Resource-Efficient Entanglement-Assisted Covert Communications over Bosonic Channels

Shi-Yuan Wang, Shang-Jen Su, Matthieu R. Bloch
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA30332
Email: {shi-yuan.wang, ssu49}@gatech.edu, matthieu.bloch@ece.gatech.edu

***Abstract*—We revisit the problem of entanglement-assisted covert communication over bosonic channels and show that the benefits of entanglement can be achieved with fewer entanglement resources than previously identified. Specifically, we show that $\mathcal{O}(\sqrt{n}\log n)$ covert and reliable bits can be exchanged using $\omega(\sqrt{n})\cap o(n)$ Two-Mode Squeezed-Vacuum (TMSV) pairs and $\omega(1)\cap o(\sqrt{n})$ secret-key bits. The conceptual approach behind the result is to combine 1) soft-covering and secret-key resources as a coordination mechanism and 2) superposition coding in the form of a two-layer On-Off Keying (OOK) and Phase Shift Keying (PSK) to index channel uses in which TMSV pairs are encoded. This approach is related to the idea of quantum trade-off coding, specialized and extended to the covert communication setting. Our technical contribution is to develop one-shot bounds then specialized to bosonic channels.**

I. Introduction

Covert communications refer to situations in which the objective is to achieve reliability while simultaneously remaining undetectable by an adversary [1]. This stringent requirement often leads to a *square-root law*, by which the number of covert and reliable bits transmitted over n uses of a channel cannot scale greater than $O(\sqrt{n})$ [2]; a notion of covert capacity capturing the constant in front of the $\sqrt{n}$ can then be appropriately defined and characterized [3], [4]. Subsequent studies have extended these characterizations to multiuser scenarios [5]–[10], Gaussian and continuous time channels [4], [11]–[13], as well as investigated situations in which the square-root law can be circumvented [14]–[19]. Specifically relevant to the present work, the problem of covert communication over quantum channels has attracted attention for its potential to offer unique quantum-secure capabilities. Following the experimental demonstration of quantum-secure communications over optical channels [20], the covert capacity of classical-quantum channels [21]–[23] and lossy bosonic channels [24]–[27] has been studied. In particular, the use of entanglement resources results in a scaling of $O(\sqrt{n}\log n)$ for the number of covert and reliable bits, strictly improving on the covert capacity without entanglement [23], [25].

One of the central questions in covert communications has been to identify the resources required to achieve covertness, i.e., in the form of secret keys or entanglement pairs. In particular, for classical channels, both the number of covert and reliable bits and the number of secret key bits required to support the communication scale as the square root of the blocklength [3], sometimes not requiring any secret key at all [3], [28]. In the case of quantum covert communications, the recent results of [29], [30] show that a $O(\sqrt{n}\log n)$ scaling can be achieved with no more than n TMSV pairs.

The main contribution of the present work is to extend the results of [25], [30] by showing that the quantum advantage of entanglement-assisted covert communications requires much fewer resources than previously identified: we show that a covert throughput $O(\sqrt{n}\log n)$ can be obtained with $\omega(\sqrt{n})\cap o(n)$ TMSV pairs and $\omega(1)\cap o(\sqrt{n})$ secret bits. Specifically, we achieve this result by combining two layers of coded modulation, including OOK and PSK, in a manner reminiscent of quantum trade-off coding [31]. Our approach leverages one-shot results subsequently specialized to the specific bosonic channel at hand.

The rest of the paper is organized as follows. We introduce necessary notation used throughout the paper in Section II and the formal system model in Section III. We present our main results in Section IV, together with a high-level proof. Finally, we outline more detailed aspects of the proof in Section V.

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 set $\mathcal{X}$ and $n\in\mathbb{N}_*$, a sequence of n elements is denoted by $x^n\triangleq(x_1,\cdots,x_n)\in\mathcal{X}^n$, and we sometimes use $\mathbf{x}\triangleq(x_1,\cdots,x_n)\in\mathcal{X}^n$ when it is clear from the context. For $\mathbf{x}\in\mathcal{X}^n$, $\hat{p}_{\mathbf{x}}$ denotes the type of $\mathbf{x}$, i.e., $\hat{p}_{\mathbf{x}}(x)=\frac{1}{n}\sum_{i=1}^n\mathbf{1}\{x_i=x\}$. 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 $[a;b]\triangleq\emptyset$. 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 $Q^{-1}(\cdot)$. Throughout the paper, $\log$ is with respect to (w.r.t.) base e, and therefore all the information quantities should be understood in *nats*.

Let $\mathcal{D}(\mathcal{H})$ denote the set of density operators acting on a *separable* Hilbert space $\mathcal{H}$, and let $\mathcal{D}_{\leqslant}(\mathcal{H})$ denote the set of subnormalized density operators with trace less than 1. Let id be the identity operator acting on $\mathcal{D}(\mathcal{H})$. For $\rho\in\mathcal{D}(\mathcal{H})$, the von Neumann entropy is $\mathbb{H}(\rho)\triangleq-\text{tr}\,(\rho\log\rho)$. The trace distance between two states ρ and

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

σ is defined as $\frac{1}{2}\|\rho-\sigma\|_1$, where $\|\sigma\|_1 \triangleq \operatorname{tr}(\sqrt{\sigma^\dagger\sigma})$. The fidelity for $\rho,\sigma \in \mathcal{D}(\mathcal{H})$ is defined as $\mathrm{F}(\rho,\sigma) \triangleq \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$. The purified distance for $\rho,\sigma \in \mathcal{D}_{\leqslant}(\mathcal{H})$ is defined as $\mathrm{P}(\rho,\sigma) \triangleq \sqrt{1-\mathrm{F}(\rho\oplus[1-\operatorname{tr}(\rho)],\sigma\oplus[1-\operatorname{tr}(\sigma)])}$. For $\sigma,\rho \in \mathcal{D}(\mathcal{H})$, the quantum relative entropy is $\mathbb{D}(\rho\,\|\,\sigma) \triangleq \operatorname{tr}(\rho(\log\rho-\log\sigma))$, the quantum relative entropy variance is $V(\rho\|\sigma) \triangleq \operatorname{tr}\left(\rho\left(\log\rho-\log\sigma-\mathbb{D}(\rho\,\|\,\sigma)\right)^2\right)$, and the fourth central moment of quantum relative entropy is $R(\rho\|\sigma) \triangleq \operatorname{tr}\left(\rho\left(\log\rho-\log\sigma-\mathbb{D}(\rho\,\|\,\sigma)\right)^4\right)$. For a Classical-Quantum (cq) state ρ_{XA}, $\mathbb{D}(\rho_{XA}\,\|\,\rho_X\otimes\rho_A)$ also represents the Holevo information. The hypothesis testing relative entropy of $\rho,\sigma \in \mathcal{D}(\mathcal{H})$ is defined as $\mathbb{D}_{\mathrm{H}}^{\epsilon}(\rho\,\|\,\sigma) \triangleq -\log\inf_{0\leqslant\Pi\leqslant I:\operatorname{tr}(\Pi\rho)\geqslant 1-\epsilon}\operatorname{tr}(\Pi\sigma)$ [32]. 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 $\mathbb{D}_{\max}(\rho\,\|\,\sigma) \triangleq \inf\{\lambda\in\mathbb{R}:\rho\leqslant e^{\lambda}\sigma\}$ [33]. 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 $\mathcal{B}^{\epsilon}(\rho) \triangleq \{\sigma\in\mathcal{D}(\mathcal{H}) : P(\rho,\sigma)\leqslant\epsilon\}$.

