Covert Joint Communication and Sensing under Variational Distance Constraint

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

***Abstract*—We consider the problem of joint communication and sensing with covertness constraint, in which the transmitter needs to design a signal to simultaneously communicate and probe a channel while escaping detection by a passive eavesdropper. Because of the covertness constraint, the weight of transmitted codewords needs to obey the square-root law. We show the existence of a coding scheme with Pulse Position Modulation that achieves the optimal performance between communication throughput and state sensing-error exponent while remaining covert with respect to a passive eavesdropper. For a binary-input discrete memoryless channel, we show that the performance is solely determined by the covertness constraint, and we characterize the optimal pair of achievable rate and error exponent.**

I. Introduction

The benefits of dual-functional waveforms simultaneously supporting radar and communication have motivated the integration of both communication and sensing [1]–[6]. Such Joint Communication and Sensing (JCS) systems, in general, incur inevitable performance tradeoffs, which have been extensively analyzed [4], [5], [7]–[13]. The dual use of signals for communication and sensing also raises security challenges that have been investigated in [14]–[16]. In particular, the growing concerns for the transmission behaviors to remain undetectable by an adversary bring the idea of *covertness* into the picture [17]. Even in a less malicious scenario, covert JCS can potentially provide a solution to the dynamic spectrum access when secondary users such as sensors are employed.

Our work is largely inspired by the recent development in *covert communications* and *covert sensing*. Following the discovery of the *square-root law* [18] for covert communication, [19]–[21] have characterized the pre-constant in front of the square-root scaling, often identified as the *covert capacity*. Subsequent works have studied covert communications under channel uncertainty, meaning either the legitimate transceiver pair or the passive eavesdropper has uncertainty regarding the channel state, to explore different strategies guaranteeing a low probability of detection [18], [22] when the eavesdropper is aware of the channel state, or even achieve a positive communication rate [23]–[26] when the eavesdropper is not aware of the channel state. Practical covert coding mechanisms have also been investigated in [27]–[32]. Of particular relevance to the present work, the Pulse Position Modulation (PPM) strategy developed in [33], [34] is a crucial component in the algorithms proposed by [28], [29], [32]. Motivated by the possibility of designing covert radar systems, [35]–[38] have studied the problem of covert sensing to characterize how to remain covert while actively performing sensing actions.

The main contributions of our work are twofold. 1) We characterize that when only an innocent symbol and a non-innocent symbol are at the transmitter's disposal, there exist no tradeoff between communication and sensing performance, as the performance is dictated by the covertness constraint. This result is in contrast to prior works [10]–[12]. 2) Our achievability proof suggests that PPM is again a good candidate for joint communication and sensing under a covertness constraint. The main reason is that PPM is a much more structured signaling strategy that is well adapted to the sensing problem.

II. Notation

Let $\mathbb{R}_+$ and $\mathbb{N}_*$ denote all non-negative real numbers and all positive integers, respectively. For any set Ω, the indicator function is defined as $\mathbf{1}(\omega \in \Omega) = 1$ if $\omega \in \Omega$ and 0 otherwise. For any discrete set $\mathcal{X}$ and $n \in \mathbb{N}_*$, 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. Following the standard method of type [39], we let $\mathcal{P}_{\mathcal{X}}$ be the set of all probability distributions on $\mathcal{X}$. For $\mathbf{x} \in \mathcal{X}^n$, $\hat{p}_{\mathbf{x}}$ denotes the type of $\mathbf{x}$, i.e., $\hat{p}_{\mathbf{x}}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{x_i = x\}$. $\mathcal{P}_{\mathcal{X}}^n$ is the set of all possible types for length n sequences in $\mathcal{X}^n$. We let $\mathbb{H}(P_X) \triangleq -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x)$ be the entropy of $X \sim P_X$. If $W_{Y|X}$ is a conditional distribution on $Y \in \mathcal{Y}$ given $X \in \mathcal{X}$, $\mathbb{H}\big(W_{Y|X}\big|P_X\big) \triangleq \mathbb{E}_{P_X}\left[\mathbb{H}\big(W_{Y|X}(\cdot|X)\big)\right]$ is the conditional entropy of $W_{Y|X}$ given an input distribution P_X and $\mathbb{I}(P_X, W_{Y|X}) \triangleq \mathbb{H}\big(W_{Y|X} \circ P_X\big) - \mathbb{H}\big(W_{Y|X}\big|P_X\big)$ is the mutual information between X and Y, where $X \sim P_X$ and $Y \sim P_X \circ W_{Y|X} \triangleq \sum_x P_X(x) W_{Y|X}(\cdot|x)$. For two distributions P_X and Q_X over the same set $\mathcal{X}$, we let $P_X \ll Q_X$ denote the absolute continuity with respect to (w.r.t.) Q_X, if for any $x \in \mathcal{X}$, $Q_X(x) = 0$ implies $P_X(x) = 0$. Also, the relative entropy is $\mathbb{D}(P_X \,\|\, Q_X) \triangleq \sum_x P_X(x) \log \frac{P_X(x)}{Q_X(x)}$, the variational distance is $\mathbb{V}(P_X, Q_X) \triangleq \frac{1}{2}\sum_x |P_X(x) - Q_X(x)|$, and the chi-squared distance is $\chi_2(P_X \,\|\, Q_X) \triangleq \sum_x \frac{(P_X(x) - Q_X(x))^2}{Q(x)}$. For $\lambda \in (0,1)$, we define the binary entropy function as $h_b(\lambda)$. Moreover, for $a, b \in \mathbb{R}$ such that $\lfloor a \rfloor \leqslant \lceil b \rceil$, we define $[a;b] \triangleq \{\lfloor a \rfloor, \lfloor a \rfloor + 1, \cdots, \lceil b \rceil - 1, \lceil b \rceil\}$; otherwise

This work has been supported by the National Science Foundation (NSF) under grants 1955401 and 1910859.

$[a;b] \triangleq \emptyset$. In addition, for any $x \in \mathbb{R}$, we let $|x|^+$ denote $\max(x,0)$. For any $x \in \mathbb{R}$, we also define the Q-function $Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx$ and its inverse function $Q^{-1}(\cdot)$. Finally, throughout the paper, $\log$ is w.r.t. base e, and therefore all the information quantities should be understood in *nats*.

