Towards a Characterization of the Covert Capacity of Bosonic Channels under Trace Distance

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

***Abstract*—We characterize upper and lower bounds for the covert capacity of lossy thermal-noise bosonic channels when measuring covertness using fidelity and trace distance. Although we fall short of characterizing the exact covert capacity, we also provide bounds on the number of secret-key bits required to achieve covertness. The bounds are established by combining recent quantum information theory results in separable Hilbert spaces, including position based coding (Oskouei et al., arXiv: 1804.08144 [1]), convex splitting (Khatri et al., arXiv: 1910.03883 [2]), and perturbation theory (Grace and Guha, arXiv: 2106.05533 [3]).**

I. Introduction

As quantum communications leap out of the lab [4], [5], there has been much interest in exploiting and demonstrating every advantage that quantum communications offer over classical ones. A prime example of such advantage is the ability to exploit the innate uncertainty of quantum measurements to achieve security levels beyond those that classical systems can offer. In particular, there has been tremendous progress in quantum key distribution showing the possibility of generating provably secret keys against quantum adversaries [6]–[8]. Most recently, the problem of ensuring low-probability of detection, also known as *covertness*, against quantum adversaries has been investigated [9].

Unlike reliable communications for which the number of reliable bits transmitted scales *linearly* with the number of channel uses, covert communications impose a sublinear scaling known as the *square-root law* [10]. This constraint is similar to what is known in steganography [11]. One can define an appropriate notion of covert capacity [12], [13], which is now known for a wide range of channels, including point-to-point discrete memoryless channels [12], [13], multiple access channels [14], some broadcast channels [15], interference channels [16], MIMO channels [17], [18], classical quantum channels [19], [20] and most recently bosonic channels [21], [22]. A subtle but crucial aspect in the definition of covert capacity is that it depends on the choice of covertness metric [23]. In particular, while using quantum relative entropy as a covertness metric offers convenient analytical tractability, detection performance is fundamentally related to the trace distance. In the classical case, the covert capacity under a trace distance metric, which then reduces to a total variation metric, is known [18], [23], [24]. The extension of these results to the quantum realm still poses challenges.

We make progress towards characterizing the covert capacity of lossy thermal-noise bosonic channels under a trace distance covertness metric by analyzing covertness using fidelity. In particular, we exploit the known relations between trace distance, fidelity, and quantum relative entropy of two density matrices ρ and σ [25]:

$$1-\sqrt{F(\rho,\sigma)} \leqslant \frac{1}{2}\left\|\rho-\sigma\right\|_1 \leqslant \sqrt{1-F(\rho,\sigma)} \leqslant \sqrt{\mathbb{D}(\rho\,\|\,\sigma)}\ . \tag{1}$$

These inequalities suggest fidelity is a natural intermediate between quantum relative entropy and trace distance. In addition, fidelity simultaneously offers analytical tractability for product states like the quantum relative entropy and retains the properties of a distance (through the purified distance) like the trace distance.

After a brief review of notation in Section II, we introduce our model and discuss our main results in Section III. We offer concluding remarks in Section IV and sketch proofs in Section V.

II. Notation

We use logarithms with base e. Let $\mathcal{D}(\mathcal{H})$ denote the set of density operators acting on a *separable* Hilbert space $\mathcal{H}$, let $\mathcal{D}_{\leqslant}(\mathcal{H})$ denote the set of subnormalized density operators with trace less than 1. The trace distance between two states ρ and σ is defined as $\frac{1}{2}\left\|\rho-\sigma\right\|_1$, where $\left\|\sigma\right\|_1 \triangleq \operatorname{tr}\left(\sqrt{\sigma^\dagger\sigma}\right)$. The fidelity for $\rho,\sigma \in \mathcal{D}(\mathcal{H})$ is defined as $\mathrm{F}(\rho,\sigma) \triangleq \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2$. The purified distance for $\rho,\sigma \in \mathcal{D}_{\leqslant}(\mathcal{H})$ is defined as $\mathrm{P}(\rho,\sigma) \triangleq \sqrt{1-\mathrm{F}(\rho\oplus[1-\operatorname{tr}(\rho)],\sigma\oplus[1-\operatorname{tr}(\sigma)]})$. The hypothesis testing relative entropy of $\rho,\sigma \in \mathcal{D}(\mathcal{H})$ is defined as [26]

$$\mathbb{D}_{\mathrm{H}}^{\epsilon}(\rho\,\|\,\sigma) \triangleq -\log \inf_{0\leqslant\Pi\leqslant I:\operatorname{tr}(\Pi\rho)\geqslant 1-\epsilon} \operatorname{tr}(\Pi\sigma)\,.$$

The hypothesis testing mutual information is defined as $\widetilde{\mathbb{I}}_{\mathrm{H}}^{\epsilon}(A;B)_\rho \triangleq \mathbb{D}_{\mathrm{H}}^{\epsilon}(\rho_{AB}\,\|\,\rho_A\otimes\rho_B)$. The max relative entropy of $\rho,\sigma \in \mathcal{D}_{\leqslant}(\mathcal{H})$ such that $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$ is defined as [27] $\mathbb{D}_{\max}(\rho\,\|\,\sigma) \triangleq \inf\left\{\lambda\in\mathbb{R} : \rho \leqslant e^{\lambda}\sigma\right\}$. The ϵ-smooth max relative entropy is defined as $\mathbb{D}_{\max}^{\epsilon}(\rho\,\|\,\sigma) \triangleq \inf_{\rho'\in\mathcal{B}^{\epsilon}(\rho)} \mathbb{D}_{\max}(\rho'\,\|\,\sigma)$, where

This work was suppoted by the National Science Foundation award 1910859.