III. Covert Communication Model

We consider the problem of covert communication over multiple uses of a single-mode lossy thermal-noise bosonic channel $\mathcal{L}_{A\to BW}^{(\kappa,N_B)}$ [34] illustrated in Fig. 1, where κ is the transmissivity and N_B is the mean photon number characterizing the background thermal noise. A transmitter (Alice) attempts to reliably transmit a message to a receiver (Bob) while avoiding detection by an adversary (Willie). Specifically, the thermal state ρ_{N_B} with mean photon number N_B is $\rho_{N_B} \triangleq \frac{1}{\pi N_B}\int\exp\left(-\frac{|\alpha|^2}{N_B}\right)d^2\alpha|\alpha\rangle\langle\alpha| = \sum_{n=0}^{\infty}\frac{N_B^n}{(N_B+1)^{n+1}}|n\rangle\langle n|$, where $\{|\alpha\rangle\}_{\alpha\in\mathbb{C}}$ is the over-complete set of coherent states and $\{|n\rangle\}_{n\in\mathbb{N}}$ is the Fock basis. The relations between annihilation operators at the input and output of the channel are described by $\hat{b}=\sqrt{\kappa}\hat{a}+\sqrt{1-\kappa}\hat{e}$ and $\hat{w}=-\sqrt{1-\kappa}\hat{a}+\sqrt{\kappa}\hat{e}$.

More formally, Alice transmits her uniform message $W\in[1;M]$ with the aid of a uniform secret key $S\in[1;K]$ and m pairs of entangled states $|\psi\rangle_{RI}^{\otimes m}$ preshared with Bob. In the present work, the entangled pair consists of the reference (R) and idler (I) of a TMSV described in the Fock basis by $|\psi\rangle_{RI} = \sum_{n=0}^{\infty}\sqrt{\frac{N_S^n}{(N_S+1)^{n+1}}}|n\rangle_R|n\rangle_I$, where N_S is the effective mean photon number, or the signal power, on each sub-system. Alice receives R^m while Bob noiselessly obtains I^m. Inspired by the trade-off coding framework [31], [35], we investigate the achievable throughput of classical bits with limited access to the entangled resource, and therefore assume the number of entangled pairs m may be different from the actual number of channel uses n. Specifically, Alice transmits her message W with a set of encoding channels $\{\mathcal{E}_{WSR^m\to A^n}^{(w,s)}\}_{(w,s)}$ mapping systems (W,S,R^m) to an n-mode state $\sigma_{A^n}(w,s) = \operatorname{tr}_{I^m}(\sigma_{A^nI^m}(w,s))$, where $\sigma_{A^nI^m}(w,s) \triangleq (\mathcal{E}_{WSR^m}^{(w,s)}\otimes\mathrm{id}_I^{\otimes m})\psi_{RI}^{\otimes m}$, and $\psi_{RI}\triangleq|\psi\rangle\langle\psi|_{RI}$. After n uses of the channel $\mathcal{L}_{A\to BW}^{(\kappa,N_B)}$, Bob observes the state $\sigma_{B^nI^m}(w,s) \triangleq \operatorname{tr}_{W^n}\left(((\mathcal{L}_{A\to BW}^{(\kappa,N_B)})^{\otimes n}\otimes\mathrm{id}_{I^m})\sigma_{A^nI^m}(w,s)\right)$ and Willie observes $\sigma_{W^n}(w,s) \triangleq \operatorname{tr}_{B^n}\left((\mathcal{L}_{A\to BW}^{(\kappa,N_B)})^{\otimes n}\sigma_{A^n}(w,s)\right)$. Bob's objective is to reliably decode the message W with his knowledge of secret key S and a collection of decoding Positive Operator-Valued Measures (POVMs) $\{\{\Pi_{B^nI^m}^{(w,s)}\}_w\}_s$. The reliability metric is the *average* probability of error of his estimate $\widehat{W}$ of W, $P_e \triangleq \mathbb{E}_S\left[\mathbb{P}(\widehat{W}\neq W|S)\right]$, where $\mathbb{P}(\widehat{W}\neq w|W=w,S=s) = \operatorname{tr}\left((\mathrm{id}_{B^nI^m}-\Pi_{B^nI^m}^{(w,s)})\sigma_{B^nI^m}(w,s)\right)$. On the other hand, Willie's objective is to detect whether Alice is transmitting or not based on his observations σ_{W^n} via a hypothesis test $\mathcal{T}_{W^n\to\{0,1\}}$ described by a POVM $\{T,\mathrm{id}-T\}$. Note that Willie has knowledge of the exact channel state and Alice's coding scheme but has neither access to the realization of secret key s nor to the entangled pairs $|\psi\rangle_{RI}^{\otimes m}$. In particular, when Alice chooses not to transmit anything, her input state to the channel is simply $|0\rangle\langle 0|^{\otimes n}$, and therefore Willie expects $\sigma_{0,W}^{\otimes n}\triangleq\rho_{\kappa N_B}^{\otimes n}$ for the null hypothesis H_0. When a transmission occurs, Willie expects $\hat{\sigma}_{W^n}$ for the alternative hypothesis H_1, where $\hat{\sigma}_{W^n}\triangleq\frac{1}{MK}\sum_{w=1}^{M}\sum_{s=1}^{K}\sigma_{W^n}(w,s)$ is the mixed state induced by the codebook. The covertness metric is then captured by the trace distance between $\sigma_{0,W}^{\otimes n}$ and $\hat{\sigma}_{W^n}$. As already pointed out in [1], [24], [26], the trace distance metric is the most operationally relevant choice since any test performed by Willie on σ_{W^n} satisfies $1\geqslant\alpha+\beta\geqslant 1-\frac{1}{2}\|\hat{\sigma}_{W^n}-\sigma_{0,W}^{\otimes n}\|_1$, where α and β are probabilities of false alarm and missed-detection, and the lower bound can be achieved by the Holevo-Helstrom test [36, Chapter IV.2], [37], [38, Lemma 9.1.1].