III. Covert Joint Communication and Sensing Model

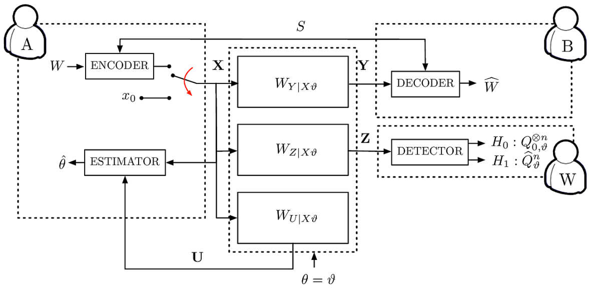

Fig. 1. Covert joint communication and sensing model.

We consider a JCS model with compound channels. As illustrated in Fig. 1, in which a transmitter, Alice, attempts to communicate with a receiver, Bob, over a state-dependent Discrete Memoryless Channel (DMC) $(\Theta, \mathcal{X}, W_{Y|X\theta}, \mathcal{Y})$ while simultaneously probing the channel state θ through a sensing channel $(\Theta, \mathcal{X}, W_{U|X\theta}, \mathcal{U})$. However, a passive eavesdropper, Willie, monitors the activity between the legitimate users through another state-dependent DMC $(\Theta, \mathcal{X}, W_{Z|X\theta}, \mathcal{Z})$. The a priori unknown state θ is assumed to be *fixed* during the whole duration of the transmission and takes value in a finite set Θ ($|\Theta| < \infty$), and Willie is fully aware of the *true realization* of θ, i.e., the channel state information is available at the eavesdropper. More formally, Alice encodes a uniformly distributed message $W \in [1;M]$, with the help of a uniform *secret key* $S \in [1;K]$ preshared with Bob, into a codeword $\mathbf{X}$ of length n as follows:

$$f : [1;M] \times [1;K] \to \mathcal{X}^n : (w,s) \mapsto \mathbf{x}_{ws} \triangleq f(w,s). \quad (1)$$

At the end of transmission, Alice also uses the estimator to detect the channel state θ as follows:

$$g : \mathcal{X}^n \times \mathcal{U}^n \to \Theta : (\mathbf{x}, \mathbf{u}) \mapsto \hat{\theta}. \quad (2)$$

Bob decodes the message with his knowledge of the secret key S as follows:

$$h : \mathcal{Y}^n \times [1;K] \to [1;M] : (\mathbf{y}, s) \mapsto \hat{w}. \quad (3)$$

Following the convention in the covert communication and sensing literature [20], [21], [37], we assume that there exists an *innocent symbol* $x_0 \in \mathcal{X}$ representing the channel input when Alice remains idle, i.e., Alice does not perform any meaningful action, and x_0 induces the following three distributions: $P_{0,\vartheta} \triangleq W_{Y|X=x_0,\theta=\vartheta}, Q_{0,\vartheta} \triangleq W_{Z|X=x_0,\theta=\vartheta}$, and $R_{0,\vartheta} \triangleq W_{U|X=x_0,\theta=\vartheta}$ for all $\vartheta \in \Theta$. Willie declares whether Alice remains idle or not by performing a test T to distinguish the following two hypotheses:

$$T : \mathcal{Z}^n \times \Theta \to \{H_0, H_1\}$$

where $H_0 : \mathbf{Z} \sim Q_{0,\vartheta}^{\otimes n}$, $H_1 : \mathbf{Z} \sim \widehat{Q}_\vartheta^n$, and $\widehat{Q}_\vartheta^n$ is the distribution induced by the encoder f:

$$\widehat{Q}_\vartheta^n(\mathbf{z}) = \frac{1}{MK} \sum_{w=1}^{M} \sum_{s=1}^{K} W_{Z|X\vartheta}^{\otimes n}(\mathbf{z}|f(w,s)).$$

We make the following assumptions:

- $\mu_0 \triangleq \min_{\vartheta \in \Theta, z \in \mathcal{Z}} Q_{0,\vartheta}(z) > 0$.
- There exists another symbol $x_1 \in \mathcal{X}$ distinct from x_0. We then define the associated distributions $P_{1,\vartheta}, Q_{1,\vartheta}$, and $R_{1,\vartheta}$ similar to $P_{0,\vartheta},\ Q_{0,\vartheta}$ and $R_{0,\vartheta}$.
- For every channel state $\vartheta \in \Theta$, $Q_{1,\vartheta} \ll Q_{0,\vartheta}$ and $Q_{1,\vartheta} \neq Q_{0,\vartheta}$.
- For every channel state $\vartheta \in \Theta$, $P_{1,\vartheta} \ll P_{0,\vartheta}$.
- For every channel state $\vartheta \in \Theta$, $R_{0,\vartheta} = R_0$ for some R_0, otherwise Alice can simply perform channel sensing by sending the innocent symbol.

The system performance is measured in terms of the asymptotic throughput of reliable communication and the asymptotic sensing-error exponent while subject to a covertness metric. Formally, we define the communication error probability and the sensing-error probability as follows:

$$P_{\mathrm{c}}^{(n)*} \triangleq \max_{\vartheta \in \Theta, s \in [1;K], w \in [1;M]} \mathbb{P}(h(\mathbf{Y}, s) \neq w | W = w, S = s, \theta = \vartheta)$$
$$P_{\mathrm{s}}^{(n)*} \triangleq \max_{\vartheta \in \Theta, s \in [1;K], w \in [1;M]} \mathbb{P}(g(\mathbf{U}) \neq s | S = s, W = w, \theta = \vartheta),$$

where the dependency on the blocklength n will be omitted when it is clear from the context. The covertness metric is defined as $\max_{\vartheta \in \Theta} \mathbb{V}\left(\widehat{Q}_\vartheta^n, Q_{0,\vartheta}^{\otimes n}\right)$, and it has been justified in [19], [22], [34] that variational distance constraint as a covertness metric is more operationally relevant and provides a 25% gain in throughput compared to the commonly used relative entropy constraint $\mathbb{D}\left(\widehat{Q}_\vartheta^n \middle\| Q_{0,\vartheta}^{\otimes n}\right)$.

