Covert Authentication Against a Myopic Adversary

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

2021 IEEE International Symposium on Information Theory (ISIT) | 978-1-5386-8209-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISIT45174.2021.9518165

***Abstract*—We consider the problem of authenticating communication over a Myopic Binary Adversarial Channel (MBAC) while maintaining covertness with respect to the myopic adversary. When the main channel between legitimate parties is degraded with respect to the adversary's channel, we show the existence of an integrated scheme that simultaneously exploits secret keys to ensure covertness and authentication. The main technical challenge we address is showing that authentication may be ensured against myopic attacks when using the low-weight codewords mandated by covert communication.**

I. Introduction

Authentication is concerned with the problem of ensuring that legitimate parties, hereafter named Alice and Bob, may not only communicate reliably but also ensure that their messages are not impersonated or substituted by a malicious adversary, hereafter named Willie. Several information-theoretic models for authentication have been considered in the literature to date, leading to different conclusions regarding the need of and use of secret keys to ensure authentication. One model of malicious adversary, which is the one investigated here, is an arbitrarily-varying channel in which Willie chooses a non-causal state sequence based on noisy observations of the transmission. The arbitrarily-varying channel can capture both the problem of *jamming*, in which case Alice and Bob try to ensure reliability against a class of adversarial state sequences [1]–[3], or *authentication*, in which case Alice and Bob can afford to not transmit reliably if they correctly detect the presence of an adversarial state [4], [5]. While ensuring authentication is perhaps a simpler task than ensuring resilience to jamming, the design and analysis of coding schemes ensuring authentication is a challenging endeavor in its own right. While state sequences are often subject to some power constraint in the jamming model, they are not in the authentication model, so that neither model necessarily supersedes the other. A key finding of [1] is that, when a jamming adversary is myopic enough, i.e., its observations are noisy enough, the adversary is effectively blind, i.e., it can do no better than if it were oblivious to the actual transmission. In the context of authentication, a key contribution of [4], [5] is the introduction of a condition, called *overwritability*, which characterizes when the authentication capacity of a channel is zero without secret keys. In addition, when the adversary is myopic enough, the authentication capacity of a binary Myopic Binary Adversarial Channel (MBAC) is shown to be equal to the regular capacity. Another model of authentication is studied in [6], [7], in which the adversary can intercept the transmitted codeword and substitute it for any another sequence. The probability of successful attacks is characterized in the presence of and absence of shared secret keys.

This work was supported in part by NSF grant 1955401

Ensuring covertness, or low probability of detection, is another concern in communication systems. The covert capacity of discrete memoryless and Gaussian channels has been thoroughly analyzed [8]–[12], highlighting the presence of a square root law and the need to use codewords with vanishing weight. The need for vanishing weight naturally questions the resilience of covert communication schemes to adversarial interference, and covert communication in the presence of malicious adversaries has already attracted some attention. Notably, [13] studies the covert communication over adversarially jammed channels and shows some degree of resilience. A similar conclusion was drawn for the more specific problem of covert key generation over adversarial channels [14]. In the present work, we investigate the problem of *covert authentication* over an MBAC. Our main contribution is to show the existence of an integrated scheme that exploits shared secret keys jointly for authentication and covertness, and is resilient to a myopic adversary without sharing extra keys except for those used already needed for covertness.

The rest of the paper is organized as follows. After a brief review of notation (Section II), we introduce the specific channel model for covert authentication in the presence of a myopic adversary (Section III) and derive our main result regarding the existence and performance of an integrated covert and authenticated scheme (Section IV). While we provide proof details in appendices, some technical details are omitted because of space constraints.

II. Notations

For two distributions P, Q on some common alphabet $\mathcal{X}$, $D(P\|Q) \triangleq \sum_x P(x) \log \frac{P(x)}{Q(x)}$ is the Kullback-Leibler (KL) divergence between P and Q. We say P is absolutely continuous with respect to (w.r.t.) Q, denoted by $P \ll Q$, if for all $x \in \mathcal{X}$ $P(x) = 0$ if $Q(x) = 0$. We denote $P^{\otimes n}$ the product distribution $\prod_{\ell=1}^n P$ on $\mathcal{X}^n$. We also define $\chi_2(P\|Q) \triangleq \sum_x \frac{(P(x)-Q(x))^2}{Q(x)}$. For any $\mathbf{x} \in \mathcal{X}^n$, $f_{\mathbf{x}}$ is the type of the sequence $\mathbf{x}$. Similarly, if $\mathbf{z} \in \mathcal{Z}^n$ for some alphabet $\mathcal{Z}$, then $f_{\mathbf{x},\mathbf{z}}$ is the joint type of the sequences $\mathbf{x}$ and $\mathbf{z}$. For any $\mathbf{x} \in \mathcal{X}^n$, $x(\ell)$ is its ℓth component of $\mathbf{x}$, and $w_H(\mathbf{x})$ is the hamming weight of $\mathbf{x}$. For any joint distribution P_{XZ} on $\mathcal{X} \times \mathcal{Z}$, X and Z are the random variables with the corresponding distribution, and $I(P_{XZ})$ is the mutual information between X and Z. We also denote τ_X the type class of P_X. The distribution of a Bernoulli random variable

with success probability p is denoted by $B(p)$. We also define the set $[M] = \{1, ..., M\}$. For any real number $a \in \mathbb{R}$, $|a|^+ = \max(a, 0)$.