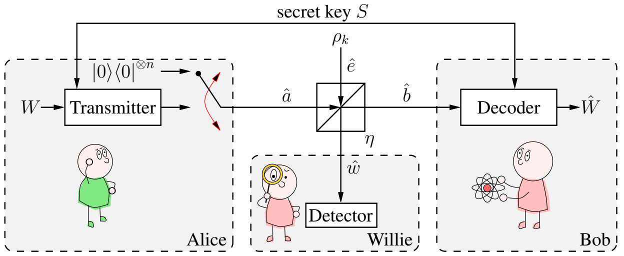

Fig. 1. Covert communication over a lossy thermal-noise bosonic channel.

$\mathcal{B}^\epsilon(\rho) \triangleq \{\sigma \in \mathcal{D}(\mathcal{H}) : P(\rho,\sigma) \leqslant \epsilon\}$. The ϵ-smooth max information is defined as

$$\tilde{\mathbb{I}}^\epsilon_{\max}(B;A)_\rho \triangleq \inf_{\rho'_{AB}:\rho'_{AB}\in\mathcal{B}^\epsilon(\rho_{AB})} \mathbb{D}_{\max}(\rho'_{AB} \,\|\, \rho_A \otimes \rho'_B)\,.$$

Moreover, for two integers $\lfloor a \rfloor$ and $\lceil b \rceil$ 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 $[\![a,b]\!] \triangleq \emptyset$. 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 $Q^{-1}(\cdot)$.

III. Covert Communication Model and Main Results

We consider the problem of covert communication over multiple uses of a single-mode lossy thermal-noise bosonic channel $\mathcal{L}^{(\eta,k)}_{A\to BW}$ [28], where η is the transmissivity and k is the mean photon number characterizing the thermal noise. A transmitter (Alice) attempts to reliably transmit a message to a receiver (Bob) while avoiding detection by an adversary (Willie).

As illustrated in Fig. 1, the channel is described by a beamsplitter with transmissivity η and an environment in a thermal bath ρ_k with mean photon number k, where $\rho_k = \frac{1}{\pi k}\int \exp\left(-\frac{|\alpha|^2}{k}\right) d^2\alpha \, |\alpha\rangle\langle\alpha| = \sum_{n=0}^\infty \frac{k^n}{(k+1)^{n+1}} |n\rangle\langle n|$, $\{|\alpha\rangle\}_{\alpha\in\mathbb{C}}$ is a set of coherent states and $\{|n\rangle\}_{n\in\mathbb{N}}$ is Fock basis. The relations between annihilation operators at the input and output of the channel are described by $\hat{b} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}$ and $\hat{w} = \sqrt{1-\eta}\hat{a} + \sqrt{\eta}\hat{e}$.

Alice aims to transmit a uniformly-distributed message $W \in [\![1,M]\!]$ with the aid of a uniformly-distributed secret key $S \in [\![1,K]\!]$ only known to Bob. Specifically, Alice's encoder is described by a set of encoding channels $\{\mathcal{E}^{(s,w)}_{SW\to A^n}\}_{s\in[\![1,K]\!],w\in[\![1,M]\!]}$ mapping a pair $(s,w) \in [\![1,K]\!] \times [\![1,M]\!]$ to an n-mode state $\rho_{A^n}(s,w)$. Upon observing $\rho_{B^n}(s,w) \triangleq \operatorname{tr}_W\left(\left(\mathcal{L}^{(\eta,k)}_{A\to BW}\right)^{\otimes n} \rho_{A^n}(s,w)\right)$ at the output of the channel, Bob uses his knowledge of the secret key s to apply Positive Operator-Valued Measures (POVMs) $\{\Pi^{(s,w)}_{B^n}\}_{w\in[\![1,M]\!]}$ and obtain an estimate $\widehat{W}$ of W. Reliability is measured by the maximal *average* probability of error: $P_e \triangleq \max_{s\in[\![1,K]\!]} \mathbb{P}\left(\widehat{W} \neq W | S = s\right)$, where $\mathbb{P}\left(\widehat{W} \neq w | W = w, S = s\right) = \operatorname{tr}\left(\left(I - \Pi^{(s,w)}_{B^n}\right)\rho_{B^n}\right)$. Alice may also choose to not transmit, in which case her input state is the n-mode vacuum state $|0\rangle\langle 0|^{\otimes n}$.

Unlike Bob, Willie's objective is merely to detect whether Alice is transmitting or not based on his observation ρ_{W^n} via a hypothesis test $\mathcal{T}_{W^n\to\{0,1\}}$ described by a POVM $\{T, I-T\}$. In particular, Willie expects $\rho^{\otimes n}_{0,W} \triangleq \rho^{\otimes n}_{\eta k}$ when there is no transmission between Alice and Bob (null hypothesis H_0), while he expects $\hat{\rho}_{W^n}$ when the transmission occurs (alternative hypothesis H_1), where $\hat{\rho}_{W^n}$ is the density operator induced by the codebook

$$\hat{\rho}_{W^n} \triangleq \frac{1}{MK}\sum_{s=1}^{K}\sum_{w=1}^{M} \operatorname{tr}_{B^n}\left(\left(\mathcal{L}^{(\eta,k)}_{A\to BW}\right)^{\otimes n} \rho_{A^n}(s,w)\right). \tag{2}$$

Note that Willie knows Alice's coding scheme but is unaware of the specific realizations of the message and the key. Any hypothesis test conducted by Willie on ρ_{W^n} satisfies $1 \geqslant \alpha + \beta \geqslant 1 - \frac{1}{2}\left\|\hat{\rho}_{W^n} - \rho^{\otimes n}_{0,W}\right\|_1$, where α and β are the probabilities of false alarm and missed detection, respectively; the lower bound is achieved by an optimal test [29, Chapter IV.2], [30, Lemma 9.1.1], and [31, Theorem 13.1.1]. Consequently, as already pointed out in [9], [21] (see also [18], [23], [24] for the classical case), the most natural covertness metric is the *trace distance* between $\rho^{\otimes n}_{0,W}$ and $\hat{\rho}_{W^n}$. Other metrics, such as the quantum relative entropy or the fidelity, are merely proxies to bound the adversary's detector performance [21], [22].