Definition 1 (Achievable pairs). *A covert throughput and sensing-error exponent pair $(R, E) \in \mathbb{R}_+ \times \mathbb{R}_+$ is achievable with corresponding key throughput $R_k \in \mathbb{R}_+$, if there exists a sequence of codes $\mathcal{C} = (n, f, g, h, [1;M], [1;K], \delta)$ with increasing blocklength n such that*

$$\lim_{n\to\infty} \frac{\log M}{\sqrt{n}} \geqslant R, \lim_{n\to\infty} \frac{\log MK}{\sqrt{n}} \leqslant R + R_k,$$
$$\lim_{n\to\infty} -\frac{\log P_{\mathrm{s}}^*}{\sqrt{n}} \geqslant E, \lim_{n\to\infty} P_{\mathrm{c}}^* = 0, \max_{\vartheta \in \Theta} \mathbb{V}\left(\widehat{Q}_\vartheta^n, Q_{0,\vartheta}^{\otimes n}\right) \leqslant \delta.$$

We then define $\mathbf{C}$ as the closure of all achievable (R, E) pairs.

Remark 1. *The above sensing strategy is called* open-loop. *A closed-loop strategy would allow Alice to adapt the codewords based on her feedback observations through the sensing* *channel [12], [13], [36], [40], [41]. We do not consider the closed-loop strategy in the present work.*

Remark 2. *We allow Willie to obtain a reliable channel state before the transmission. This assumption not only captures a situation in which a more powerful eavesdropper can control the channel but also prevents the legitimate pair from exploiting the channel uncertainty to go beyond the square-root law [23], [24].*

IV. Main Result

We restrict ourselves to the binary case $\mathcal{X} = \{x_0, x_1\}$ in order to obtain a closed-form result that reflects how the covertness constraint affects JCS performance.

Theorem 2. *The closure region* $\mathbf{C}$ *for* $\mathcal{X} = \{x_0, x_1\}$ *degrades to a rectangular region characterized by the vertex* (R^*, E^*)*:*

$$R^* = \frac{2\min_\vartheta \mathbb{D}(P_{1,\vartheta} \,\|\, P_{0,\vartheta})}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}} Q^{-1}\left(\frac{1-\delta}{2}\right),$$

$$E^* = \frac{2\min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} C\left(\vartheta\|\vartheta'\right)}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}} Q^{-1}\left(\frac{1-\delta}{2}\right),$$

where $C(\vartheta\|\vartheta') \triangleq \sup_{l\in[0,1]} -\log\left(\sum_u R_{1,\vartheta}(u)^l R_{1,\vartheta'}(u)^{1-l}\right)$. *The above region is achievable with key throughput*

$$R_k = \frac{2Q^{-1}\left(\frac{1-\delta}{2}\right)}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}} \times \left|\max_\vartheta \mathbb{D}(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta}) - \min_\vartheta \mathbb{D}(P_{1,\vartheta} \,\|\, P_{0,\vartheta})\right|^+.$$

The above result is conceptually different from prior work [12] in that no tradeoff appears between the covert throughput and sensing-error exponent. The main reason is that the covertness constraint effectively controls the numbers of non-innocent symbol x_1 allowed in the codebook.

Remark 3. *The tradeoff between the covert throughput and sensing-error exponent reappears again in the case* $|\mathcal{X}| > 2$ *with more than one non-innocent symbol. The closure region* $\mathbf{C}$ *is then governed by the distribution among non-innocent symbols, i.e.,* $\overline{P} \in \mathcal{P}_{\mathcal{X}\setminus\{x_0\}}$. *Our conjecture is that* $\mathbf{C} = \cup_{\overline{P}\in\mathcal{P}_{\mathcal{X}\setminus\{x_0\}}}$

$$\left\{\begin{array}{l} (R,E) \in \mathbb{R}_+ \times \mathbb{R}_+ : \\ R \leqslant \frac{2Q^{-1}\left(\frac{1-\delta}{2}\right)\min_\vartheta \sum_{x_i\neq x_0} \overline{P}(x_i)\mathbb{D}\left(P_{i,\vartheta} \,\middle\|\, P_{0,\vartheta}\right)}{\sqrt{\max_\vartheta \chi_2\left(\sum_{x_i\neq x_0} \overline{P}(x_i) Q_{i,\vartheta} \,\middle\|\, Q_{0,\vartheta}\right)}} \\ E \leqslant \frac{2Q^{-1}\left(\frac{1-\delta}{2}\right)\min_{\vartheta\neq\vartheta'} C\left(\vartheta\|\vartheta'|\overline{P}\right)}{\sqrt{\max_\vartheta \chi_2\left(\sum_{x_i} \overline{P}(x_i) Q_{i,\vartheta} \,\middle\|\, Q_{0,\vartheta}\right)}} \end{array}\right\},$$

where $C(\vartheta\|\vartheta'|\overline{P}) \triangleq \sup_{l\in[0,1]} -\sum_{x_i\neq x_0} \overline{P}(x_i) \times \log\left(\sum_u R_{i,\vartheta}(u)^l R_{i,\vartheta'}(u)^{1-l}\right)$.

A. Achievability for Theorem 2

Our achievability proof reuses several results from [33], [34] to analyze the performance of the codebook constructed from PPM symbols. The advantage of utilizing PPM in the present work is to control the type of codewords very precisely. In particular, for the binary input alphabet, this is equivalent to a *constant-composition* codebook characterized by the number of non-innocent symbols x_1, which allows us to precisely analyze the sensing-error exponent.

1) Codebook Construction via Pulse Position Modulation: We first define the notion of PPM super symbol.