III. Problem Formulation and Main Result

We use the same adversarial channel model as in [5]. Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{S}$ be the transmitter, receiver and state alphabets, respectively, and $W_{Y|X,S}$ be the discrete memoryless adversarial channel between the legitimate users. For all length n sequences $\mathbf{x} \in \mathcal{X}^n$, $\mathbf{y} \in \mathcal{Y}^n$ and $\mathbf{s} \in \mathcal{S}^n$, the transition probability of $\mathbf{y}$ given $\mathbf{x}$ and $\mathbf{s}$ can be written as

$$W_{Y|X,S}(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{\ell=1}^{n} W_{Y|X,S}(y(\ell)|x(\ell), s(\ell)). \quad (1)$$

Let $\mathcal{Z}$ be the output alphabet at the adversary. The adversary observes the transmission through the discrete memoryless channel $W_{Z|X}$. A *myopic* adversary can choose the state sequence $\mathbf{s}$ by observing the received sequence $\mathbf{z} \in \mathcal{Z}^n$ according to a policy $J_{S^n|Z^n}$, which depends on the channel $W_{Z|X}$ and the codebook used by the legitimate parties. Let $s_0 \in \mathcal{S}$ be the state of no interference, i.e., $W_{Y|X,S=s_0} \triangleq W_{Y|X}$, where $W_{Y|X}$ is the channel between legitimate parties without adversaries. Then, the state sequence $\mathbf{s}_0 \in \mathcal{S}^n$ represents the situation in which there is no adversary. Let M be the size of the message set, so that a message belongs to the set $[M]$. Besides, let K be the size of keys. An authentication code consists of an encoder/decoder pair (ψ, ϕ) with

$$\psi : \{1, ..., M\} \times \{1, ..., K\} \mapsto \mathcal{X}^n, \quad (2)$$

$$\phi : \mathcal{Y}^n \times \{1, ..., K\} \mapsto \{0, 1, ..., M\}, \quad (3)$$

where the decoded symbol 0 indicates the existence of an adversary. The decoder ϕ is successful if either the decoder output is equal to the transmitted message, or if the decoder output is 0 and $\mathbf{s} \neq \mathbf{s}_0$. Therefore, given the transmitted message i, the key k and the corresponding codeword $\mathbf{x}_{i,k} = \psi(i, k)$, the probability of error of the code (ψ, ϕ) under the attacking policy $J_{S^n|Z^n}$ is

$$\begin{aligned} &e(i, k, J_{S^n|Z^n}) \\ &= \sum_{\mathbf{z}} W_{Z|X}(\mathbf{z}|\mathbf{x}_{ik}) J_{S^n|Z^n}(\mathbf{s}_0|\mathbf{z}) W_{Y|X,S}(\phi^{-1}(i)^c|\mathbf{x}_{ik}, \mathbf{s}_0) \\ &+ \sum_{\mathbf{s}\neq\mathbf{s}_0,\mathbf{z}} W_{Z|X}(\mathbf{z}|\mathbf{x}_i) J_{S^n|Z^n}(\mathbf{s}|\mathbf{z}) W_{Y|X,S}(\phi^{-1}(\{i, 0\})^c|\mathbf{x}_{ik}, \mathbf{s}), \end{aligned}$$

where $\phi^{-1}(A)$ denotes the set of channel outputs that are decoded, with the information of the shared key k, to an element in the set A, and $\phi^{-1}(A)^c$ is the complement of the set in $\mathcal{Y}^n$. Assuming that every message is transmitted with equal probability, then the average probability of error for the policy $J_{S^n|Z^n}$ is

$$e(J_{S^n|Z^n}) = \frac{1}{M}\frac{1}{K}\sum_{i=1}^{M}\sum_{k=1}^{K} e(i, k, J_{S^n|Z^n}). \quad (4)$$

The objective of Alice and Bob is not only to communicate reliably over the DMC $W_{Y|X}$ but also to avoid detection by Willie. We assume next that $\mathcal{X} = \{0, 1\}$ where the symbol 0 corresponds to the input to the channel when no communication takes place. In the absence of communication, the output distributions of the two channels induced by the symbol 0 are then given by

$$P_0 = W_{Y|X=0} \quad \text{and} \quad Q_0 = W_{Z|X=0}. \quad (5)$$

If symbol 1 is transmitted, the output distributions of the two channels are given by

$$P_1 = W_{Y|X=1} \quad \text{and} \quad Q_1 = W_{Z|X=1}. \quad (6)$$

We assume that $Q_1 \ll Q_0$ and $P_1 \ll P_0$ to simplify the analysis. Willie can perform any hypothesis test on his observation $\mathbf{z}$ to decide whether Alice and Bob communicate (hypothesis H_1) or not (hypothesis H_0). Let α be the type I error probability (rejecting H_0 when true) and β be the type II error probability (accepting H_0 when wrong). It has been shown in [15] that Willie's optimal hypothesis test satisfies the tradeoff $\alpha + \beta \geqslant 1 - \sqrt{D(\widehat{Q}^n \| Q_0^{\otimes n})}$, where $\widehat{Q}^n$ is the expected distribution of $\mathbf{z}^n$ when the communication takes place, and $Q_0^{\otimes n} = \prod_{\ell=1}^{n} Q_0$ is the distribution of $\mathbf{z}^n$ when no communication happens. Therefore, covert communication is achieved by making the divergence $D(\widehat{Q}^n \| Q_0^{\otimes n})$ small. Hence, the objectives of Alice and Bob are two-fold: the transmitted codeword $\mathbf{x}$ should be correctly decoded from the interfered noisy sequence $\mathbf{y}$; Willie should not identify the presence of a transmission. To be more specific, the code design needs to satisfy

$$\lim_{n\to\infty} \max_{J} e(J) = 0 \quad (7)$$

$$\lim_{n\to\infty} D(\widehat{Q}^n \| Q_0^{\otimes n}) < \eta \quad (8)$$

for some $\eta > 0$, where η is the parameter reflecting how covert the communication should be. We assume that the channel between Alice and Willie is a binary symmetric channel with crossover probability q and denoted by BSC_q. The adversary Willie can decide on a state sequence $\mathbf{s}$ according to some policy $J(\mathbf{s}|\mathbf{z})$ and adds it to the transmitted codeword $\mathbf{x}$. The interfered sequence $\mathbf{x} \oplus \mathbf{s}$ is passed through another binary symmetric channel BSC_p with crossover probability p, and the channel output $\mathbf{y}$ is received by Bob. Therefore, $W_{Y|X,S}(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{l=1}^{n} W_{Y|X}(y(\ell)|x(\ell) \oplus s(\ell))$, and $W_{Y|X}$ is a binary symmetric channel with cross over probability p. This myopic binary adversarial channel with parameter p and q is denoted by $\text{MBAC}_{p,q}$. For all $0 \leqslant \alpha_n \leqslant 1$, we define the distribution Π_{α_n} on $\mathcal{X}$ such that $\Pi_{\alpha_n}(1) = 1 - \Pi_{\alpha_n}(0) = \alpha_n$ as well as the distribution

$$Q_{\alpha_n}(z) = \alpha_n Q_1(z) + (1 - \alpha_n) Q_0(z) \quad (9)$$

for all $z \in \mathcal{Z}$. Our main theorem is stated below.