A code achieving reliability and covertness for the channel model of Fig. 1 is formally defined as follows.

Definition 1. *An $(M,K,m,n,\epsilon,\delta)$ code defined by $\left(\{\mathcal{E}_{WSR^m\to A^n}^{(w,s)}\}_{(w,s)},\{\Pi_{B^nI^m}^{(w,s)}\}_{(w,s)}\right)$ is both ϵ-reliable and δ-covert if $P_e\leqslant\epsilon$ and $\frac{1}{2}\|\hat{\sigma}_{W^n}-\sigma_{0,W}^{\otimes n}\|_1\leqslant\delta$.*

IV. Main Result

We propose a protocol achieving a reliable and covert communication with a two-layer encoding, inspired by the classical-entanglement trade-off coding framework [31], [35]. The intuition is that Alice and Bob may reduce the number of TMSV pairs by *indexing* a subset of positions in which the pairs are phase-coded. Although this indexing might partially rely on secret-key bits, it also carries information bits as the indexing may be viewed as OOK modulation. Specifically, Alice splits her message W into two message layers W_1 and W_2; the first message layer W_1 is coded for reliability and

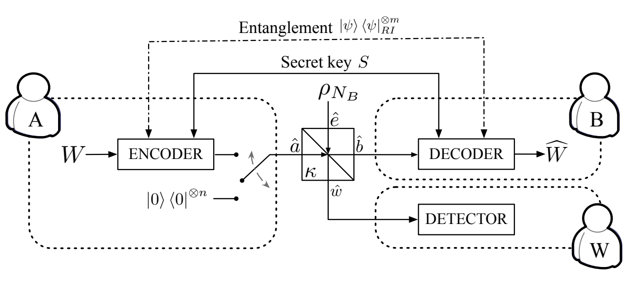

Fig. 1. Covert communication model with secret keys and entangled resources over a lossy thermal-noise bosonic channel.

covertness using OOK, while the second message layer W_2 is phase-coded for reliability onto the references of TMSV pairs, which are transmitted in the ON timeslots of the OOK modulation. Bob subsequently attempts to decode the first layer W_1 using the preshared secret key S but without any entanglement resources, and then decodes the second layer W_2 based on the estimate $\widehat{W}_1$ and the entanglement resources I^m.

Our protocol achieves covertness by combining a *sparse* OOK encoding together with *diffuse* power when phase modulating. This simultaneous use of sparse and a diffuse power is central to efficiently use resources, which is unlike existing coding schemes [25], [30] that solely rely on diffuse power to achieve covertness. We shall see that there exist trade-offs between the sparsity and the power levels required for covertness.

Formally, we shall use the following two-layer codes.

Definition 2. *Consider a sub-family of $(M, K, m, n, \epsilon, \delta)$ codes where $M = M_1 M_2$. An $(M_1, M_2, K, m, n, \epsilon, \delta)$ 2-layer code with encoding channels $\{\mathcal{E}^{(w_1,w_2,s)}_{W_1 W_2 S R^m \to A^n}\}_{(w_1,w_2,s)}$ and collections of 2-stage decoding POVMs $\left(\{\Pi^{(w_1,s)}_{B^n}\}_{(w_1,s)}, \{\Gamma^{(w_1,w_2)}_{B^n I^m}\}_{(w_1,w_2)}\right)$ is ϵ-reliable and δ-covert if*

$$P_e = \mathbb{E}_S \left[\mathbb{P}\left(\widehat{W}_1 \neq W_1 \text{ or } \widehat{W}_2 \neq W_2 | S\right)\right] \leqslant \epsilon,$$
$$\frac{1}{2}\|\hat{\sigma}_{W^n} - \sigma_{0,W}^{\otimes n}\|_1 \leqslant \delta.$$

Our main result is the characterization of $(M_1, M_2, K, m, n, \epsilon, \delta)$ code as follows.

Theorem 3. *Let $\epsilon, \delta > 0$. Let $\{\alpha_n\}_{n\in\mathbb{N}_*}$ and $\{s_n\}_{n\in\mathbb{N}_*}$ be sequences of positive real numbers satisfying $\alpha_n \in o(1) \cap \omega(n^{-\frac{1}{2}})$, $s_n \in o(1) \cap \omega(n^{-\frac{1}{2}})$, and such that $\alpha_n s_n \leqslant \frac{2\sqrt{\kappa N_B(1+\kappa N_B)}}{(1-\kappa)} Q^{-1}\left(\frac{1-\delta}{2}\right) n^{-\frac{1}{2}} - o(n^{-\frac{1}{2}})$. There exist $\mu \in (0,1)$ and a sequence of $(M_1, M_2, K, m, n, \epsilon, \delta)$ codes such that for n large enough,*

$$\log M_1 = \frac{n\kappa^2 \alpha_n s_n^2}{2(1-\kappa)N_B(1+(1-\kappa)N_B)} + o(\sqrt{n}),$$
$$\log M_1 K = \frac{n(1-\kappa)^2 \alpha_n s_n^2}{2\kappa N_B(1+\kappa N_B)} + o(\sqrt{n}),$$
$$\log M_2 = -(1-\mu)\frac{\kappa n \alpha_n s_n \log s_n}{1+(1-\kappa)N_B} + o(\sqrt{n}\log n),$$

and $m = n\alpha_n(1-\mu)$.

A couple of comments are in order. The parameters α_n and s_n capture the sparsity and the diffuse signal power level of our coding scheme, respectively. There is a joint constraint that restricts the scaling of the product $\alpha_n s_n$ with the blocklength n but different choices of the sparsity and the diffuse signal power level result in different scalings for the number of bits $\log M_1$, $\log M_2$, the number of secret-key bits $\log K$, and the number of TMSV pairs m. As expected, the quantum entanglement advantage presents itself in the second layer of coding $\log M_2$ and not in the first layer $\log M_1$, but the benefit of the first layer presents itself in the reduced number of TMSV pairs scaling as $O(n\alpha_n)$ instead of n.