Definition III.1. *An (M,K,n,ϵ,δ) code with $\left(\{\mathcal{E}^{(s,w)}_{WK\to A^n}\}, \{\Pi^{(s,w)}_{B^n}\}\right)$ is both ϵ-reliable and δ-covert if $P_e \leqslant \epsilon$ and $\frac{1}{2}\left\|\hat{\rho}_{W^n} - \rho^{\otimes n}_{0,W}\right\|_1 \leqslant \delta$. The maximum number of messages that can be transmitted by an (M,K,n,ϵ,δ) code is denoted by $M^*(n,\epsilon,\delta)$ and the covert capacity is*

$$C_{\text{cov}} \triangleq \lim_{n\to\infty} \frac{\log M^*(n,\epsilon,\delta)}{\sqrt{n}}.$$

For a sequence of codes achieving the covert capacity, the associated secret key throughput is $\lim_{n\to\infty} \frac{\log K}{\sqrt{n}}$.

In the sequel, we further constrain the coding scheme encoders to use *n-mode coherent states*, which essentially precludes the use of entanglement across modes. We briefly discuss in Section IV how to approach a more general scheme with n-mode Gaussian states. Our main results are the following.

Proposition III.2. *The covert capacity of the lossy thermal-noise bosonic channel under a trace distance covertness constraint δ is lower bounded as*

$$C_{\text{cov}} \geqslant \frac{2\sqrt{\eta k(\eta k+1)}}{1-\eta}\eta\log\left(1 + \frac{1}{(1-\eta)k}\right)\delta.$$

The lower bound is achieved using a secret key throughput of $2\sqrt{\eta k(\eta k+1)}\left(\log\left(1+\frac{1}{\eta k}\right) - \frac{\eta}{1-\eta}\log\left(1+\frac{1}{(1-\eta)k}\right)\right)^+ \delta$, where $(x)^+ \triangleq \max(x,0)$.

As expected, no secret key is required when $\eta \geqslant \frac{1}{2}$, in which case Bob has a better channel than Willie.

Proposition III.3. *The covert capacity of the lossy thermal-noise bosonic channel under a trace distance covertness constraint δ when encoding is limited to n-mode coherent state encoding is upper bounded as*

$$C_{\text{cov}} \leqslant \frac{2\sqrt{\eta k(\eta k+1)}}{1-\eta}\eta\log\left(1+\frac{1}{(1-\eta)k}\right)Q^{-1}\left(\frac{1-\delta}{2}\right).$$

We now briefly discuss the meaning and consequences of the results, especially in relation with the existing characterization of the covert capacity under a quantum relative entropy metric [21], [22]. Specifically, [22, Theorem 1] shows that under a quantum relative entropy constraint δ, the covert capacity is

$$\frac{\sqrt{2\eta k(\eta k+1)}}{1-\eta}\eta\log\left(1+\frac{1}{(1-\eta)k}\right)\sqrt{\delta}. \tag{3}$$

The lower bound for C_{cov} in Proposition III.2 can be derived from [22, Theorem 1] using the quantum Pinsker inequality. However, our approach is to study covertness using a fidelity metric $F(\hat{\rho}_{W^n}, \rho_{0,W^n}^{\otimes n})$ and then exploit the inequalities in (1). The study of fidelity as an intermediate metric is what allows us to characterize the associated key throughput and requires a distinct analysis. Specifically, our result exploit position based coding [1], convex splitting [2], and perturbation theory results for fidelity [3] in separable Hilbert spaces. Proof sketches are given in Section V.

We next illustrate the bounds obtained for a lossy thermal-noise bosonic channel with mean photon number $k = 20$ and transmissivity $\eta = 0.7$. The covert capacity characterized in [22] under a quantum relative entropy constraint $\mathbb{D}(\hat{\rho}_{W^n} \,\|\, \rho_{0,W^n}^{\otimes n}) \leqslant \delta$ is denoted $f_D(\delta)$. Because of the quantum Pinsker's inequality, we have $C_{\text{cov}} \geqslant f_D(2\delta^2)$. The achievable covert throughput under a fidelity constraint $F(\hat{\rho}_{W^n}, \rho_{0,W^n}^{\otimes n}) \geqslant \delta$ characterized in our proof of Proposition III.2 is denoted $f_F(\delta)$. Because of the inequalities in (1), we have $C_{\text{cov}} \geqslant f_F(\delta^2)$ and it turns out that $f_F(\delta^2) = f_D(2\delta^2)$. Finally, the upper bound on the covert capacity under a trace distance constraint $\frac{1}{2}\left\|\hat{\rho}_{W^n} - \rho_{0,W^n}^{\otimes n}\right\|_1 \leqslant \delta$ derived in Proposition III.3 is denoted by $f_V(\delta)$. As expected and seen in Fig. 2, $f_V(\delta) \geqslant f_F(\delta^2) = f_D(2\delta^2)$. The gap between the bounds converges to $\sqrt{\frac{\pi}{2}} - 1$ as $\delta \to 0$, which one can check analytically and as illustrated in Fig. 3 through a plot of the relative increase of the upper bound with respect to (w.r.t.) the lower bound as a function of δ.