Theorem 1. *Consider the MBAC with parameter $1/2 > p > q$. Let $\omega = \sqrt{\frac{2\eta}{\chi_2(Q_1\|Q_0)}}$, and $\alpha_n = \frac{\omega}{\sqrt{n}}$. For any $\delta > 0$, there exist $\xi_1, \xi_2 > 0$ and a coding scheme such that*

$$\frac{\log M}{\sqrt{n}} = \omega(D(P_1\|P_0) - \delta) \quad (10)$$

$$\frac{\log K}{\sqrt{n}} = \omega(D(Q_1\|Q_0) - D(P_1\|P_0) + 2\delta), \tag{11}$$

and

$$\max_J e(J) < e^{-\sqrt{n}\omega\xi_1} \tag{12}$$

$$|D(\widehat{Q}^n\|Q_0^{\otimes n}) - D(Q_{\alpha_n}^{\otimes n}\|Q_0^{\otimes n})| < e^{-\sqrt{n}\omega\xi_2} \tag{13}$$

when n is large enough.

By Theorem 1, there exists a code such that (7) is satisfied when $n \to \infty$, and the divergence satisfies

$$\begin{aligned} D(\widehat{Q}^n\|Q_0^{\otimes n}) &\leqslant |D(\widehat{Q}^n\|Q_0^{\otimes n}) - D(Q_{\alpha_n}^{\otimes n}\|Q_0^{\otimes n})| + D(Q_{\alpha_n}^{\otimes n}\|Q_0^{\otimes n}) \\ &\leqslant e^{-\omega\xi_2\sqrt{n}} + n \times \left(\frac{\alpha_n^2}{2}\chi_2(Q_1\|Q_0)(1+o(1))\right) \quad (14) \\ &\leqslant e^{-\omega\xi_2\sqrt{n}} + \eta(1+o(1)), \quad (15) \end{aligned}$$

where (14) is from Lemma 1 of [11]. Therefore, $\lim_{n\to\infty} D(\widehat{Q}^n\|Q_0^{\otimes n}) \leqslant \eta$. Theorem 1 shows that there exists an integrated scheme that jointly ensures covertness and authentication with the same asymptotically optimal message throughput and secret key throughput as schemes that only ensure covertness. In other words, any authentication put in place at the message level is automatically strengthened by the intrinsic authentication provided by the integrated scheme.

IV. Proof of Main result

A. Encoder and Decoder Design

Let $\omega = \sqrt{\frac{2\eta}{\chi_2(Q_1\|Q_0)}}$ and define the relative rate of messages and keys for any $\delta > 0$ as

$$\frac{\log M}{\sqrt{n}} = \omega(D(P_1\|P_0) - \delta), \tag{16}$$

$$\frac{\log K}{\sqrt{n}} = \omega(D(Q_1\|Q_0) - D(P_1\|P_0) + 2\delta), \tag{17}$$

so that the total number of codewords $N = MK = 2^{\sqrt{n}\omega(D(Q_1\|Q_0)+\delta)}$ is large enough to form a resolvability code for the channel $W_{Z|X}$. We denote by $\mathbf{x}_{ik}$ the codeword when the message/key pair is (i,k). Let $\alpha_n = \frac{\omega}{\sqrt{n}}$. The codewords $\mathbf{x}_{11}, ..., \mathbf{x}_{MK}$ are drawn from the product distribution $\Pi_{\alpha_n}^{\otimes n}$, and we denote by $\mathbb{P}_C$ the probability measure induced by this random coding. Then, by [11], for n large enough

$$\mathbb{E}_C\left\{D(\widehat{Q}^n\|Q_{\alpha_n}^{\otimes n})\right\} < e^{-\rho_3\omega\sqrt{n}} \tag{18}$$

for some $\rho_3 > 0$, and

$$\mathbb{P}_C\left\{D(\widehat{Q}^n\|Q_{\alpha_n}^{\otimes n}) > e^{-0.5\rho_3\omega\sqrt{n}}\right\} \leqslant e^{-0.5\rho_3\omega\sqrt{n}}. \tag{19}$$

Let $\mathbf{z}$ be the received sequence of Willie. For each sequence $\mathbf{z} \in \{0,1\}^n$ and each joint distribution P_{XZ} on $\mathcal{X} \times \mathcal{Z}$, we define $S(\mathbf{z}, P_{XZ})$ as

$$S(\mathbf{z}, P_{XZ}) \triangleq \{(i,k) \in [M] \times [K] : f_{\mathbf{x}_{ik},\mathbf{z}} = P_{XZ}\}. \tag{20}$$

Furthermore, for any $\epsilon > 0$, we define the set A_ϵ as

$$\begin{aligned} A_\epsilon = \{P_{XZ} : |P_X(1) - \frac{\omega}{\sqrt{n}}| \leqslant \frac{\omega\epsilon}{\sqrt{n}}, \\ |P_{Z|X=1}(0) - q| \leqslant \epsilon, |P_{Z|X=0}(1) - q| \leqslant \epsilon\} \end{aligned} \tag{21}$$

so that whenever $P_{XZ} \in A_\epsilon$, the distribution P_{XZ} is typical. When $P_{XZ} \in A_\epsilon$, we obtain the following useful inequalities. First,

$$\begin{aligned} I(P_{XZ}) &= P_X(1)D\left(P_{Z|X=1}\|P_{Z|X=0}\right) - D\left(P_Z\|P_{Z|X=0}\right) \quad (22) \\ &\leqslant P_X(1)D\left(B(1-q+\epsilon)\|B(q-\epsilon)\right) + O(1/n) \quad (23) \\ &\leqslant \frac{\omega}{\sqrt{n}}(1+\epsilon)\left(D(Q_1\|Q_0) + \epsilon O(1)\right) + O(1/n) \quad (24) \\ &\leqslant \frac{\omega}{\sqrt{n}}D(Q_1\|Q_0) + \frac{\omega}{\sqrt{n}}\epsilon O(1), \quad (25) \end{aligned}$$

