing," *Communications in Mathematical Physics*, vol. 392, no. 3, pp. 993–1027, Apr. 2022.

[18] Y. Li and V. Y. F. Tan, "Second-order asymptotics of sequential hypothesis testing," *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 7222–7230, 2020.

[19] I. Csiszár and J. Körner, *Information theory: Coding theorems for discrete memoryless systems*. Cambridge: Cambridge University Press, Jan. 2011.