We also point out that the analysis of [22] does not subsume the steps involved in our analysis of fidelity for the proof of Proposition III.2 in Section V-A. In fact, because the relation between quantum relative entropy and fidelity in (1) is weaker than Pinsker's inequality [25], one can only ensure that $f_D(\delta^2) \leqslant f_F(\delta^2)$, which is illustrated in Fig. 2. One can check that the relative gap between $f_D(\delta^2)$ and $f_F(\delta^2)$ w.r.t. $f_D(\delta^2)$ converges to $\sqrt{\pi} - 1$ as $\delta \to 0$, which is illustrated in Fig. 4.

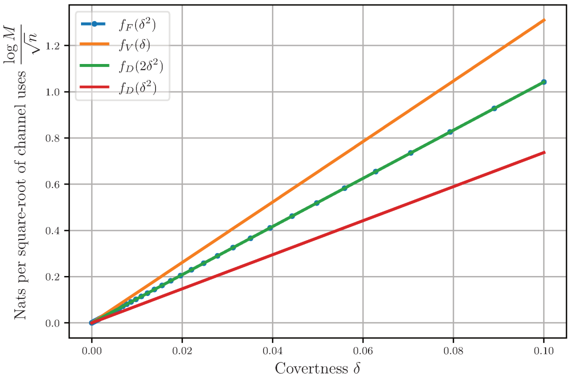

Fig. 2. Comparison of throughputs under different covertness metrics.

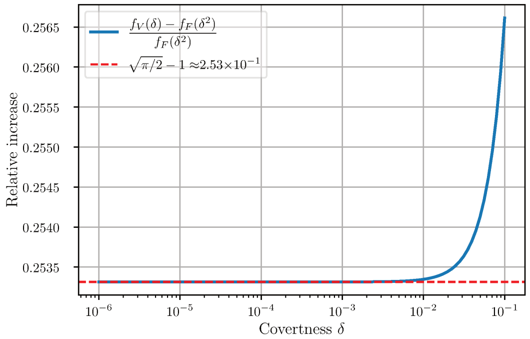

Fig. 3. Relative increase of the upper bound to the lower bound obtained by fidelity.

IV. Concluding Remarks

While our results fall short of establishing the covert capacity for a trace distance metric, we conjecture that our upper bound in Proposition III.3 is the correct characterization, which would give a 25% improvement. In fact, the main weakness of our achievability proof is to rely on fidelity merely because of its convenient factorization for product states. As conjectured in [22], we also expect that using a Gaussian ensemble of coherent states $\left\{\frac{1}{\pi s_n}\exp\left(-\frac{|\alpha|^2}{s_n}\right), |\alpha\rangle\langle\alpha|\right\}$ with

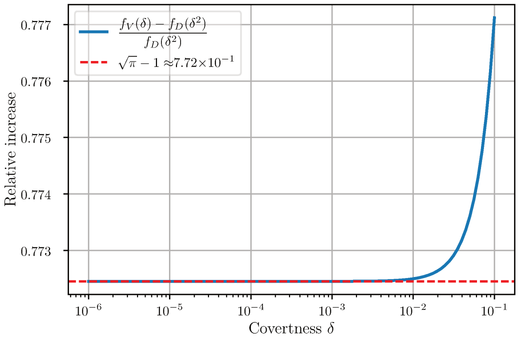

Fig. 4. Relative increase of the upper bound to the lower bound obtained by quantum relative entropy through fidelity.

$s_n = \frac{2\sqrt{\eta k(\eta k+1)}}{1-\eta} Q^{-1}\left(\frac{1-\delta}{2}\right)$ would achieve the upper bound in Proposition III.3. Indeed, the fact that thermal states are diagonal in Fock basis simplifies the direct analysis of trace distance $\frac{1}{2}\left\|\rho_{n,W}^{\otimes n} - \rho_{0,W}^{\otimes n}\right\|_1$ in Lemma IV.1.

Lemma IV.1. *The trace distance between two thermal states* $\rho_{\eta k}^{\otimes n}$ *and* $\rho_{\eta k+(1-\eta)s_n}^{\otimes n}$ *is*

$$\frac{1}{2}\left\|\rho_{\eta k}^{\otimes n} - \rho_{\eta k+(1-\eta)s_n}^{\otimes n}\right\|_1 \leqslant 1 - 2Q\left(\frac{\sqrt{n}(1-\eta)s_n}{2\sqrt{\eta k(1+\eta k)}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right). \tag{4}$$

Proof: The proof is similar to [23, Lemma 8] and [18, Lemma 9]. ■

The technical issues are: (i) the above Gaussian ensemble has an uncountable alphabet $\mathbb{C}$, which prevents us from using the known results of [1] and derandomizing the code; and, (ii) since the displaced thermal state cannot be diagonalized in Fock basis, a direct analysis on the trace distance between the Quadrature Phase-Shift Keying (QPSK) ensemble and a thermal state is challenging.

Similarly, while our converse is presently claimed only for encoding with product coherent states, we conjecture that the result holds for general n-mode Gaussian states or even more generally. In fact, for general n-mode Gaussian state, one can analyze a suboptimal test by first applying the following entanglement-breaking channel for each mode independently: $\rho_W \mapsto \sum_m \text{tr}\left(|m\rangle\langle m|\rho_W\right) \otimes |m\rangle\langle m|_M$. This operation reduces the observation into n-independent single-mode Gaussian states. Since a single-mode Gaussian state can be decomposed into a thermal state with a symplectic transform and displacement, the analysis of photon-number statistics would be more tractable. Our ongoing work is exploring ways to consolidate the above argument and address the aforementioned technical challenges.