where (23) follows from the definition of A_ϵ and the fact that $D(P_Z\|P_{Z|X=0}) \leqslant O(\alpha_n^2) = O(1/n)$ by (13) in [11]. Secondly,

$$\begin{aligned} &D(P_X\|\Pi_{\alpha_n}) \\ &\leqslant \Pi_{\alpha_n}(1)(1+\epsilon)\log\frac{\Pi_{\alpha_n}(1)(1+\epsilon)}{\Pi_{\alpha_n}(1)} + \log\frac{1-\Pi_{\alpha_n}(1)(1-\epsilon)}{1-\Pi_{\alpha_n}(1)} \quad (26) \\ &\leqslant \Pi_{\alpha_n}(1)(1+\epsilon)\log(1+\epsilon) + \log\left(1 + \frac{\Pi_{\alpha_n}(1)\epsilon}{1-\Pi_{\alpha_n}(1)}\right) \quad (27) \\ &\leqslant \epsilon\Pi_{\alpha_n}(1)O(1), \quad (28) \end{aligned}$$

where (28) follows because $\log(1+x) \leqslant x - \frac{x^2}{2} + \frac{x^3}{3}$ for $x > -1$. The inequality (25) shows that the mutual information induced by the joint type P_{XZ} is not far from $\frac{\omega}{\sqrt{n}}D(Q_1\|Q_0)$, and (28) says the divergence between P_X and Π_{α_n} is smaller than $O(\epsilon\frac{\omega}{\sqrt{n}})$. Furthermore, we can upper bound the number of codewords whose types do not satisfy the first condition in (21). By the Chernoff bound,

$$\mathbb{P}_C\left\{\left|\frac{w_H(\mathbf{x}_{ik})}{n} - \frac{\omega}{\sqrt{n}}\right| > \frac{\epsilon\omega}{\sqrt{n}}\right\} \leqslant 2\exp\left\{-\frac{\epsilon^2\omega\sqrt{n}}{3}\right\} \tag{29}$$

for all (i,k) pairs. Then, by Markov's inequality

$$\begin{aligned} &\mathbb{P}_C\left\{\sum_{i=1}^{M}\sum_{k=1}^{K} 1\left(\left|\frac{w_H(\mathbf{x}_{ik})}{n} - \frac{\omega}{\sqrt{n}}\right| > \frac{\epsilon\omega}{\sqrt{n}}\right) > 2MK\exp\left(-\frac{\epsilon^2\omega\sqrt{n}}{6}\right)\right\} \\ &\leqslant \exp\left(-\frac{\epsilon^2\omega\sqrt{n}}{6}\right). \end{aligned} \tag{30}$$

Next, we introduce the following two lemmas, which are adapted from Lemma IV.2 of [16] and Lemma 3 of [17] whose proofs are omitted.

Lemma 2. *For any $\delta > 0$, let $N = 2^{\sqrt{n}\omega(D(Q_1\|Q_0)+\delta)}$ be the number of codewords and let codewords $\{\mathbf{x}_i\}_{i=1}^N$ be drawn independently from the product distribution $\Pi_{\alpha_n}^{\otimes n}$. There exists $\epsilon, \epsilon_1 > 0$ small enough such that when $P_{XZ} \in A_\epsilon$, the number of codewords in the set $S(\mathbf{z}, P_{XZ})$ is lower bounded by*

$$N2^{-nI(P_{XZ})-nD(P_X\|\Pi_{\alpha_n})-\epsilon_1\sqrt{n}} \tag{31}$$

with probability at least $1 - e^{-e^{\rho_1\sqrt{n}}}$ for some $\rho_1 > 0$ when n is large enough.

Lemma 3. *Let $\epsilon_2 > 0$ and let $\mathbf{x}_1, ..., \mathbf{x}_N$ be drawn independently from the product distribution $\Pi_{\alpha_n}^{\otimes n}$, where $N = 2^{nR}$. With probability at least $1 - e^{-e^{\rho_2\sqrt{n}}}$ for some $\rho_2 > 0$, this code satisfies the following. For any type class $\tau_{XX'S}$, any sequence $\mathbf{s} \in \mathcal{S}^n$ and any alphabet set $\mathcal{S}$,*

$$\begin{aligned} &|\{i : \exists j \neq i \; s.t \; (\mathbf{x}_i, \mathbf{x}_j, \mathbf{s}) \in \tau_{XX'S}\}| \\ &\leqslant 2^{n|R-I(X;X'S)-D(P_X\|\Pi_{\alpha_n})+|R-I(X';S)-D(P_{X'}\|\Pi_{\alpha_n})|^+|^+ + \epsilon_2 n^{1/2}}. \end{aligned} \tag{32}$$

By Lemma 2, Lemma 3, (19) and (30), for n large enough, there exist codewords $\mathbf{x}_{11}, ..., \mathbf{x}_{MK}$ drawn from the product distribution $\Pi_{\alpha_n}^{\otimes n}$ such that

$$D(\widehat{Q}^n \| Q_{\alpha_n}^{\otimes n}) < e^{-0.5\rho_3 \omega \sqrt{n}} \tag{33}$$

$$\sum_{ik} 1\left(\left|\frac{w_H(\mathbf{x}_{ik})}{n} - \frac{\omega}{\sqrt{n}}\right| > \frac{\epsilon\omega}{\sqrt{n}}\right) < 2MK \exp\left(-\frac{\epsilon^2 \omega \sqrt{n}}{6}\right), \tag{34}$$

and (31) and (32) are satisfied for all joint type $P_{XZ} \in A_\epsilon$ because the number of joint types in A_ϵ is at most polynomial in n. When $D(\widehat{Q}^n \| Q_{\alpha_n}^{\otimes n}) < e^{-0.5\rho_3\omega\sqrt{n}}$, one can show that

$$|D(\widehat{Q}^n \| Q_0^{\otimes n}) - D(Q_{\alpha_n}^{\otimes n} \| Q_0^{\otimes n})| < e^{-\sqrt{n}\omega\xi_2}$$

for some $\xi_2 > 0$ by [11, (77)-(79)]. Therefore, it remains to analyze the error probability $e(J)$. We define our encoding/decoding rules as follows.