Theorem 3 can be further specialized to reveal constants associated to the scaling.

Corollary 4. *For any $\tau \in (0, \frac{1}{2})$, fix some $\alpha, s > 0$ such that $\alpha s = \frac{2\sqrt{\kappa N_B(1+\kappa N_B)}}{1-\kappa} Q^{-1}\left(\frac{1-\delta}{2}\right)$, and let $\lim_{n\to\infty} \frac{\alpha_n}{n^{-\frac{1}{2}+\tau}} = \alpha$, and $\lim_{n\to\infty} \frac{s_n}{n^{-\tau}} = s$. Then*

$$\lim_{n\to\infty} \frac{\log M_1}{n^{\frac{1}{2}-\tau}} = \frac{\kappa^2 s\sqrt{\kappa N_B(1+\kappa N_B)}}{(1-\kappa)^2 N_B(1+(1-\kappa)N_B)} Q^{-1}\left(\frac{1-\delta}{2}\right),$$
$$\lim_{n\to\infty} \frac{\log M_1 K}{n^{\frac{1}{2}-\tau}} = \frac{(1-\kappa)s}{\sqrt{\kappa N_B(1+\kappa N_B)}} Q^{-1}\left(\frac{1-\delta}{2}\right),$$
$$\lim_{n\to\infty} \frac{\log M_2}{n^{\frac{1}{2}}\log n} = \frac{2\tau\kappa\sqrt{\kappa N_B(1+\kappa N_B)}}{(1-\kappa)(1+(1-\kappa)N_B)} Q^{-1}\left(\frac{1-\delta}{2}\right),$$
$$\lim_{n\to\infty} \frac{m}{n^{\frac{1}{2}+\tau}} = \alpha,$$

and, in particular, the number of entangled nats consumed by this protocol is scaling as

$$\lim_{n\to\infty} \frac{m g(s_n)}{n^{\frac{1}{2}}\log n} = \frac{2\tau\sqrt{\kappa N_B(1+\kappa N_B)}}{1-\kappa} Q^{-1}\left(\frac{1-\delta}{2}\right),$$

where $g(x) = (x+1)\log(x+1) - x\log x$.

Note that for channels for which $\kappa > 0.5$, no secret-key bits are needed to ensure that the first layer of coding remains covert. In such cases, our protocol offers strict savings of resources compared to [25], [30]. In other cases, our protocol strictly reduces the number of TMSV pairs required at the expense of the use of secret-key bits. An illustration of the resource trade-off in terms of τ is provided in Fig. 2.

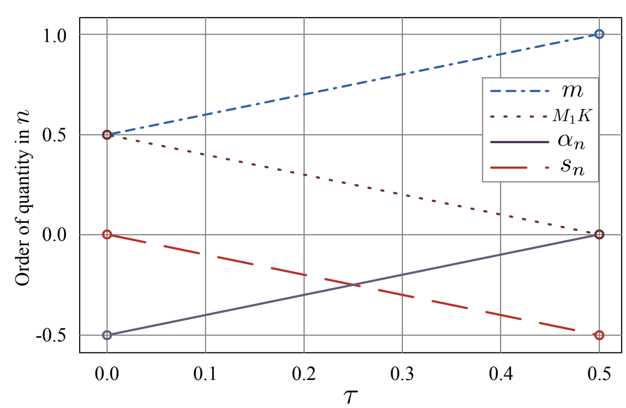

Fig. 2. Illustration of resource trade-off in terms of orders of parameters with blocklength. The sparsity is represented by α_n and the signal power is represented by s_n.

Remark 1. *Even though our results do not extend directly to the case in which either $\alpha_n = \Theta(n^{-\frac{1}{2}})$ or $s_n = \Theta(n^{-\frac{1}{2}})$, we offer some insights on how our idea may be applied to the two extreme regimes.*

1) *When $s_n = \Theta(n^{-\frac{1}{2}})$, our scheme corresponds to the entanglement-assisted covert capacity investigated*

in [25], in which at least n TMSV pairs are consumed and no indexing is required.

2) *When $\alpha_n = \Theta(n^{-\frac{1}{2}})$, our scheme corresponds to the sparse OOK strategy investigated in [22]. In addition, the benefits of the additional $\log n$ scaling in covert and reliable throughput is lost. The reason is that, strictly speaking, the phase encoding effectively utilizes TMSV pairs as cq states, and the associated rate is therefore limited by the Holevo information instead of the entanglement-assisted capacity. This is consistent with [29], [30], which both show that the benefit of phase encoding on TMSV pairs to approach entanglement-assisted capacity only shows up when the signal power is small enough.*

V. Proof of Theorem 3

The trade-off between the sparsity of OOK and the diffuse signal power in terms of the mean photon number N_S of the shared TMSV pairs ψ_{RI} translates into the the two-layer codebook for covertness, in which we set $\alpha_n \in o(1)\cap\omega(n^{-\frac{1}{2}})$ and let $N_S = s_n \in o(1)\cap\omega(n^{-\frac{1}{2}})$.

A Codebook Construction

Let $M_1, M_2, K \in \mathbb{N}_*$. Alice generates two codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ independently, with $|\mathcal{C}_1| = M_1K$ and $|\mathcal{C}_2| = M_2$. Let $\mathcal{X} \triangleq \{x_0, x_1\}$ and $P_X(x) \triangleq (1-\alpha_n)\mathbf{1}\{x=x_0\} + \alpha_n\mathbf{1}\{x=x_1\}$, where x_0 and x_1 represent the off and on symbols of OOK, respectively. The above distribution P_X captures the sparsity level of OOK and generates $n\alpha_n$ pulses on average for n channel uses, but we require a finer control over the number of pulses within each codeword; we create the following auxiliary n-fold distribution $\widetilde{P}_{X^n}$ for $\mathbf{x}\in\mathcal{X}^n$ to throw away the codewords with too small weights:

$$\widetilde{P}_{X^n}(\mathbf{x}) \triangleq \frac{P_X^{\otimes n}(\mathbf{x})\mathbf{1}\{\hat{p}_{\mathbf{x}}(x_1)\geqslant(1-\bar{\mu})\alpha_n\}}{\mathbb{P}\left(\hat{p}_{\mathbf{x}}(x_1)\geqslant(1-\bar{\mu})\alpha_n\right)},$$

where $\bar{\mu}\in(0,1)$. Alice then generates M_1K codewords $\mathbf{x}_{w_1s}$ with $w_1\in[1;M_1]$ and $s\in[1;K]$ according to $\widetilde{P}_{X^n}$ independently at random for $\mathcal{C}_1$. Note that there are at least $\ell \triangleq \lfloor n(1-\bar{\mu})\alpha_n\rfloor$ x_1-pulses for every codeword $\mathbf{x}_{w_1s}$. For codebook $\mathcal{C}_2$, Alice generates M_2 codewords $\boldsymbol{\theta}_{w_2}\in\Theta_n^\ell$ of length ℓ with $w_2\in[1;M_2]$, where Θ_n represents $L_n\triangleq 2^n$-PSK, according to a ℓ-product uniform distribution $P_{\Theta_n}^{\otimes\ell}$ over Θ_n independently at random.