V. Proof of Main results

A. Achievability

We first recall one-shot channel reliability and resolvability results via position-based coding and convex splitting lemma.

Lemma V.1 (one-shot channel reliability and resolvability adapted from [1], [2], [32], [33])**.** *Fix* $\epsilon \in (0,1)$, $\kappa \in (0, \frac{\delta}{2})$, $\gamma_1 \in (0, \frac{\epsilon^2}{10})$, $\gamma_2 \in (0, \frac{\kappa}{2})$, *and* $\gamma_3 \in (0, \frac{\kappa}{2} - \gamma_2)$. *Then for a bi-partite state* ρ_{XA} *and a channel* $\mathcal{G}: \rho_{XA} \mapsto \rho_{XBW}$, *there exists a coding scheme such that*

$$\log M \geqslant \mathbb{D}_{\text{H}}^{\epsilon^2/10-\gamma_1}(\rho_{XB} \,\|\, \rho_X \otimes \rho_B) - \log\left(\frac{4\epsilon^2}{10\gamma_1^2}\right), \tag{5}$$

$$\log MK \leqslant \mathbb{D}_{\text{max}}^{\kappa/2-\gamma_2-\gamma_3}(\rho_{XW} \,\|\, \rho_X \otimes \rho_W) + 2\log\left(\frac{1}{\gamma_2}\right) + \log\left(\frac{8}{\gamma_3^2}\right), \tag{6}$$

$\mathbb{E}_{\mathcal{C},S}\left\{\mathbb{P}\left(\widehat{W} \neq W|S\right)\right\} \leqslant \frac{\epsilon^2}{10}, \mathbb{E}_{\mathcal{C}}\left\{\frac{1}{2}\|\hat{\rho}_W - \rho_W\|_1\right\} \leqslant \kappa - \gamma_2$, *where* $\mathcal{C}$ *is the codebook.*

We now specialize the above results to n-channel uses for n sufficiently large and define the following state:

$$\rho_{n,XA} \triangleq \sum_{x\in[\![0,3]\!]} |x\rangle\langle x|_X \otimes \left|u_n e^{j\pi x/2}\right\rangle\left\langle u_n e^{j\pi x/2}\right|_A, \tag{7}$$

where $\left|u_n e^{j\pi x/2}\right\rangle$ is a coherent state and $u_n^2 = s_n \triangleq \frac{\sqrt{4\eta k(\eta k+1)(\delta-2\kappa)^2}}{(1-\eta)\sqrt{n}}$. Essentially, the above ensemble represents the QPSK symbols used in [22]. The received states at Bob's and Willie's systems therefore become $\rho_{n,XB} \triangleq \sum_{x\in[\![0,3]\!]} |x\rangle\langle x|_X \otimes \rho_{(1-\eta)k,B}(\sqrt{\eta}u_n e^{j\pi x/2})$ and $\rho_{n,XW} \triangleq \sum_{x\in[\![0,3]\!]} |x\rangle\langle x|_X \otimes \rho_{\eta k,W}(\sqrt{1-\eta}u_n e^{j\pi x/2})$, respectively, where $\rho_{\bar{N}}(\alpha)$ is a displaced thermal state with mean photon number $\bar{N}$ and displacement $|\alpha\rangle$.

It then follows from [22, Lemma 1 and 2] and [2, Corollary 11] that there exists a coding scheme with

$$\log M \geqslant n\mathbb{D}(\rho_{n,XB} \,\|\, \rho_X \otimes \rho_{n,B}) - \sqrt{nV(\rho_{n,XB}\|\rho_X \otimes \rho_{n,B})}Q^{-1}(\epsilon) + \mathcal{O}\left(\frac{R(\rho_{n,XB}\|\rho_X \otimes \rho_{n,B})^{3/4}}{V(\rho_{n,XB}\|\rho_X \otimes \rho_{n,B})}\right) \tag{8}$$

$$\log MK \leqslant n\mathbb{D}(\rho_{n,XW} \,\|\, \rho_X \otimes \rho_{n,W}) + \sqrt{nV(\rho_{n,XW}\|\rho_X \otimes \rho_{n,W})}Q^{-1}\left(\frac{\kappa^2}{4}\right) + \mathcal{O}\left(\frac{R(\rho_{n,XW}\|\rho_X \otimes \rho_{n,W})^{3/4}}{V(\rho_{n,XB}\|\rho_X \otimes \rho_{n,W})}\right), \tag{9}$$

such that $\mathbb{E}_{\mathcal{C},S}\left\{\mathbb{P}\left(\widehat{W} \neq W|S\right)\right\} \leqslant \frac{\epsilon^2}{10}$ and $\mathbb{E}_{\mathcal{C}}\left\{\frac{1}{2}\left\|\hat{\rho}_W^n - \rho_{n,W}^{\otimes n}\right\|_1\right\} \leqslant \kappa - \frac{1}{\sqrt{n}}$, where $R(\rho_{XB}\|\rho_X \otimes \rho_B)$ is the fourth central moment of quantum relative entropy, $\gamma_1 = \gamma_2 = \gamma_3 = \frac{1}{\sqrt{n}}$ and $\rho_{n,W}^{\otimes n} \triangleq \left(\text{tr}_X\left(\rho_{n,XW}\right)\right)^{\otimes n}$.

We next show that the specific choice of u_n leads to $\frac{1}{2}\left\|\rho_{n,W}^{\otimes n} - \rho_{0,W}^{\otimes n}\right\|_1 \leqslant \sqrt{1 - \text{F}\left(\rho_{n,W}^{\otimes n}, \rho_{0,W}^{\otimes n}\right)} \leqslant \delta - 2\kappa$ with Lemma V.2.