- Encoding: For each message $i \in [M]$ and key $k \in [K]$, transmit $\mathbf{x}_{ik}$.
- Decoding: For each received sequence $\mathbf{y}$ and given the shared key k, decode to a message i if i is the unique index with the properties

$$p - \epsilon \leqslant \bar{d}_1(\mathbf{y}, \mathbf{x}_{ik}) \triangleq \frac{\sum_{\ell=1}^n 1(x_{ik}(\ell) = 1, y(\ell) = 0)}{\sum_{l=1}^n 1(x_{ik}(\ell) = 1)} \leqslant p + \epsilon \tag{35}$$

$$p - \epsilon \leqslant \bar{d}_0(\mathbf{y}, \mathbf{x}_{ik}) \triangleq \frac{\sum_{\ell=1}^n 1(x_{ik}(\ell) = 0, y(\ell) = 1)}{\sum_{\ell=1}^n 1(x_{ik}(\ell) = 0)} \leqslant p + \epsilon,$$

$$\left|\frac{w_H(\mathbf{x}_{ik})}{n} - \frac{\omega}{\sqrt{n}}\right| \leqslant \frac{\omega\epsilon}{\sqrt{n}}, \tag{36}$$

  and declare adversarial interference 0 otherwise.

B. Error Probability Analysis

In the analysis below, we fix the code and assume it satisfies (31)-(34). Let $\mathbf{x}$ be the transmitted codeword and $f_{\mathbf{x},\mathbf{z}}$ be the joint type of the transmitted codeword and Willie's received sequence $\mathbf{z}$. We consider a strengthened attacker that knows the noise sequence $\mathbf{n}$ generated from BSC_p and the joint type $f_{\mathbf{x},\mathbf{z}}$. Therefore, Willie can select the noisy attack sequence $\bar{\mathbf{s}} = \mathbf{s} + \mathbf{n}$ according to some policy $J(\bar{\mathbf{s}}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}})$ such that $\mathbf{y} = \mathbf{x} + \bar{\mathbf{s}}$. Since this assumption only makes the attacker stronger, the error probability is upper bounded by the one obtained with this strengthened attacker. For any policy J, we can upper bound the error probability $e(J)$ by

$$e(J) = \sum_{i=1}^M \sum_{k=1}^K \sum_{\mathbf{z}} \sum_{\bar{\mathbf{s}}} \sum_{f_{\mathbf{x},\mathbf{z}}} \mathbb{P}(\text{error}, \mathbf{x} = \mathbf{x}_{ik}, \mathbf{z}, \bar{\mathbf{s}}, f_{\mathbf{x},\mathbf{z}}) \tag{37}$$

$$= \sum_{i,k,\mathbf{z},\bar{\mathbf{s}},f_{\mathbf{x},\mathbf{z}}} \mathbb{P}(\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \mathbb{P}(\bar{\mathbf{s}}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \mathbb{P}(\mathbf{x} = \mathbf{x}_{ik}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \times \mathbb{P}(\text{error}|\mathbf{x} = \mathbf{x}_{ik}, \bar{\mathbf{s}}, \mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \tag{38}$$

$$\leqslant \sum_{i,k,\mathbf{z},\bar{\mathbf{s}},f_{\mathbf{x},\mathbf{z}}} \mathbb{P}(\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) J(\bar{\mathbf{s}}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \mathbb{P}(\mathbf{x} = \mathbf{x}_{ik}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \times \mathbb{P}(\text{error}|\mathbf{x} = \mathbf{x}_{ik}, \bar{\mathbf{s}}, \mathbf{z}, f_{\mathbf{x},\mathbf{z}}) + \mathbb{P}(f_{\mathbf{x},\mathbf{z}} \notin A_\epsilon). \tag{39}$$

By a union bound, the term $\mathbb{P}(f_{\mathbf{x},\mathbf{z}} \notin A_\epsilon)$ on the right hand side of (39) can be upper bounded by

$$\mathbb{P}(f_{\mathbf{x},\mathbf{z}} \notin A_\epsilon) \leqslant \frac{1}{MK} \sum_{i=1}^M \sum_{k=1}^K 1\left(\left|\frac{w_H(\mathbf{x}_{ik})}{n} - \frac{\omega}{\sqrt{n}}\right| > \frac{\epsilon\omega}{\sqrt{n}}\right)$$

$$+ \mathbb{P}\left(\left|\frac{\sum_{\ell:\mathbf{x}(\ell)=1} 1(\mathbf{x}(\ell) = 1, \mathbf{z}(\ell) = 0)}{w_H(\mathbf{x})} - q\right| > \epsilon \middle| \left|\frac{w_H(\mathbf{x})}{n} - \frac{\omega}{\sqrt{n}}\right| \leqslant \frac{\epsilon\omega}{\sqrt{n}}\right)$$

$$+ \mathbb{P}\left(\left|\frac{\sum_{\ell:\mathbf{x}(\ell)=0} 1(\mathbf{x}(\ell) = 0, \mathbf{z}(\ell) = 1)}{n - w_H(\mathbf{x})} - q\right| > \epsilon \middle| \left|\frac{w_H(\mathbf{x})}{n} - \frac{\omega}{\sqrt{n}}\right| \leqslant \frac{\epsilon\omega}{\sqrt{n}}\right) \tag{40}$$

$$\leqslant 2\exp\left(\frac{-\epsilon^2\omega\sqrt{n}}{6}\right) + 2\exp\left(-2\epsilon^2 w_H(\mathbf{x})\right) + 2\exp\left(-2\epsilon^2(n - w_H(\mathbf{x}))\right) \tag{41}$$