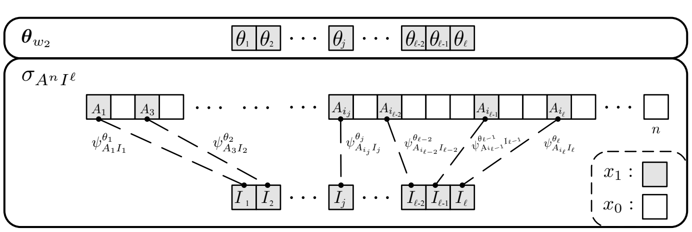

Fig. 3. Illustration of two-layered encoding scheme

The two-layer encoding scheme is illustrated in Fig. 3 and works as follows. Alice first encodes w_1 and s into layer 1 codeword $\mathbf{x}_{w_1s}$, and also encodes w_2 into layer 2 codeword $\boldsymbol{\theta}_{w_2}$. She then performs phase modulation by applying the unitary operators, characterized by $\widehat{U}_{\theta_i}\triangleq\exp\{j\theta_i\hat{a}^\dagger\hat{a}\}$, according to codeword $\boldsymbol{\theta}$ for each $i\in[1;\ell]$ on her share of the entanglement sub-systems R^ℓ, which requires a consumption of ℓ entangled pairs, i.e., $\psi_{R^\ell I^\ell}^{\boldsymbol{\theta}_{w_2}}\triangleq\bigotimes_{i=1}^\ell\psi_{R_iI_i}^{\theta_i}=\bigotimes_{i=1}^\ell((\widehat{U}_{\theta_i}\otimes\mathrm{id}_{I_i})\psi_{R_iI_i}(\widehat{U}_{\theta_i}^\dagger\otimes\mathrm{id}_{I_i}))$. Finally, Alice spreads the state $\psi_{R^\ell I^\ell}^{\boldsymbol{\theta}_{w_2}}$ according to $\mathbf{x}_{w_1s}$ by placing the states in the first ℓ positions of x_1, i.e.,

$$\begin{aligned}\sigma_{A^nI^\ell}^{\mathbf{x}_{w_1s},\boldsymbol{\theta}_{w_2}}&\triangleq\psi_{A_1I_1}^{\theta_1}\otimes|0\rangle\langle0|_{A_2}\otimes\psi_{A_3I_2}^{\theta_2}\otimes|0\rangle\langle0|_{A_4}\otimes\cdots\\&\otimes\psi_{A_{i_j}I_j}^{\theta_j}\otimes|0\rangle\langle0|_{A_{i_j+1}}\otimes\psi_{A_{i_{\ell-2}}I_{\ell-2}}^{\theta_{\ell-2}}\otimes\cdots\otimes|0\rangle\langle0|_{A_n},\end{aligned}$$

where $\{i_j\}_j\subset[1;n]$ represents the first ℓ indices of the x_1-pulse positions with $j\in[1;\ell]$. Equivalently, the two-layer encoder maps the message and the secret key to the following cq state: $\sigma_{X^n\Theta^\ell A^nI^\ell}(w_1,w_2,s)\triangleq|\mathbf{x}_{w_1s}\rangle\langle\mathbf{x}_{w_1s}|_{X^n}\otimes|\boldsymbol{\theta}_{w_2}\rangle\langle\boldsymbol{\theta}_{w2}|_{\Theta^\ell}\otimes\sigma_{A^nI^\ell}^{\mathbf{x}_{w_1s},\boldsymbol{\theta}_{w_2}}$. Note that without access to I^ℓ, the reduced state

$$\begin{aligned}\sigma_{A^n}^{\mathbf{x}_{w_1s},\boldsymbol{\theta}_{w_2}}&=\mathrm{tr}_{I^\ell}\left(\sigma_{A^nI^\ell}^{\mathbf{x}_{w_1s},\boldsymbol{\theta}_{w_2}}\right)\\&=\rho_{s_n,A_1}\otimes|0\rangle\langle0|_{A_2}\otimes\rho_{s_n,A_3}\otimes|0\rangle\langle0|_{A_4}\otimes\cdots\\&\otimes\rho_{s_n,A_{i_j}}\otimes|0\rangle\langle0|_{A_{i_j+1}}\otimes\rho_{s_n,A_{i_{\ell-2}}}\otimes\cdots\otimes|0\rangle\langle0|_{A_n},\end{aligned}$$

where ρ_{s_n} is a thermal state with mean photon number s_n, and $\sigma_{A^n}^{\mathbf{x}_{w_1s},\boldsymbol{\theta}_{w_2}}$ completely loses the phase information $\boldsymbol{\theta}_{w_2}$.

Note also that the expected cq state over the generation of random codebook $\mathcal{C}_1$ is $\mathbb{E}_{\mathcal{C}_1}\left[|\mathbf{X}\rangle\langle\mathbf{X}|_{X^n}\otimes\sigma_{A^n}^{\mathbf{X}}\right]=\sum_{\mathbf{x}}\widetilde{P}_{X^n}(\mathbf{x})|\mathbf{x}\rangle\langle\mathbf{x}|_{X^n}\otimes\sigma_{A^n}^{\mathbf{x}}\triangleq\tilde{\sigma}_{X^nA^n}$. Lemma 5 characterizes the trace distance between $\tilde{\sigma}_{X^nA^n}$ and $\sigma_{XA}^{\otimes n}$, where $\sigma_{XA}\triangleq\sum_xP_X(x)|x\rangle\langle x|_X\otimes\sigma_A^x=(1-\alpha_n)|x_0\rangle\langle x_0|_X\otimes|0\rangle\langle0|_A+\alpha_n|x_1\rangle\langle x_1|_X\otimes\rho_{s_n,A}$.