Lemma V.2. *The fidelity between QPSK ensemble* $\rho_{n,W}$ *and thermal state* $\rho_{\eta k+(1-\eta)s_n}$ *is*

$$\text{F}\left(\rho_{n,W}, \rho_{\eta k+(1-\eta)s_n}\right) = 1 - \frac{(1-\eta)^2 s_n^2}{4\eta k(\eta k+1)} + \mathcal{O}(s_n^3). \tag{10}$$

Proof: The proof follows from the techniques developed by [29, Chapter VI. eq(1.31)] and [21] with the perturbation theory applied to fidelity [3, Theorem 5]. ■

By the above lemma and tensor product property of fidelity, we immediately obtain $\text{F}\left(\rho_{n,W}^{\otimes n}, \rho_{0,W}^{\otimes n}\right) = \text{F}\left(\rho_{n,W}, \rho_{0,W}\right)^n \geqslant 1 - (\delta - 2\kappa)^2 + \frac{(\delta-2\kappa)^4}{2}$.

To identify a specific codebook, Markov's inequality shows that there exists a codebook $\mathcal{C}$ with n large enough and $\mathbb{E}_S\left(\mathbb{P}\left(\widehat{W} \neq W|S\right)\right) \leqslant \frac{\epsilon^2}{5}$ and $\frac{1}{2}\left\|\hat{\rho}_W^n - \rho_{n,W}^{\otimes n}\right\|_1 \leqslant 2\kappa - \frac{1}{\sqrt{n}}$, so that $\frac{1}{2}\left\|\hat{\rho}_W^n - \rho_{0,W}^{\otimes n}\right\|_1 \leqslant \delta$. To control the maximal average probability of error without affecting the covertness, we apply the rearranging argument in [24] and [18, Lemma 10] as follows.

Lemma V.3. *Suppose a code $\mathcal{C}$ contains K sub-codebooks of size M such that $\mathbb{E}_S\left(\mathbb{P}\left(\widehat{W} \neq W|S\right)\right) \leqslant \frac{\epsilon^2}{5}$ and $\frac{1}{2}\left\|\hat{\rho}_W^n - \rho_{0,W}^{\otimes n}\right\|_1 \leqslant \delta$. Then, there exists a code $\mathcal{C}'$ containing K' sub-codebooks of size M' such that $\max_{s\in[\![1,K]\!]} \mathbb{P}\left(\widehat{W} \neq W|S=s\right) \leqslant \epsilon$ and $\frac{1}{2}\left\|\hat{\rho}_W^n - \rho_{0,W}^{\otimes n}\right\|_1 \leqslant \delta$. In particular, $M'K' = MK$ and $\log M' = \log M + \mathcal{O}(\epsilon)$.*

Eventually, the information quantity in (8) and (9) can be obtained by [22, Lemma 2 and Appendix A] and

$$
\begin{aligned}
&\log M^*(n,\epsilon,\delta) \geqslant \\
&\quad \frac{2\sqrt{\eta k(\eta k+1)}}{1-\eta}\eta\log\left(1+\frac{1}{(1-\eta)k}\right)\sqrt{n}(\delta-2\kappa) \\
&\quad -\sqrt{\frac{2\sqrt{\eta k(\eta k+1)}\eta(1+2(1-\eta)k)(\delta-2\kappa)}{1-\eta}}n^{1/4} \\
&\qquad \times\log\left(1+\frac{1}{(1-\eta)k}\right)Q^{-1}(\frac{\epsilon^2}{10})+\mathcal{O}(n^{1/8}),
\end{aligned}
$$

$$
\begin{aligned}
&\log MK(n,\epsilon,\delta) \leqslant 2\sqrt{\eta k(\eta k+1)}\log\left(1+\frac{1}{\eta k}\right)\sqrt{n}(\delta-2\kappa) \\
&+\sqrt{2\sqrt{\eta k(\eta k+1)}(1+2\eta k)(\delta-2\kappa)} \\
&\times n^{1/4}\log\left(1+\frac{1}{\eta k}\right)Q^{-1}\left(\frac{\kappa^2}{4}\right)+\mathcal{O}(n^{1/8}).
\end{aligned}
$$

Proposition III.2 follows with $n \to \infty$ and κ arbitrarily small.

B. Converse proof for coherent-state codebook

We now provide a converse under a trace distance metric. We assume the codebook used by Alice is only comprised of coherent states, i.e., each codeword has the form: $\left|\psi^{(c)}\right\rangle^n = \left|\alpha_1^{(c)}\right\rangle \otimes \cdots \otimes \left|\alpha_n^{(c)}\right\rangle$, without entanglement between modes.

Proposition V.4. *Consider a sequence of covert communication schemes with coherent-state codebook and increasing blocklength $n \in \mathbb{N}^*$ characterized by $\epsilon_n \triangleq P_e^{max}$ and $\delta \geqslant \frac{1}{2}\left\|\hat{\rho}_W^n - \rho_{0,W}^{\otimes n}\right\|$. If $\lim\limits_{n\to\infty}\epsilon_n = 0$ and $\lim\limits_{n\to\infty} M = \infty$, then we have $\liminf\limits_{n\to\infty}\frac{\log M}{\sqrt{n}} \leqslant \frac{2\sqrt{\eta k(\eta k+1)}}{1-\eta}\eta\log\left(1+\frac{1}{(1-\eta)k}\right)Q^{-1}\left(\frac{1-\delta}{2}\right)$.*