$$\leqslant \exp(-c_1\omega\sqrt{n}) \tag{42}$$

for some $c_1 > 0$, where (41) follows from the assumption of the code and Hoeffding's inequality. Therefore, it suffices to upper bound the summation terms in (39). We first fix the sequences $\mathbf{z}$ and the joint type $f_{\mathbf{x},\mathbf{z}}$, and upper bound the sum $\sum_{\bar{\mathbf{s}}} J(\bar{\mathbf{s}}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \sum_{i,k} \mathbb{P}(\mathbf{x} = \mathbf{x}_{ik}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \mathbb{P}(\text{error}|\mathbf{x} = \mathbf{x}_{ik}, \bar{\mathbf{s}}, \mathbf{z}, f_{\mathbf{x},\mathbf{z}})$. Define the set $E(\bar{\mathbf{s}})$ as

$$E(\bar{\mathbf{s}}) \triangleq \Big\{(i,k) : \exists j \neq i \ s.t \left|\frac{w_H(\mathbf{x}_{jk})}{n} - \frac{\omega}{\sqrt{n}}\right| \leqslant \frac{\omega\epsilon}{\sqrt{n}}, |\bar{d}_1(\mathbf{x}_{ik} \oplus \bar{\mathbf{s}}, \mathbf{x}_{jk}) - p| < \epsilon, |\bar{d}_0(\mathbf{x}_{ik} \oplus \bar{\mathbf{s}}, \mathbf{x}_{jk}) - p| < \epsilon\Big\}.$$

Notice that if $\bar{\mathbf{s}} \neq \mathbf{n}$, the message/key pair $(i,k) \notin E(\bar{\mathbf{s}})$ implies that only i or the adversarial interference symbol 0 can be decoded, and no error occurs in both cases. On the other hand, if $\bar{\mathbf{s}} = \mathbf{n}$, $\mathbf{x}_{ik} \oplus \bar{\mathbf{s}} = \mathbf{x}_{ik} \oplus \mathbf{n}$. Then, the fact that the message/key pair $(i,k) \notin E(\bar{\mathbf{s}})$ guarantees that no other message $j \neq i$ can be decoded. Furthermore, by definition of the decoder and the use of concentration inequalities on the noise sequence $\mathbf{n}$, the message i is decoded with probability at least $1 - e^{-c_2\omega\sqrt{n}}$ for some $c_2 > 0$ when $f_{\mathbf{x},\mathbf{z}} \in A_\epsilon$. Therefore,

$$\sum_{\bar{\mathbf{s}},(i,k)\notin E(\bar{\mathbf{s}})} J(\bar{\mathbf{s}}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \mathbb{P}(\mathbf{x} = \mathbf{x}_{ik}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \mathbb{P}(\text{error}|\mathbf{x} = \mathbf{x}_{ik}, \bar{\mathbf{s}}, \mathbf{z}, f_{\mathbf{x},\mathbf{z}}) \leqslant e^{-c_2\omega\sqrt{n}} \tag{43}$$

for some $c_2 > 0$. Next, we consider the case when $(i,k) \in E(\bar{\mathbf{s}})$ for some fixed $\bar{\mathbf{s}}$. Notice that

$$\mathbb{P}(\mathbf{x} = \mathbf{x}_{ik}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}} = P_{XZ}) = \begin{cases} \frac{1}{|S(\mathbf{z}, P_{XZ})|} & \text{if } (i,k) \in S(\mathbf{z}, P_{XZ}) \\ 0 & \text{otherwise} \end{cases}$$

because all the codewords in $S(\mathbf{z}, P_{XZ})$ have the same distance from $\mathbf{z}$. Therefore, for each $\mathbf{z}, \bar{\mathbf{s}}$ and joint type $f_{\mathbf{x},\mathbf{z}} = P_{XZ}$, the sum $\sum_{(i,k)\in E(\bar{\mathbf{s}})} \mathbb{P}(\mathbf{x} = \mathbf{x}_{ik}|\mathbf{z}, f_{\mathbf{x},\mathbf{z}} = P_{XZ}) \mathbb{P}(\text{error}|\mathbf{x} = \mathbf{x}_{ik}, \bar{\mathbf{s}}, \mathbf{z}, f_{\mathbf{x},\mathbf{z}} = P_{XZ})$ can upper bounded by

$$\sum_{(i,k)\in S(\mathbf{z},P_{XZ})\cap E(\bar{\mathbf{s}})} \frac{\mathbb{P}(\text{error}|\mathbf{x} = \mathbf{x}_{ik}, \bar{\mathbf{s}}, \mathbf{z}, f_{\mathbf{x},\mathbf{z}} = P_{XZ})}{|S(\mathbf{z}, P_{XZ})|} \leqslant \frac{|S(\mathbf{z}, P_{XZ}) \cap E(\bar{\mathbf{s}})|}{|S(\mathbf{z}, P_{XZ})|},$$

where we upper bound $\mathbb{P}(\text{error}|\mathbf{x} = \mathbf{x}_{ik}, \bar{\mathbf{s}}, \mathbf{z}, f_{\mathbf{x},\mathbf{z}} = P_{XZ})$ by 1 if $(i,k) \in E(\bar{\mathbf{s}})$. The size of the set $S(\mathbf{z}, P_{XZ}) \cap E(\bar{\mathbf{s}})$ can be upper bounded by the lemma below.

Lemma 4. *Let $P_{XZ} \in A_\epsilon$, and $\epsilon_2 > 0$ be defined in Lemma 3, it holds with probability at least $1 - e^{-\rho_4\omega\sqrt{n}}$ for some $\rho_4 > 0$ that*

$$|E(\bar{\mathbf{s}}) \cap S(\mathbf{z}, P_{XZ}))| \leqslant 2^{\sqrt{n}\omega D(Q_1\|Q_0) - nI(P_{XZ}) + 0.5\omega\sqrt{n}\delta + O(\epsilon\omega\sqrt{n}) + 2n^{1/2}\epsilon_2} \tag{44}$$

when n is sufficiently large.