Lemma 5. *For n large enough, there exists some $c_{\bar{\mu}}>0$ such that $\frac{1}{2}\left\|\tilde{\sigma}_{X^nA^n}-\sigma_{XA}^{\otimes n}\right\|_1\leqslant\exp(-c_{\bar{\mu}}n\alpha_n)$ and $\frac{1}{2}\left\|\tilde{\sigma}_{X^n}\otimes\tilde{\sigma}_{A^n}-\sigma_X^{\otimes n}\otimes\sigma_A^{\otimes n}\right\|_1\leqslant2\exp(-c_{\bar{\mu}}n\alpha_n)$*

Proof: This is a consequence of the Chernoff bound and the monotonicity of the trace distance. ■

B. Reliability and Soft-Covering Analysis

Observe that after n uses of $\mathcal{L}_{A\to BW}^{(\kappa,N_B)}$, the global joint state containing Bob's systems is $\sigma_{X^n\Theta^\ell B^nI^\ell}(w_1,w_2,s)\triangleq\mathrm{tr}_{W^n}\left(((\mathcal{L}_{A\to BW}^{(\kappa,N_B)})^{\otimes n}\otimes\mathrm{id}_{I^\ell})\sigma_{X^n\Theta^\ell A^nI^\ell}(w_1,w_2,s)\right)$, while the global joint state containing Willie's terminal is $\sigma_{X^n\Theta^\ell W^n}(w_1,w_2,s)\triangleq\mathrm{tr}_{B^n}\left((\mathcal{L}_{A\to BW}^{(\kappa,N_B)})^{\otimes n}\sigma_{X^n\Theta^\ell A^n}(w_1,w_2,s)\right)=|\boldsymbol{\theta}_{w_2}\rangle\langle\boldsymbol{\theta}_{w_2}|_{\Theta^\ell}\otimes\sigma_{X^nW^n}(w_1,s)$, which shows that Willie's observation W^n is decoupled from Θ^ℓ. Similarly, when Bob chooses to ignore systems I^ℓ, his observations are also decoupled from Θ^ℓ as $\sigma_{X^nB^n}(w_1,s)$. We then recall the following one-shot channel reliability and soft-covering results via position-based coding and convex splitting lemma.

Lemma 6 (One-shot Channel Reliability and Soft-covering Adapted from [26], [39]–[42])**.** *Fix $\epsilon_1\in(0,1)$, $\eta\in(0,\delta)$, $\gamma_1\in(0,\epsilon_1)$, $\gamma_2\in(0,\frac{\eta}{2})$, and $\gamma_3\in(0,\frac{\eta}{2}-\gamma_2)$. Then for a cq* *state ρ_{XA} and a channel $\mathcal{G}:\rho_{XA}\mapsto\rho_{XBW}$, there exists a coding scheme such that* $\log M \geqslant \mathbb{D}_{\mathrm{H}}^{\epsilon_1-\gamma_1}(\rho_{XB}\,\|\,\rho_X\otimes\rho_B) - \log\left(4\epsilon_1\gamma_1^{-2}\right)$, $\log MK \leqslant \mathbb{D}_{\max}^{\eta/2-\gamma_2-\gamma_3}(\rho_{XW}\,\|\,\rho_X\otimes\rho_W) - 2\log(\gamma_2) + \log\left(8\gamma_3^{-2}\right)$, $\mathbb{E}_{\mathcal{C},S}[\mathbb{P}(\widehat{W}\neq W|S)] \leqslant \epsilon_1$, *and* $\mathbb{E}_{\mathcal{C}}[\frac{1}{2}\|\hat{\rho}_W-\rho_W\|_1] \leqslant \eta-\gamma_2$, *where $\mathcal{C}$ is the codebook and $\hat{\rho}_W$ is induced by the codebook $\mathcal{C}$.*

Proof: Omitted due to space constraint. ■

We now specialize the above result for layer 1 codebook $\mathcal{C}_1$. Bob first ignores his share of entangled systems I^ℓ, and the expected cq state of Bob over the generation of random codebook $\mathcal{C}_1$ is $\tilde{\sigma}_{X^nB^n} \triangleq \mathbb{E}_{\mathcal{C}_1}[|\mathbf{X}\rangle\langle\mathbf{X}|_{X^n}\otimes \mathrm{tr}_{W^n}((\mathcal{L}_{A\to BW}^{(\kappa,N_B)})^{\otimes n}\sigma_{A^n}^{\mathbf{X}})]$. Similarly, since Willie has no access to I^ℓ, the expected cq state of Willie over the generation of random codebook $\mathcal{C}_1$ is $\tilde{\sigma}_{X^nW^n} \triangleq \mathbb{E}_{\mathcal{C}_1}[|\mathbf{X}\rangle\langle\mathbf{X}|_{X^n}\otimes \mathrm{tr}_{B^n}((\mathcal{L}_{A\to BW}^{(\kappa,N_B)})^{\otimes n}\sigma_{A^n}^{\mathbf{X}})]$. On the other hand, the observation induced by $\mathcal{C}_1$ at Willie's terminal is $\hat{\sigma}_{W^n}$. We can then apply Lemma 6 and obtain

$$\log M_1 \geqslant \mathbb{D}_{\mathrm{H}}^{\epsilon_1-n^{-1/2}}(\tilde{\sigma}_{X^nB^n}\,\|\,\tilde{\sigma}_{X^n}\otimes\tilde{\sigma}_{B^n}) - \mathcal{O}(\log n), \tag{1}$$

$$\log M_1K \leqslant \mathbb{D}_{\max}^{\eta/2-2n^{-1/2}}(\tilde{\sigma}_{X^nW^n}\,\|\,\tilde{\sigma}_{X^n}\otimes\tilde{\sigma}_{W^n}) + \mathcal{O}(\log n), \tag{2}$$

such that

$$\mathbb{E}_{\mathcal{C}_1,S}[\mathbb{P}(\widehat{W}_1\neq W_1|S)] \leqslant \epsilon_1, \tag{3}$$

$$\mathbb{E}_{\mathcal{C}_1}\left[\frac{1}{2}\|\hat{\sigma}_{W^n}-\tilde{\sigma}_{W^n}\|_1\right] \leqslant \eta - n^{-\frac{1}{2}}, \tag{4}$$

where we have chosen $\gamma_1=\gamma_2=\gamma_3=n^{-\frac{1}{2}}$. We then specify α_n and s_n to ensure that $\sigma_{n,W}^{\otimes n}\triangleq\sigma_W^{\otimes n}$ and $\sigma_{0,W}^{\otimes n}$ are close.