The proof extends techniques in [9], [10], [18], [23], [24] by constructing a test for Willie that is sub-optimal yet powerful enough to obtain a tight upper bound. The main idea is to show that there cannot be too many codewords with high photon number, for otherwise the covertness would be violated. We start by establishing a lower bound relating the covertness metric to the minimum received photon number of codewords at Willie within a given code $\mathcal{M}$. Consider a hypothesis testing problem with two hypotheses H_0 and H_1 corresponding to two ensembles of states $\rho_{0,W}^{\otimes n} = \rho_{\eta k}^{\otimes n}$ and $\hat{\rho}_{W^n}$. We define a sub-optimal photon number detector $\{T, I-T\}$, where $T \triangleq \sum_{m_1+m_2+\cdots+m_n\geqslant\tau} |m_1\rangle\langle m_2|\otimes|m_2\rangle\langle m_2|\otimes\cdots\otimes|m_n\rangle\langle m_n|$, and the threshold τ will be specified later. As the codewords are n-mode coherent states with no entanglement, the received observations are n-independent displaced thermal state $\rho_{\eta k}(\alpha_i^{(m)})$. Equivalently, this test counts the photon numbers *independently* for each mode, and reduces to a classical hypothesis testing problem $T(m^n) = \mathbb{1}\{\sum_{i=1}^n m_i \geqslant \tau\}$.

With the photon number statistics, we then choose $\tau = \frac{(1-\eta)N_*}{2} + n\eta k$, where $N_* \triangleq \min_{m\in\mathcal{M}}\sum_{i=1}^n\left|\alpha_i^{(m)}\right|^2$, and use Berry-Esseen Theorem to obtain

$$
\begin{aligned}
&\frac{1}{2}\left\|\hat{\rho}_{W^n} - \rho_{0,W}^{\otimes n}\right\|_1 \geqslant 1-\alpha-\beta \geqslant 1-\frac{B_0+B_1}{\sqrt{n}} \\
&-\frac{(1-\eta)^2(1+2\eta k)N_*^2}{4\sqrt{2\pi}n^{3/2}[\eta k(\eta k+1)]^{3/2}} - 2Q\left(\frac{(1-\eta)N_*}{2\sqrt{n\eta k(\eta k+1)}}\right).
\end{aligned} \tag{11}
$$

We next show that for a covert codebook $\mathcal{C}$, the maximum photon number of a non-empty low-photon-number sub-codebook is bounded. The key idea is to use (11) to analyze the covertness [23, Lemma 12] and [18, Lemma 6].

Lemma V.5. *For any covert codebook $\mathcal{C}$, given a decreasing sequence $\{\gamma_n\}$ with $\gamma_n \in (0,1)$ and $\lim\limits_{n\to\infty}\gamma_n = 0$, there exists a subset of codewords $\mathcal{C}^{(\ell)}$ such that $\left|\mathcal{C}^{(\ell)}\right| \geqslant \gamma_n|\mathcal{C}|$ and $N^{(c)} \leqslant A\sqrt{n}$, where $A \triangleq \frac{2\sqrt{\eta k(\eta k+1)}}{1-\eta}Q^{-1}\left(\frac{1-\delta}{2} - \frac{\nu^2(1-\eta)^2(1+2\eta k)}{4\sqrt{2\pi n}[\eta k(\eta k+1)]^{3/2}} - \gamma_n\right)$, and ν depends on the channel.*

The codebook $\mathcal{C}$ can be partitioned into K_n sub-codebooks $\mathcal{C}_s$ indexed by the key value s for all $s \in [\![1, K_n]\!]$ such that $\mathcal{C} = \cup_{s\in[\![1,K_n]\!]}\mathcal{C}_s$, and the size of each sub-codebook is M_n. Let $\mathcal{C}_s^{(\ell)} \triangleq \mathcal{C}_s \cap \mathcal{C}^{(\ell)}$. By the pigeonhole principle, there exists a sub-codebook $\mathcal{C}_s$ satisfying $\left|\mathcal{C}_s^{(\ell)}\right| \geqslant \gamma_n M_n$. Furthermore, since the average probability of error of $\mathcal{C}_s$ is at most ϵ_n, we have $\mathbb{P}\left(W \neq \widehat{W}|S=s\right) \leqslant \frac{\epsilon_n}{\gamma_n}$ for the codebook $\mathcal{C}_s^{(\ell)}$, which vanishes in the limit of large n upon choosing $\{\gamma_n\}_{n=1}^\infty$ such that $\lim\limits_{n\to\infty}\frac{\epsilon_n}{\gamma_n} = 0$.

By Fano's inequality, Holevo bound, and the fact that the capacity of bosonic channel is additive [34],

$$
\log\left|\mathcal{C}_s^{(\ell)}\right|(1-\frac{\epsilon_n}{\gamma_n}) - 1 \overset{(a)}{\leqslant} nC(\frac{A}{\sqrt{n}},\eta,k) \tag{12}
$$

$$
\leqslant \eta A\sqrt{n}\log\left(1+\frac{1}{(1-\eta)k}\right), \tag{13}
$$

where (a) follows from the fact that capacity of bosonic channel is additive [34] and $C(N,\overline{N},\eta) \triangleq g(\eta N+(1-\eta)\overline{N}) - g((1-\eta)\overline{N})$, where $g(x) = (x+1)\log(x+1) - x\log x$. Combining Lemma V.5 and (13), and the fact that $\lim\limits_{n\to\infty}\gamma_n = 0$, we then have

$$
\log\left|\mathcal{C}_s^{(\ell)}\right| \leqslant \frac{\frac{2\eta\sqrt{n\eta k(\eta k+1)}\log\left(1+\frac{1}{(1-\eta)k}\right)}{1-\eta}Q^{-1}\left(\frac{1-\delta}{2}\right)+\mathcal{O}(1)}{1-\frac{\epsilon_n}{\gamma_n}}.
$$

We further choose the sequence $\{\gamma_n\}$ such that $\lim\limits_{n\to\infty} -\frac{\log\gamma_n}{\sqrt{n}} = 0$, then the result follows.