Definition 3. *For* $m \in \mathbb{N}_*$, *a PPM super symbol of order* m *is a sequence* $x^m \in \mathcal{X}^m$ *of weight one, i.e.,* $\sum_{i=1}^m \mathbf{1}\{x_i = x_1\} = 1$. *We then denote these PPM super symbols by* $\{\tilde{\mathbf{x}}^{(i)}\}_{i=1}^m$ *and* $\tilde{\mathbf{x}}^{(i)}$ *is the PPM super symbol with* x_1 *at its* i*-th position. The distribution induced by choosing all PPM super symbols of order* m *uniformly at random is* $\widetilde{\Pi}^m_{\text{PPM}}(x^m) = \frac{1}{m}\sum_{i=1}^m \mathbf{1}\{x^m = \tilde{\mathbf{x}}^{(i)}\}$. *PPM super symbols are effectively transmitted over super channels* $(\Theta, \{\tilde{\mathbf{x}}^{(i)}\}_{i=1}^m, W^{\otimes m}_{Y|X\vartheta}, \widetilde{Y} = \mathcal{Y}^m)$, $(\Theta, \{\tilde{\mathbf{x}}^{(i)}\}_{i=1}^m, W^{\otimes m}_{Z|X\vartheta}, \widetilde{Z} = \mathcal{Z}^m)$, *and* $(\Theta, \{\tilde{\mathbf{x}}^{(i)}\}_{i=1}^m, W^{\otimes m}_{U|X}, \widetilde{U} = \mathcal{U}^m)$, *and subsequently induce distributions* $\widetilde{P}^m_{\text{PPM},\vartheta} \triangleq W^{\otimes m}_{Y|X\vartheta} \circ \widetilde{\Pi}^m_{\text{PPM}}$, $\widetilde{Q}^m_{\text{PPM},\vartheta} \triangleq W^{\otimes m}_{Z|X\vartheta} \circ \widetilde{\Pi}^m_{\text{PPM}}$, *and* $\widetilde{R}^m_{\text{PPM},\vartheta} \triangleq W^{\otimes m}_{U|X\vartheta} \circ \widetilde{\Pi}^m_{\text{PPM}}$.

We then construct a random code of blocklength ℓ over PPM super symbols of order m, and the overall blocklength is $n \triangleq m\ell$. The choices of m and ℓ are not independent and both scale like $\mathcal{O}(\sqrt{n})$. By choosing

$$\ell = 2Q^{-1}\left(\frac{1-\delta}{2}\right)\sqrt{\frac{n}{\max_\vartheta \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}} - \frac{2\sqrt{\pi}e^{\frac{Q^{-1}\left(\frac{1-\delta}{2}\right)^2}{2}} n^{\frac{1}{4}}}{\sqrt{Q^{-1}\left(\frac{1-\delta}{2}\right)\max_\vartheta \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})^{\frac{1}{4}}}} + C,$$

where C is some constant independent of n, then by Lemma 4 below, we can show that

$$\max_{\vartheta\in\Theta} \mathbb{V}\left((\widetilde{Q}^m_{\text{PPM},\vartheta})^{\otimes\ell}, (Q^{\otimes m}_{0,\vartheta})^{\otimes\ell}\right) \leqslant \delta - n^{-\frac{1}{2}}, \tag{4}$$

which characterizes the covertness performance of our signaling strategy based on PPM super symbols after ℓ super channel uses.

Lemma 4 (Lemma 8 from [34])**.** *Let* $n, \ell \in \mathbb{N}_*$ *with* $m = \lfloor \frac{n}{\ell} \rfloor$ *large enough and* $\ell = \Theta(m)$. *We have*

$$\max_{\vartheta\in\Theta} \mathbb{V}\left((\widetilde{Q}^m_{PPM,\vartheta})^{\otimes\ell}, Q^{\otimes m\ell}_{0,\theta}\right) \leqslant 1 - 2Q\left(\frac{\ell}{2}\sqrt{\frac{\max_\vartheta \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}{n}}\right) + \frac{2}{\sqrt{\ell}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

Intuitively, the ratio of weight to overall blocklength for each codeword is $1/m$, and the order of PPM becomes larger as we increase the overall blocklength n; the increasing sparsity level is to ensure the covertness constraint is satisfied.

Alice generates MK codewords $\{\tilde{\mathbf{x}}^\ell_{ws} \triangleq f(w,s)\}$ of length ℓ from $\widetilde{\Pi}^m_{\text{PPM}}$ in an independent and identically distributed (i.i.d.) fashion, where $\tilde{\mathbf{x}} \in \{\tilde{\mathbf{x}}^{(i)}\}_{i=1}^m$, $w \in [1;M]$ is the message and $s \in [1;K]$ is the secret key preshared with Bob. Given channel state ϑ, the distribution induced by the codebook over the PPM super channel at Willie's terminal is $\widehat{Q}^n_\vartheta = \widehat{Q}^{m,\ell}_{\text{PPM},\vartheta}$.

2) Channel Reliability: We first introduce the average probability of error induced by the code as

$$\overline{P}_{e,\vartheta,s} \triangleq \mathbb{P}\left(W \neq \widehat{W} | S = s, \theta = \vartheta\right).$$

Lemma 5. *By choosing $\log M = \min_{\vartheta}(1 - 2\xi)\ell\mathbb{D}(P_{1,\vartheta} \,\|\, P_{0,\vartheta})$, the expected average probability of error satisfies*

$$\mathbb{E}_C\left[\overline{P}_{e,\vartheta,s}\right] \leqslant \exp(-\rho_1 n^{1/2-\epsilon_1}) \tag{5}$$

for every key s and every ϑ and for some $\xi, \rho_1 > 0$, $\epsilon_1 \in (0, 1/2)$ and n large enough.

Proof: Modified from [33, Proposition 1]. ■

3) Channel Resolvability:

Lemma 6. *By choosing $\log MK = \max_{\vartheta}(1 + \xi)\ell\mathbb{D}(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})$, the expected variational distance between the codebook-induced distribution and PPM signaling distribution satisfies*

$$\mathbb{E}_C\left[\mathbb{V}\left(\widehat{Q}^{m,\ell}_{\text{PPM},\vartheta}, (\widetilde{Q}^{m}_{PPM,\vartheta})^{\otimes \ell}\right)\right] \leqslant \exp(-\rho_2 n^{1/2-\epsilon_2}), \tag{6}$$

for every ϑ and for some $\xi, \rho_2 > 0$, $\epsilon_2 \in (0, 1/2)$ and n large enough.