Lemma 7. *The trace distance between $\sigma_{n,W}^{\otimes n}$ and $\sigma_{0,W}^{\otimes n}$ is*

$$\begin{aligned}\frac{1}{2}\left\|\sigma_{n,W}^{\otimes n}-\sigma_{0,W}^{\otimes n}\right\|_1 \leqslant 1-2Q\,(&\frac{\sqrt{n}(1-\kappa)\alpha_n s_n}{2\sqrt{\kappa N_B(1+\kappa N_B)}})\\ &+D_0\sqrt{n}\alpha_n s_n^3 + D_1 n^{-1/2}.\end{aligned} \tag{5}$$

Proof: Adapted from [26, Lemma IV.1]. ■

Choose $\alpha_n s_n = \frac{2\sqrt{\kappa N_B(1+\kappa N_B)}}{(1-\kappa)}Q^{-1}\left(\frac{1-(\delta-\eta)}{2}\right)n^{-\frac{1}{2}} - \bar{D}s_n^2 n^{-1/2} - \mathcal{O}(n^{-1})$, where $\bar{D}$ is some constant depending on the channel to ensure that

$$\frac{1}{2}\|\sigma_{n,W}^{\otimes n}-\sigma_{0,W}^{\otimes n}\|_1 \leqslant \delta-\eta-s_n^2+\mathcal{O}(n^{-1/2}). \tag{6}$$

Therefore, by combining (6), triangle inequality, and Lemma 5, we obtain

$$\frac{1}{2}\|\tilde{\sigma}_{W^n}-\sigma_{0,W}^{\otimes n}\|_1 \leqslant \delta-\eta. \tag{7}$$

Similar to the analysis for layer 1, we use Lemma 6 for reliability of layer 2 codebook $\mathcal{C}_2$. However, because of the perturbation caused by the decoding POVM for layer 1, we have to account for the perturbation by the *gentle measurement lemma* [43] to ensure compatibility and include it as part of the decoding error for layer 2. Consequently,

$$\log M_2 \geqslant \mathbb{D}_{\mathrm{H}}^{\epsilon_2-2\sqrt{\epsilon_1}-n^{-1/2}}(\sigma_{\Theta BI}^{\otimes\ell}\,\|\,\sigma_{\Theta}^{\otimes\ell}\otimes\sigma_{BI}^{\otimes\ell}) - \mathcal{O}(\log n), \tag{8}$$

such that

$$\mathbb{E}_{\mathcal{C}_2}[\widehat{W}_2\neq W_2\Big|\widehat{W}_1=W_1] \leqslant \epsilon_2-2\sqrt{\epsilon_1}, \tag{9}$$

where $\sigma_{\Theta BI}\triangleq\mathbb{E}_{P_\Theta}[\mathrm{tr}_W(|\Theta\rangle\langle\Theta|_\Theta\otimes(\mathcal{L}_{A\to BW}^{(\kappa,N_B)}\otimes\mathrm{id}_I)\psi_{AI}^{\Theta})]$. By combining (3), (9) and the gentle measurement lemma, $\mathbb{E}_{\mathcal{C}_1\mathcal{C}_2}\left[\mathbb{P}\left(\hat{W}_1\neq W_1 \text{ or } \widehat{W}_2\neq W_2\Big|S\right)\right] \leqslant \epsilon_1+\epsilon_2\leqslant\epsilon$. By combining (4), (6) and (7), $\mathbb{E}_{\mathcal{C}_1\mathcal{C}_2}\left[\frac{1}{2}\|\hat{\sigma}_{W^n}-\sigma_{0,W}^{\otimes n}\|_1\right] \leqslant \delta$. Therefore, by applying a derandomization argument similar to the one in [40, Section 4.2], there exists a two-layer coding scheme with codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ as desired.

C. Throughput Analysis

To obtain the asymptotics of (1) and (2), we ignore the second order asymptotics and the associated information quantities due to space constraint. The following lemma provides information quantities involved in the characterization of first-order asymptotics. Note that the truncation effect characterized by Lemma 5 creates negligible difference from the asymptotics computed by the corresponding n-product states.

Lemma 8. *Let $\sigma_{XBW}\triangleq\mathcal{L}_{A\to BW}^{(\kappa,N_B)}(\sigma_{XA})$. Then*

$$\begin{aligned}\mathbb{D}(\sigma_{XB}\,\|\,\sigma_X\otimes\sigma_B) &= \frac{\kappa^2\alpha_n s_n^2}{2(1-\kappa)N_B(1+(1-\kappa)N_B)}\\ &\quad-\mathcal{O}(\alpha_n^2 s_n^2,\alpha_n s_n^3),\\ \mathbb{D}(\sigma_{XW}\,\|\,\sigma_X\otimes\sigma_W) &= \frac{(1-\kappa)^2\alpha_n s_n^2}{2\kappa N_B(1+\kappa N_B)} - \mathcal{O}(\alpha_n^2 s_n^2,\alpha_n s_n^3),\end{aligned}$$

Proof: Omitted due to space constraint. ■

Therefore,

$$\log M_1 \geqslant \frac{n\kappa^2\alpha_n s_n^2}{2(1-\kappa)N_B(1+(1-\kappa)N_B)} + o(\sqrt{n}), \tag{10}$$

$$\log M_1K \leqslant \frac{n(1-\kappa)^2\alpha_n s_n^2}{2\kappa N_B(1+\kappa N_B)} + o(\sqrt{n}). \tag{11}$$

For the phase modulation on layer 2, we show that as we enlarge the set of phase modulation L_n, there exists a state $\tilde{\sigma}_{BI}$ that represents the mixed state coming from a uniform and continuous phase ensemble over $[0,2\pi]$. Then the code rate also converges to the Holevo information corresponding to this ensemble [29, Eq. (11)].