REFERENCES

[1] S. K. Oskouei, S. Mancini, and M. M. Wilde, "Union bound for quantum information processing," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 475, no. 2221, p. 20180612, Apr. 2018.

[2] S. Khatri, E. Kaur, S. Guha, and M. M. Wilde, "Second-order coding rates for key distillation in quantum key distribution," *arXiv e-prints*, vol. 1910.03883, Oct. 2019.

[3] M. R. Grace and S. Guha, "Perturbation Theory for Quantum Information," *arXiv preprint*, vol. 2106.05533, Jun. 2021.

[4] J. Qiu, "Quantum communications leap out of the lab," *Nature*, vol. 508, no. 7497, pp. 441–442, Apr. 2014.

[5] Y. A. Chen, Q. Zhang, T. Y. Chen, W. Q. Cai, S. K. Liao, J. Zhang, K. Chen, J. Yin, J. G. Ren, Z. Chen, S. L. Han, Q. Yu, K. Liang, F. Zhou, X. Yuan, M. S. Zhao, T. Y. Wang, X. Jiang, L. Zhang, W. Y. Liu, Y. Li, Q. Shen, Y. Cao, C. Y. Lu, R. Shu, J. Y. Wang, L. Li, N. L. Liu, F. Xu, X. B. Wang, C. Z. Peng, and J. W. Pan, "An integrated space-to-ground quantum communication network over 4,600 kilometres," *Nature*, vol. 589, no. 7841, pp. 214–219, Jan. 2021.

[6] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 145–195, Mar. 2002.

[7] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Reviews of Modern Physics*, vol. 81, no. 3, pp. 1301–1350, Sep. 2009.

[8] F. Xu, X. Ma, Q. Zhang, H. K. Lo, and J. W. Pan, "Secure quantum key distribution with realistic devices," *Reviews of Modern Physics*, vol. 92, no. 2, p. 025002, Jun. 2020.

[9] 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.

[10] 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.

[11] A. D. Ker, "A Capacity Result for Batch Steganography," *IEEE Signal Processing Letters*, vol. 14, no. 8, pp. 525–528, Aug. 2007.

[12] 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.

[13] 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.

[14] K. S. K. Arumugam and M. R. Bloch, "Covert Communication Over a K -User Multiple-Access Channel," *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7020–7044, Nov. 2019.

[15] V. Y. F. Tan and S.-H. Lee, "Time-Division is Optimal for Covert Communication Over Some Broadcast Channels," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1377–1389, May 2019.

[16] K.-H. Cho and S.-H. Lee, "Treating Interference as Noise Is Optimal for Covert Communication Over Interference Channels," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 322–332, 2021.

[17] 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.

[18] 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, 2021.

[19] A. Sheikholeslami, B. A. Bash, D. Towsley, D. Goeckel, and S. Guha, "Covert communication over classical-quantum channels," in *Proc. of 2016 IEEE International Symposium on Information Theory*, Jul. 2016, pp. 2064–2068.

[20] L. Wang, "Optimal throughput for covert communication over a classical-quantum channel," in *Proc. of IEEE Information Theory Workshop*, Sep. 2016, pp. 364–368.

[21] M. S. Bullock, C. N. Gagatsos, S. Guha, and B. A. Bash, "Fundamental Limits of Quantum-Secure Covert Communication Over Bosonic Channels," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 3, pp. 471–482, Mar. 2020.

[22] C. N. Gagatsos, M. S. Bullock, and B. A. Bash, "Covert Capacity of Bosonic Channels," *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 2, pp. 555–567, Feb. 2020.

[23] 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.

[24] Q. E. Zhang, M. R. Bloch, M. Bakshi, and S. Jaggi, "Undetectable Radios: Covert Communication under Spectral Mask Constraints," in *Proc. of IEEE International Symposium on Information Theory*, Paris, France, Jul. 2019, pp. 992–996.

[25] M. Hayashi, *Quantum Information Theory*, 2nd ed., ser. Graduate Texts in Physics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017.

[26] L. Wang and R. Renner, "One-Shot Classical-Quantum Capacity and Hypothesis Testing," *Physical Review Letters*, vol. 108, no. 20, p. 200501, May 2012.

[27] N. Datta, M. Mosonyi, M.-H. Hsieh, and F. G. S. L. Brandao, "A Smooth Entropy Approach to Quantum Hypothesis Testing and the Classical Capacity of Quantum Channels," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8014–8026, Dec. 2013.

[28] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Reviews of Modern Physics*, vol. 84, no. 2, pp. 621–669, May 2012.

[29] C. W. Helstrom, *Quantum detection and estimation theory*, 1st ed. New York, NY, USA: Academic Press, 1976.

[30] M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge: Cambridge University Press, 2017.

[31] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, ser. Springer Texts in Statistics. Springer New York, 2006.

[32] A. Anshu, R. Jain, and N. A. Warsi, "Building Blocks for Communication Over Noisy Quantum Networks," *IEEE Transactions on Information Theory*, vol. 65, no. 2, pp. 1287–1306, Feb. 2019.

[33] M. M. Wilde, "Position-based coding and convex splitting for private communication over quantum channels," *Quantum Information Processing*, vol. 16, no. 10, p. 264, Oct. 2017.

[34] V. Giovannetti, A. S. Holevo, and R. García-Patrón, "A Solution of Gaussian Optimizer Conjecture for Quantum Channels," *Communications in Mathematical Physics*, vol. 334, no. 3, pp. 1553–1571, Dec. 2015.