Proof: Modified from [33, Proposition 1]. ■

4) Identification of a Specific Universal Code: We will use an idea similar to the step in [34, Lemma 4] to identify the existence of a joint reliability and resolvability code that is universal w.r.t. the compound channel. We first define the following events:

$$\begin{aligned}
&\mathcal{E}_1 \triangleq \{\max_{\vartheta,s} \overline{P}_{e,\vartheta,s} \leqslant \lambda_1 e^{-\rho_1 n^{1/2-\epsilon_1}}\}, \\
&\mathcal{E}_2 \triangleq \Bigg\{\forall \mathcal{I} \subset [1;M] \times [1;K] \text{ such that } |\mathcal{I}| = \lambda_2 MK, \\
&\qquad \max_{\vartheta} \mathbb{V}\left(\widehat{Q}^{m,\ell}_{\mathcal{I},\vartheta}, (\widetilde{Q}^{m}_{\text{PPM},\vartheta})^{\otimes \ell}\right) \leqslant e^{-\rho_2 n^{1/2-\epsilon_2}} + \lambda_3\Bigg\},
\end{aligned}$$

where $\widehat{Q}^{m,\ell}_{\mathcal{I},\vartheta}$ is the distribution induced by codewords in $\mathcal{I}$.

Lemma 7 (Lemma 4 from [34])**.** *Suppose that conditions in Lemma 5 and Lemma 6 are satisfied. If $1 > \exp\left(\log|\Theta| - M\left(\lambda_2\lambda_3^2 + h_b\left(\lambda_2\right)\right)\right) + |\Theta|\lambda_1^{-1}$, then $\mathbb{P}_C\left(\mathcal{E}_1 \cap \mathcal{E}_2\right) > 0$.*

We then choose $\lambda_1 = n, \lambda_2 = 1 - n^{-1}, \lambda_3 = n^{-1}$; we can therefore guarantee the existence of a code $\mathcal{C}$ satisfying that $\max_{\vartheta,s} \overline{P}_{e,\vartheta,s} \leqslant n e^{-\rho_1 n^{1/2-\epsilon_1}}$, and that for every $\mathcal{I} \subset [1;M] \times [1;K]$ such that $|\mathcal{I}| = (1 - n^{-1})MK$,

$$\max_{\vartheta} \mathbb{V}\left(\widehat{Q}^{m,\ell}_{\mathcal{I},\vartheta}, (\widetilde{Q}^{m}_{\text{PPM},\vartheta})^{\otimes \ell}\right) \leqslant e^{-\rho_2 n^{1/2-\epsilon_2}} + n^{-1},$$

since all the conditions in Lemma 7 are satisfied for n large enough. To guarantee the maximal probability of error, we then expurgate the codewords by removing $(1 - \lambda_2)M$ codewords within every sub-codebook $\mathcal{C}_S$ indexed by s, i.e., $\mathcal{C}_s \triangleq \{\tilde{x}^{\ell}_{sw}\}_{w\in[1;M]}$. The codewords which are removed are those with top-$(1 - \lambda_2)$ error probability within each C_s; then by Markov's inequality, the expurgated code $\mathcal{C}'$ then guarantees that for n large enough

$$P^*_{\text{c}} \leqslant n^2 e^{-\rho_1 n^{1/2-\epsilon_1}} \leqslant e^{-\bar{\rho}_1 n^{1/2-\bar{\epsilon}_1}},$$

and the maximal probability of error can be made arbitrarily small as $n \to \infty$. The expurgated code $\mathcal{C}'$ satisfies

$$\max_{\vartheta} \mathbb{V}\left(\widehat{Q}^{n}_{\vartheta}, (\widetilde{Q}^{m}_{\text{PPM},\vartheta})^{\otimes \ell}\right) \leqslant e^{-\rho_2 n^{1/2-\epsilon_2}} + n^{-1}. \tag{7}$$

We then proceed to analyze the covertness metric for $\mathcal{C}'$ as follows:

$$\begin{aligned}
\max_{\vartheta} \mathbb{V}\left(\widehat{Q}^{n}_{\vartheta}, Q^{\otimes n}_{0,\vartheta}\right) &\overset{\text{(a)}}{\leqslant} \max_{\vartheta} \mathbb{V}\left(\widehat{Q}^{n}_{\vartheta}, (\widetilde{Q}^{m}_{\text{PPM},\vartheta})^{\otimes \ell}\right) \\
&\quad + \max_{\vartheta} \mathbb{V}\left((\widetilde{Q}^{m}_{\text{PPM},\vartheta})^{\otimes \ell}, Q_0^{\otimes m\ell}\right) \\
&\overset{\text{(b)}}{\leqslant} e^{-\rho_2 n^{1/2-\epsilon_2}} + n^{-1} + \delta - n^{-\frac{1}{2}} \\
&\overset{\text{(c)}}{\leqslant} \delta,
\end{aligned}$$

where (a) follows from the triangle inequality, (b) follows from (4) and (7), and (c) follows for n large enough.

5) Sensing-Error Exponent: The sensing-error exponent is given by Lemma 8.

Lemma 8 (Lemma 10 from [12], Theorem 1 from [40])**.** *Suppose that the codeword $\{\mathbf{x}_{ws}\}_{w[1;M],s\in[1;K]}$ has type $P_X \in \mathcal{P}_{\mathcal{X}}$, then the maximal sensing-error probability $P^{(n)*}_{\text{s}}$ in an open-loop scheme is lower bounded by $\max_{\vartheta} P^{(n)*}_{\text{s}} \geqslant \Theta(1)e^{-n\phi(P_X)}$, where $\phi(P_X) \triangleq \min_{\vartheta} \min_{\vartheta'\neq\vartheta} \max_{l\in[0,1]} -\sum_x P_X(x)$*
$\times \log\left(\sum_u W_{U|X\vartheta}(u|x)^l W_{U|X\vartheta'}(u|x)^{1-l}\right)$. Moreover, it is also asymptotically achievable by a maximum likelihood estimator g_{ML}, i.e.,

$$\max_{\vartheta} \mathbb{P}\left(g_{\text{ML}}(\mathbf{Z}) \neq \vartheta | \theta = \vartheta, S = s, W = w\right) \leqslant \Theta(1)e^{-n\phi(P_X)}.$$

Indeed, since the code $\mathcal{C}'$ is a constant composition code, every codeword has exactly ℓ x_1 symbols and our assumption that $\forall\vartheta\ R_{0,\vartheta} = R_0$ precludes x_0 from contributing to the sensing-error exponent, we obtain that for n large enough

$$\begin{aligned}
\log P^*_{\text{s}} &= \ell \min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} \max_{l\in[0,1]} -\log\left(\sum_u R_{1,\vartheta}(u)^l R_{1,\vartheta'}(u)^{1-l}\right) - \xi \\
&= \ell \min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} C(\vartheta\|\vartheta') - \xi,
\end{aligned}$$

where ξ is some positive constant.