Lemma 9. *There exists $\tilde{\sigma}_{\Theta BI}=\lim_{L_n\to\infty}\sigma_{\Theta BI}$ in the sense of trace distance, and* $\mathbb{D}(\sigma_{\Theta BI}\,\|\,\sigma_\Theta\otimes\sigma_{BI}) \geqslant -\frac{\kappa}{1+(1-\kappa)N_B}s_n\log s_n + \mathcal{O}(s_n) + \mathcal{O}(2^{4n-3\times 2^n\log_2 s_n^{-1}-3\times 2^n\log_2\frac{(1-\kappa)N_B+1}{\kappa}})$ *for large enough n.*

Proof: The proof follows from combining [44, Theorem 2 and 5] and [29, Appendix B]. ■

Finally,

$$\log M_2 \geqslant (1-\mu)\frac{-\kappa n\alpha_n s_n\log s_n}{1+(1-\kappa)N_B} + o(\sqrt{n}\log n). \tag{12}$$

The result follows by taking $m=\ell=\lfloor n(1-\bar{\mu})\alpha_n\rfloor$, μ such that $(1-\mu)=\frac{m}{n\alpha_n}$, and η from (4) and (7) is arbitrarily small.

References

[1] B. A. Bash, D. Goeckel, D. Towsley, and S. Guha, “Hiding information in noise: fundamental limits of covert wireless communication,” *IEEE Communications Magazine*, vol. 53, no. 12, pp. 26–31, Dec. 2015.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] K. S. K. Arumugam and M. R. Bloch, “Embedding covert information in broadcast communications,” *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2787–2801, Oct. 2019.

[7] ——, “Covert communication over a k-user multiple access channel,” *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7020–7044, Nov. 2019.

[8] K. S. K. Arumugam, M. R. Bloch, and L. Wang, “Covert communication over a physically degraded relay channel with non-colluding wardens,” in *Proc. of IEEE International Symposium on Information Theory*, Vail, CO, Jun. 2018, pp. 766–770.

[9] 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.

[10] A. Bounhar, M. Sarkiss, and M. Wigger, “Mixing a covert and a non-covert user,” in *Proc. of IEEE International Symposium on Information Theory*. Taipei, Taiwan: IEEE, Jun. 2023, pp. 2577–2582.

[11] L. Wang, “On gaussian covert communication in continuous time,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, p. 283, Dec. 2019.

[12] Q. Zhang, M. 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.

[13] C. Bouette, L. Luzzi, and L. Wang, “Covert communication over two types of additive noise channels,” in *Proc. of IEEE Information Theory Workshop*, Saint-Malo, France, Apr. 2023.

[14] P. H. Che, M. Bakshi, C. Chan, and S. Jaggi, “Reliable deniable communication with channel uncertainty,” in *Proc. of IEEE Information Theory Workshop*, Hobart, Tasmania, November 2014, pp. 30–34.

[15] B. A. Bash, D. Goeckel, and D. Towsley, “Covert communication gains from adversary’s ignorance of transmission time,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8394–8405, Dec. 2016.

[16] K. S. K. Arumugam and M. R. Bloch, “Keyless asynchronous covert communication,” in *Proc. of IEEE Information Theory Workshop*, Cambridge, United Kingdom, Sep. 2016, pp. 191–195.

[17] 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, 2017.

[18] S. H. Lee, L. Wang, A. Khisti, and G. W. Wornell, “Covert communication with channel-state information at the transmitter,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2310–2319, Sep. 2018.

[19] H. Zivari-Fard, M. Bloch, and A. Nosratinia, “Keyless covert communication via channel state information,” *IEEE Transactions on Information Theory*, vol. 68, no. 8, pp. 5440–5474, Aug. 2022.

[20] 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, pp. –, October 2015.

[21] L. Wang, “Optimal throughput for covert communication over a classical-quantum channel,” in *Proc. of IEEE Information Theory Workshop*, Cambridge, UK, September 2016, pp. 364–368.

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

[23] E. Zlotnick, B. Bash, and U. Pereg, “Entanglement-assisted covert communication via qubit depolarizing channels,” in *Proc. of IEEE International Symposium on Information Theory*. Taipei, Taiwan: IEEE, Jun. 2023, pp. 198–203.

[24] M. S. Bullock, C. N. Gagatsos, S. Guha, and B. A. Bash, “Fundamental limits of quantum-secure covert communication over bosonic channels,” *IEEE Journal of Selected Areas in Communications*, vol. 38, no. 3, pp. 471–482, mar 2020.

[25] 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, aug 2020.

[26] S.-Y. Wang, T. Erdoğan, and M. R. Bloch, “Towards a characterization of the covert capacity of bosonic channels under trace distance,” in *Proc. of IEEE International Symposium on Information Theory*, Helsinki, Finland, Jun. 2022, pp. 354–359.

[27] R. Di Candia, H. Yiğitler, G. Paraoanu, and R. Jäntti, “Two-way covert quantum communication in the microwave regime,” *PRX Quantum*, vol. 2, no. 2, p. 020316, may 2021.

[28] 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.

[29] H. Shi, Z. Zhang, and Q. Zhuang, “Practical route to entanglement-assisted communication over noisy bosonic channels,” *Physical Review Applied*, vol. 13, no. 3, p. 034029, mar 2020.

[30] A. Cox, Q. Zhuang, C. N. Gagatsos, B. Bash, and S. Guha, “Transceiver designs approaching the entanglement-assisted communication capacity,” *Physical Review Applied*, vol. 19, no. 6, p. 064015, jun 2023.

[31] M. M. Wilde, P. Hayden, and S. Guha, “Quantum trade-off coding for bosonic communication,” *Phys. Rev. A*, vol. 86, p. 062306, Dec 2012.

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

[33] 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.

[34] 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.

[35] P. W. Shor, “The classical capacity achievable by a quantum channel assisted by a limited entanglement,” *Quantum Information & Computation*, vol. 4, no. 6, pp. 537–545, Dec. 2004.

[36] C. W. Helstrom, *Quantum Detection and Estimation Theory*. New York, NY, USA: Academic Press, 1976.

[37] A. S. Holevo, “Statistical decision theory for quantum systems,” *Journal of Multivariate Analysis*, vol. 3, no. 4, pp. 337–394, Dec. 1973.

[38] M. M. Wilde, *Quantum Information Theory*. Cambridge: Cambridge University Press, 2017.

[39] 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.

[40] 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.

[41] 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.

[42] S. Khatri, E. Kaur, S. Guha, and M. M. Wilde, “Second-order coding rates for key distillation in quantum key distribution,” *arXiv preprint*, vol. 1910.03883, Oct. 2019.

[43] A. Winter, “Coding theorem and strong converse for quantum channels,” *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2481–2485, Nov. 1999.

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