By a union bound, the probability that (31), (33), (34) and (44) are satisfied is greater than $1-e^{-\rho\omega\sqrt{n}}$ for some $\rho>0$. Hence, there exists a code that satisfies (31), (33), (34) and (44). Then, by combining Lemma 2 and Lemma 4, this specific code satisfies

$$\frac{|S(\mathbf{z},P_{XZ})\cap E(\mathbf{s})|}{|S(\mathbf{z},P_{XZ})|}\leqslant 2^{-0.5\omega\sqrt{n}\delta+O(\epsilon\omega\sqrt{n})+n^{1/2}\epsilon_1+2n^{1/2}\epsilon_2}.$$

For each $\delta>0$, we can choose ϵ,ϵ_1 and ϵ_2 small enough such that $-0.5\omega\sqrt{n}\delta+O(\epsilon\omega\sqrt{n})+n^{1/2}\epsilon_1+2n^{1/2}\epsilon_2$ is negative, and

$$\frac{|S(\mathbf{z},P_{XZ})\cap E(\mathbf{s})|}{|S(\mathbf{z},P_{XZ})|}\leqslant e^{-c_3\omega\sqrt{n}}$$

for some $c_3>0$ when n is large enough. Then,

$$\begin{aligned}e(J)&\leqslant e^{-c_1\omega\sqrt{n}}+e^{-c_2\omega\sqrt{n}}+e^{-c_3\omega\sqrt{n}} &(45)\\ &\leqslant e^{-\xi_1\omega\sqrt{n}} &(46)\end{aligned}$$

for some $\xi_1>0$ when n is sufficient large. It remains to prove Lemma 4, as outlined in Appendix A.

Appendix A
Proof of Lemma 4

We first define the set $\tilde{E}(\bar{\mathbf{s}})$ as

$$\tilde{E}(\bar{\mathbf{s}})\triangleq\Big\{(i,k):\exists(j,l)\neq(i,k)\ s.t\left|\frac{w_H(\mathbf{x}_{jl})}{n}-\frac{\omega}{\sqrt{n}}\right|\leqslant\frac{\omega\epsilon}{\sqrt{n}},\\ \left|\bar{d}_1(\mathbf{x}_{ik}\oplus\bar{\mathbf{s}},\mathbf{x}_{jl})-p\right|<\epsilon,\ \left|\bar{d}_0(\mathbf{x}_{ik}\oplus\bar{\mathbf{s}},\mathbf{x}_{jl})-p\right|<\epsilon\Big\}.$$

$\tilde{E}(\bar{\mathbf{s}})$ is the set of message/key pairs (i,k) for which there exists a pair $(j,l)\neq(i,k)$, without the restriction of $l=k$, such that the decoding rules are satisfied. By splitting into different type classes, we can upper bound $S(\mathbf{z},P_{XZ})\cap\tilde{E}(\bar{\mathbf{s}})$ by

$$\begin{aligned}&|S(\mathbf{z},P_{XZ})\cap\tilde{E}(\bar{\mathbf{s}})|\\ &\leqslant\sum_{X',X'',S,Z'}|\{(i,k):\exists(j,l)\neq(i,k)\ s.t\ (\mathbf{x}_{ik},\mathbf{x}_{jl},\bar{\mathbf{s}},\mathbf{z})\in\tau_{X'X''SZ'}\}|,\end{aligned}$$

where X',X'',S, and Z' are dummy random variables describing each type class and satisfying following inequalities

$$\begin{aligned}|P_{X'}(1)-\frac{\omega}{\sqrt{n}}|\leqslant\frac{\omega\epsilon}{\sqrt{n}},\quad |P_{X''}(1)-\frac{\omega}{\sqrt{n}}|,\leqslant\frac{\omega\epsilon}{\sqrt{n}}\\ |P_{S\oplus X'|X''=1}-p|\leqslant\epsilon,\quad |P_{S\oplus X'|X''=0}-p|,\leqslant\epsilon,\\ P_{X'Z'}=P_{XZ}.\end{aligned}$$

By viewing $(\bar{\mathbf{s}},\mathbf{z})$ as $\mathbf{s}$ in Lemma 3 and assuming (32) holds, we have

$$|\{(i,k):\exists(j,l)\neq(i,k)\ s.t\ (\mathbf{x}_{ik},\mathbf{x}_{jl},\bar{\mathbf{s}},\mathbf{z})\in\tau_{X'X''SZ'}\}|\leqslant 2^{n\gamma+\epsilon_2 n^{1/2}},$$

where we define $\gamma\triangleq|R-I(X';X''SZ')-D(P_{X'}\|\Pi_{\alpha_n})+|R-I(X'';SZ')-D(P_{X''}\|\Pi_{\alpha_n})|^+|^+$. To calculate the value of γ, we need to consider the two cases, $R\geqslant I(X'';SZ')+D(P_{X''}\|\Pi_{\alpha_n})$ and $R<I(X'';SZ)+D(P_{X''}\|\Pi_{\alpha_n})$.

Case 1: $R\geqslant I(X'';SZ')+D(P_{X''}\|\Pi_{\alpha_n})$

$$\begin{aligned}\gamma&\leqslant|2R-I(X';SZ')-I(X';X''|SZ')-I(X'';SZ')-D(P_{X'}\|\Pi_{\alpha_n})|^+\\ &\leqslant|2R-I(X';Z')-I(X'';X'\oplus S)-D(P_{X'}\|\Pi_{\alpha_n})|^+ &(47)\end{aligned}$$

by data processing inequalities and

$$\begin{aligned}&I(X'';X'\oplus S)\\ &=P_{X''}(1)D(P_{X'\oplus S|X''=1}\|P_{X'\oplus S|X''=0})-D(P_{X'\oplus S}\|P_{X'\oplus S|X''=0})\\ &\geqslant\frac{\omega}{\sqrt{n}}(1-\epsilon)D(B(1-p-\epsilon)\|B(p+\epsilon))-O(1/n)\\ &\geqslant\frac{\omega}{\sqrt{n}}(1-\epsilon)\left(D(P_1\|P_0)-0.4\delta\right)\end{aligned}$$

when we choose ϵ small enough and n is sufficient large. Then,

$$\begin{aligned}\gamma&\leqslant\frac{2\omega}{\sqrt{n}}(D(Q_1\|Q_0)+\delta)-I(P_{XZ})\\ &\quad-\frac{\omega}{\sqrt{n}}(1-\epsilon)\left(D(P_1\|P_0)-0.4\delta\right)-D(P_{X'}\|\Pi_{\alpha_n}), &(48)\end{aligned}$$

where (48) is positive when we choose ϵ sufficient small because $D(Q_1\|Q_0)>D(P_1\|P_0)$, $P_{X'Z'}=P_{XZ}$, $I(P_{XZ})\leqslant\frac{\omega}{\sqrt{n}}D(Q_1\|Q_0)+O(\epsilon\frac{\omega}{\sqrt{n}})$, and $D(P_X\|\Pi_{\alpha_n})\leqslant O(\epsilon\frac{\omega}{\sqrt{n}})$ by (25) and (28).