6) Throughput and Exponent Analysis:

$$\begin{aligned}
\lim_{n\to\infty} \frac{\log M}{\sqrt{n}} &= \lim_{n\to\infty} (1 - 2\xi) \frac{\ell \min_{\vartheta} \mathbb{D}(P_{1,\vartheta} \,\|\, P_{0,\vartheta})}{\sqrt{n}} \\
&= (1 - 2\xi) \frac{2Q^{-1}\left(\frac{1-\delta}{2}\right) \min_{\vartheta} \mathbb{D}(P_{1,\vartheta} \,\|\, P_{0,\vartheta})}{\sqrt{\max_{\vartheta} \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}} \\
\lim_{n\to\infty} \frac{\log MK}{\sqrt{n}} &= \lim_{n\to\infty} (1 + \xi) \frac{\ell \max_{\vartheta} \mathbb{D}(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}{\sqrt{n}} \\
&= (1 + \xi) \frac{2Q^{-1}\left(\frac{1-\delta}{2}\right) \max_{\vartheta} \mathbb{D}(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}{\sqrt{\max_{\vartheta} \chi_2(Q_{1,\vartheta} \,\|\, Q_{0,\vartheta})}} \\
\lim_{n\to\infty} -\frac{\log P_{\text{s}}}{\sqrt{n}} &= \lim_{n\to\infty} \frac{\ell \min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} C(\vartheta\|\vartheta') - \xi}{\sqrt{n}}
\end{aligned}$$

$$= \frac{2Q^{-1}\left(\frac{1-\delta}{2}\right)\min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} C(\vartheta\|\vartheta')}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}}.$$

B. Converse for Theorem 2

The proof largely follows from [34, Section IV.B] and [22, Section III.A] to conclude that there exists a good sub-codebook in the sense that the weight of each codeword is low. We reiterate key steps to provide a self-contained proof.

1) Lower Bound on Covertness Metric: We start by characterizing the covertness metric from the minimum weight of codewords within a given code $\mathcal{M}$. Since Willie is fully aware of the true channel state ϑ, his goal is to distinguish the two hypotheses H_0 and H_1 associated to distributions $Q_{0,\vartheta}^{\otimes n}$ and $\widehat{Q}_\vartheta^n$, respectively, based on his observations $\mathbf{Z}$ of length n. He employs a sub-optimal test T as follows: $T(\mathbf{z},\vartheta) \triangleq \mathbf{1}\{\sum_{i=1}^n A(z_i) > \tau\}$, where $A(z) \triangleq \frac{Q_{1,\vartheta}(z)-Q_{0,\vartheta}(z)}{Q_{0,\vartheta}(z)}$ and τ is some threshold to be determined later. The idea is that using a sub-optimal test potentially provides a looser constraint on our converse result but as we shall see later, it turns out to match the achievability; test T captures the weight requirement within a code subject to the covertness constraint. Lemma 9 then characterizes the probabilities of false alarm α and missed-detection β of a given code $\mathcal{M}$.

Lemma 9 (Lemma 11 from [34])**.** *Consider a specific code $\mathcal{M}$. Let wt_* be the minimum weight of the codewords in $\mathcal{M}$. Then by setting the threshold to $\tau \triangleq \frac{\mathrm{wt}_*}{2}\chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})$ of test T, for any channel state ϑ,*

$$\alpha_\vartheta \leqslant Q\left(\frac{\mathrm{wt}_*}{2}\sqrt{\frac{\chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}{n}}\right) + B_1 n^{-\frac{1}{2}},$$

$$\beta_\vartheta \leqslant Q\left(\frac{\mathrm{wt}_*}{2}\sqrt{\frac{\chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}{n}}\right) + \mathrm{wt}_*^2 B_3 n^{-\frac{3}{2}} + B_2 n^{-\frac{1}{2}},$$

where B_1, B_2 and B_3 are some positive constants independent of n.

We therefore obtain

$$\begin{aligned}\max_\vartheta \mathbb{V}\left(\widehat{Q}_\vartheta^n, Q_{0,\vartheta}^{\otimes n}\right) &\geqslant \max_\vartheta 1-\alpha_\vartheta-\beta_\vartheta \\ &\geqslant 1-2Q\left(\frac{\mathrm{wt}_*}{2}\sqrt{\frac{\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}{n}}\right) \\ &\quad -(B_1+B_2)n^{-\frac{1}{2}} - \mathrm{wt}_*^2 B_3 n^{-\frac{3}{2}}, \qquad (8)\end{aligned}$$

which characterizes the lower bound on covertness metric and relates it to the minimum weight as promised.

2) Existence of a Low-weight Sub-codebook: Now we consider a covert code $\mathcal{C}$ that indeed satisfies the covertness constraint $\max_\vartheta \mathbb{V}\left(\widehat{Q}_\vartheta^n, Q_{0,\vartheta}^{\otimes n}\right) \leqslant \delta$. Lemma 10 then shows that there must exist a low-weight sub-codebook with substantial size to satisfy the covertness constraint.

Lemma 10 (Lemma 12 from [34])**.** *For any code $\mathcal{C}$ satisfies covertness constraint, there exists a sub-codebook $\mathcal{C}^{(\ell)}$ of $\mathcal{C}$ such that $|\mathcal{C}^{(\ell)}| \geqslant \gamma_n|\mathcal{C}|$ and*

$$\max_{\mathbf{x}\in\mathcal{C}^{(\ell)}} \hat{p}_\mathbf{x}(x_1) \leqslant \frac{2Q^{-1}\left(\frac{1-\delta}{2}-\frac{C}{\sqrt{n}}-\gamma_n\right)}{\sqrt{n\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}}, \qquad (9)$$

where $\gamma_n \in (0,1)$, $\lim_{n\to\infty}\gamma_n = 0$ and C is some constant depends on the compound channel.

3) Upper Bound on Covert Throughput: The code $\mathcal{C}$ can be partitioned into K sub-codebooks $\mathcal{C}_s$ indexed by the key s, and each of $\mathcal{C}_s$ has size M. Now consider the sets $\mathcal{C}_s^{(\ell)} \triangleq \mathcal{C}_s \cap \mathcal{C}^{(\ell)}$. The pigeonhole principle would therefore imply that at least one $\mathcal{C}_s^{(\ell)}$ satisfying $|\mathcal{C}_s^{(\ell)}| \geqslant \gamma_n M$. Since the code $\mathcal{C}$ satisfies the maximal probability of error $P_\mathrm{c}^{(n)*}$, the sub-code formed by $\mathcal{C}_s^{(\ell)}$ also satisfies the same reliability constraint. Then by the standard converse techniques for reliable communication, for every channel state ϑ,

$$\log|\mathcal{C}_s^{(\ell)}| \leqslant \frac{n\mathbb{I}(\overline{\Pi}_X, W_{Y|X\vartheta})+1}{1-P_\mathrm{c}^{(n)*}},$$

where $\overline{\Pi}_X(x) \triangleq \frac{1}{n}\sum_{i=1}^n \frac{1}{|\mathcal{C}_s^\ell|}\sum_{\mathbf{x}\in\mathcal{C}_s^{(\ell)}} \mathbf{1}\{x=x_i\} = \frac{1}{|\mathcal{C}_s^{(\ell)}|}\sum_{\mathbf{x}\in\mathcal{C}_s^{(\ell)}} \hat{p}_\mathbf{x}(x)$. Then by Lemma 10, [20, Lemma 1] and [34, (294)], we have for every channel state ϑ,

$$\begin{aligned}\mathbb{I}(\overline{\Pi}_X, W_{Y|X\vartheta}) \leqslant{}& \frac{2\mathbb{D}(P_{1,\vartheta}\,\|\,P_{0,\vartheta})}{\sqrt{n\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}}Q^{-1}\left(\frac{1-\delta}{2}\right) \\ &+\mathcal{O}\left(n^{-1}\right)+\mathcal{O}(n^{-1/2}\gamma_n).\end{aligned}$$

Therefore,

$$\begin{aligned}&\log|\mathcal{C}_s^{(\ell)}| \\ &\leqslant \frac{\frac{2\sqrt{n}\min_\vartheta \mathbb{D}(P_{1,\vartheta}\,\|\,P_{0,\vartheta})}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}}Q^{-1}\left(\frac{1-\delta}{2}\right)+\mathcal{O}\left(1\right)+\mathcal{O}(n^{1/2}\gamma_n)}{1-P_\mathrm{c}^{(n)*}}.\end{aligned}$$

Finally, by choosing $\{\gamma_n\}_n$ such that $\lim_{n\to\infty} -\frac{\log\gamma_n}{\sqrt{n}} = 0$,

$$\begin{aligned}\lim_{n\to\infty}\frac{\log M}{\sqrt{n}} &\leqslant \lim_{n\to\infty}\frac{\log|\mathcal{C}_s^{(\ell)}|-\log\gamma_n}{\sqrt{n}} \\ &= \frac{2\min_\vartheta \mathbb{D}(P_{1,\vartheta}\,\|\,P_{0,\vartheta})}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}}Q^{-1}\left(\frac{1-\delta}{2}\right).\end{aligned}$$

4) Upper Bound on Sensing-Error Exponent: We then characterize an upper bound on sensing-error exponent for $\mathcal{C}^{(\ell)}$, and it subsequently serves as an upper bound for $\mathcal{C}$. We first partition $\mathcal{C}$ into $\mathcal{C}_i^{(\ell)}$, each of them corresponds to a specific type $P_i \in \mathcal{P}_\mathcal{X}^n$ such that $\mathcal{C}_i^{(\ell)} \triangleq \{\mathbf{x}\in\mathcal{C}^{(\ell)} : \hat{p}_\mathbf{x}(x_1) = i\}$, where $i\in\mathbb{N}$ represents the weight of codewords. Finally,

$$\begin{aligned}&-\log P_\mathrm{s}^{(n)*} \\ &\leqslant -\log \max_{\vartheta,(w,s)\in f^{-1}(\mathcal{C}_i^{(\ell)})} \mathbb{P}(g(\mathbf{U})\neq s|S=s,W=w,\theta=\vartheta) \\ &\overset{\text{(a)}}{\leqslant} \min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} iC(\vartheta\|\vartheta')-\xi \\ &\overset{\text{(b)}}{\leqslant} \frac{2\sqrt{n}\min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} C(\vartheta\|\vartheta')}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}}Q^{-1}\left(\frac{1-\delta}{2}\right)-\xi,\end{aligned}$$

where (a) follows from Lemma 8 and (b) follows from Lemma 10. We obtain

$$\lim_{n\to\infty} -\frac{\log P_\mathrm{s}^{(n)*}}{\sqrt{n}} \leqslant \frac{2\min_{\vartheta,\vartheta':\vartheta\neq\vartheta'} C(\vartheta\|\vartheta')}{\sqrt{\max_\vartheta \chi_2(Q_{1,\vartheta}\,\|\,Q_{0,\vartheta})}}Q^{-1}\left(\frac{1-\delta}{2}\right).$$

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] 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.

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

[4] 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, Nov. 2021.

[5] X. Fang, W. Feng, Y. Chen, N. Ge, and Y. Zhang, "Joint Communication and Sensing Toward 6G: Models and Potential of Using MIMO," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4093–4116, Mar. 2023.

[6] H. He, L. Jiang, Y. Pan, A. Yi, X. Zou, W. Pan, A. E. Willner, X. Fan, Z. He, and L. Yan, "Integrated sensing and communication in an optical fibre," *Light: Science & Applications*, vol. 12, no. 1, p. 25, Jan. 2023.