Case 2: $R<I(X'';SZ')+D(P_{X''}\|\Pi_{\alpha_n})$

$$\begin{aligned}\gamma&\leqslant|R-I(X';Z')-D(P_{X'}\|\Pi_{\alpha_n})|^+ &(49)\\ &=\frac{\omega}{\sqrt{n}}(D(Q_1\|Q_0)+\delta)-I(P_{X'Z'})-D(P_{X'}\|\Pi_{\alpha_n}), &(50)\end{aligned}$$

where (50) is positive for the same reason as (48) when ϵ is small enough. The difference between (50) and (48) is the term $\frac{\omega}{\sqrt{n}}(D(Q_1\|Q_0)+\delta)-\frac{\omega}{\sqrt{n}}(1-\epsilon)\left(D(P_1\|P_0)-0.4\delta\right)$, which is always positive. Therefore, in both cases, we can upper bound γ by (48). Then,

$$\begin{aligned}&|S(\mathbf{z},P_{XZ})\cap\tilde{E}(\bar{\mathbf{s}})|\leqslant\sum_{X',X'',S,Z'}2^{n\gamma+n^{1/2}\epsilon_2}\\ &\leqslant 2^{\sqrt{n}\omega[2(D(Q_1\|Q_0)+\delta)-D(P_1\|P_0)+0.4\delta+O(\epsilon)]-nI(f_{XZ})+2\sqrt{n}\epsilon_2} &(51)\end{aligned}$$

when (32) holds and n is sufficient large because there is at most a polynomial number of types. We have obtained the upper bound on the set $\tilde{E}(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ})$ but what we need is the upper bound on the size of the set $E(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ})$. Lemma 5, whose proof is omitted, helps us relate the size of $\tilde{E}(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ})$ and $E(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ})$. In fact, the lemma below says that the size of $E(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ})$ is roughly $1/K$ of $\mathbb{E}_C\left\{|\tilde{E}(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ})|\right\}$ with high probability.

Lemma 5. *It holds with probability at least $1-e^{-\rho_4\omega\sqrt{n}}$ for some $\rho_4>0$ that*

$$|E(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ}))|<\frac{\mathbb{E}_C\left\{|S(\mathbf{z},f_{XZ})\cap\tilde{E}(\bar{\mathbf{s}})|\right\}}{K}\times 2^{0.1\sqrt{n}\omega\delta}$$

when n is large enough.

Notice that (51) holds when (32) is true. Combining the fact that (51) holds with probability at least $1-e^{-e^{n^{1/2}\rho_2}}$ and the fact that $|E(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ})|$ is at most an exponential function of $\sqrt{n}$, the term $\mathbb{E}\left\{|S(\mathbf{z},f_{XZ})\cap\tilde{E}(\bar{\mathbf{s}})|\right\}$ is also upper bounded by (51) when n is sufficient large. Substituting the size of K, which is $2^{\sqrt{n}\omega(D(Q_1\|Q_0)-D(P_1\|P_0)+2\delta)}$, into Lemma 5, we obtain

$$\begin{aligned}&|E(\bar{\mathbf{s}})\cap S(\mathbf{z},P_{XZ}))|\\ &\leqslant 2^{\sqrt{n}\omega D(Q_1\|Q_0)-nI(P_{XZ})+0.5\omega\sqrt{n}\delta+O(\epsilon\omega\sqrt{n})+2n^{1/2}\epsilon_2} &(52)\end{aligned}$$

with probability at least $1-e^{-\rho_4\omega\sqrt{n}}$ when n is large enough, which completes the proof.

REFERENCES

[1] B. K. Dey, S. Jaggi, and M. Langberg, "Sufficiently myopic adversaries are blind," *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5718–5736, sep 2019.

[2] Y. Zhang, S. Vatedka, S. Jaggi, and A. D. Sarwate, "Quadratically constrained myopic adversarial channels," in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, jun 2018.

[3] A. D. Sarwate, "Coding against myopic adversaries," in *2010 IEEE Information Theory Workshop*, 2010, pp. 1–5.

[4] O. Kosut and J. Kliewer, "Authentication capacity of adversarial channels," in *2018 IEEE Information Theory Workshop (ITW)*. IEEE, nov 2018.

[5] A. Beemer, O. Kosut, J. Kliewer, E. Graves, and P. Yu, "Authentication against a myopic adversary," in *2019 IEEE Conference on Communications and Network Security (CNS)*. IEEE, jun 2019.

[6] L. Lai, H. E. Gamal, and H. V. Poor, "Authentication over noisy channels," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 906–916, feb 2009.

[7] W. Tu and L. Lai, "Keyless authentication and authenticated capacity," *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3696–3714, 2018.

[8] B. 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, September 2013.

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

[10] 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.

[11] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334–2354, 2016.

[12] 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.

[13] Q. E. Zhang, M. Bakshi, and S. Jaggi, "Covert communication over adversarially jammed channels," in *2018 IEEE Information Theory Workshop (ITW)*. IEEE, nov 2018.

[14] M. Tahmasbi and M. R. Bloch, "Covert secret key generation with an active warden," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1026–1039, Jun. 2020.

[15] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*, 3rd ed., ser. Springer Texts in Statistics. New York: Springer, 2005.

[16] B. K. Dey, S. Jaggi, and M. Langberg, "Sufficiently myopic adversaries are blind," in *2015 IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 1164–1168.

[17] I. Csiszar and P. Narayan, "The capacity of the arbitrarily varying channel revisited: positivity, constraints," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 181–193, 1988.