[7] M. Kobayashi, G. Caire, and G. Kramer, "Joint State Sensing and Communication: Optimal Tradeoff for a Memoryless Case," in *2018 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2018, pp. 111–115.

[8] 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*, Jul. 2019, pp. 270–274.

[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] H. Joudeh and F. M. J. Willems, "Joint Communication and Binary State Detection," *IEEE Journal on Selected Areas in Information Theory*, pp. 1–1, 2022.

[11] H. Wu and H. Joudeh, "On Joint Communication and Channel Discrimination," in *Proc. of 2022 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2022, pp. 3321–3326.

[12] 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 Journal on Selected Areas in Information Theory*, vol. 4, pp. 245–259, May 2023.

[13] M.-C. Chang, S.-Y. Wang, and M. R. Bloch, "Sequential Joint Communication and Sensing of Fixed Channel States," in *proc. of 2023 IEEE Information Theory Workshop (ITW)*, Saint-Malo, France, Apr. 2023, pp. 462–467.

[14] O. Günlü, M. R. Bloch, R. F. Schaefer, and A. Yener, "Secure Integrated Sensing and Communication," *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 40–53, 2023.

[15] H. Wu, Y. Zhang, Y. Shen, X. Jiang, and T. Taleb, "Achieving Covertness and Secrecy: The Interplay Between Detection and Eavesdropping Attacks," *IEEE Internet of Things Journal*, pp. 1–1, 2023.

[16] Z. Wei, F. Liu, C. Masouros, N. Su, and A. P. Petropulu, "Toward Multi-Functional 6G Wireless Networks: Integrating Sensing, Communication, and Security," *IEEE Communications Magazine*, vol. 60, no. 4, pp. 65–71, Apr. 2022.

[17] J. Hu, Q. Lin, S. Yan, X. Zhou, Y. Chen, and F. Shu, "Covert Transmission via Integrated Sensing and Communication Systems," *IEEE Transactions on Vehicular Technology*, pp. 1–6, 2023.

[18] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of Reliable Communication with Low Probability of Detection on AWGN Channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.

[19] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *Proc. of IEEE International Symposium on Information Theory*, Jul. 2013, pp. 2945–2949.

[20] M. R. Bloch, "Covert Communication over Noisy Channels: A Resolvability Perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.

[21] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental Limits of Communication With Low Probability of Detection," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493–3503, Jun. 2016.

[22] S. Y. Wang and M. R. Bloch, "Covert MIMO Communications under Variational Distance Constraint," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4605–4620, Sep. 2021.

[23] P. H. Che, M. Bakshi, C. Chan, and S. Jaggi, "Reliable Deniable Communication with Channel Uncertainty," in *Proc. of IEEE Information Theory Workshop*, Nov. 2014, pp. 30–34.

[24] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert Communication in the Presence of an Uninformed Jammer," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6193–6206, Sep. 2017.

[25] M. Ahmadipour, S. Salehkalaibar, M. H. Yassaee, and V. Y. F. Tan, "Covert Communication Over a Compound Discrete Memoryless Channel," in *proc. of 2019 IEEE International Symposium on Information Theory*, Jul. 2019, pp. 982–986.

[26] A. Bendary, A. Abdelaziz, and C. E. Koksal, "Achieving Positive Covert Capacity Over MIMO AWGN Channels," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 149–162, Mar. 2021.

[27] B. A. Bash, A. H. Gheorghe, M. Patel, J. L. Habif, D. Goeckel, D. Towsley, and S. Guha, "Quantum-secure covert communication on bosonic channels," *Nature Communications*, vol. 6, no. 1, p. 8626, Dec. 2015.

[28] I. A. Kadampot, M. Tahmasbi, and M. R. Bloch, "Multilevel-Coded Pulse-Position Modulation for Covert Communications Over Binary-Input Discrete Memoryless Channels," *IEEE Transactions on Information Theory*, vol. 66, no. 10, pp. 6001–6023, Oct. 2020.

[29] ——, "Codes for Covert Communication over Additive White Gaussian Noise Channels," in *Proc. of IEEE International Symposium on Information Theory*, Paris, France, Jul. 2019, pp. 977–981.

[30] M. Lamarca and D. Matas, "A non-linear channel code for covert communications," in *proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakesh, Morocco, Morocco, Apr. 2019, pp. 1–7.

[31] Q. Zhang, M. Bakshi, and S. Jaggi, "Covert Communication with Polynomial Computational Complexity," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1354–1384, Mar. 2020.

[32] S.-Y. Wang and M. R. Bloch, "Explicit Design of Provably Covert Channel Codes," in *Proc. of IEEE International Symposium on Information Theory*, Melbourne, Australia, Jul. 2021, pp. 190–195.

[33] M. R. Bloch and S. Guha, "Optimal covert communications using pulse-position modulation," in *proc. of IEEE International Symposium on Information Theory*, Aachen, Germany, Jun. 2017, pp. 2825–2829.

[34] M. Tahmasbi and M. R. Bloch, "First- and Second-Order Asymptotics in Covert Communication," *IEEE Transactions on Information Theory*, vol. 65, no. 4, pp. 2190–2212, Apr. 2019.

[35] B. A. Bash, C. N. Gagatsos, A. Datta, and S. Guha, "Fundamental limits of quantum-secure covert optical sensing," in *Proc. of IEEE International Symposium on Information Theory*, Jun. 2017, pp. 3210–3214.

[36] M. Tahmasbi and M. R. Bloch, "Active Covert Sensing," in *Proc. of 2020 IEEE International Symposium on Information Theory*, Los Angeles, CA, Jun. 2020, pp. 840–845.

[37] ——, "On Covert Quantum Sensing and the Benefits of Entanglement," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 352–365, Mar. 2021.

[38] M.-C. Chang and M. R. Bloch, "Covert Online Decision Making: From Sequential Hypothesis Testing to Stochastic Bandits," *arXiv preprint*, vol. 2311.12176, Nov. 2023.

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

[40] 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.

[41] M. Hayashi, "Discrimination of Two Channels by Adaptive Methods and Its Application to Quantum System," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3807–3820, Aug. 2009.