Covert Entanglement Generation over Bosonic Channels

Evan J. D. Anderson, Michael S. Bullock, Ohad Kimelfeld, Christopher K. Eyre, Filip Rozpędek, Uzi Pereg, and Boulat A. Bash

***Abstract*—We explore covert entanglement generation over the lossy thermal-noise bosonic channel, which is a quantum-mechanical model of many practical settings, including optical, microwave, and radio-frequency (RF) channels. Covert communication ensures that an adversary is unable to detect the presence of transmissions, which are concealed in channel noise. We show that a *square root law* (SRL) for covert entanglement generation similar to that for classical: $L_{\rm EG}\sqrt{n}$ entangled bits (ebits) can be generated covertly and reliably over n uses of a bosonic channel. We report a single-letter expression for optimal $L_{\rm EG}$ as well as an achievable method. We additionally analyze the performance of covert entanglement generation using single- and dual-rail photonic qubits, which may be more practical for physical implementation.**

I. Introduction

Standard communication security protects the transmission content from unauthorized access using cryptography [3] or information-theoretic secrecy [4]. On the other hand, covert, or low probability of detection/intercept (LPD/LPI), signaling prevents detection of the transmission in the first place. Over the last decade, the fundamental limits of covert communication were explored for classical [5]–[8] and classical-quantum channels [9]–[13]. Covert communication over these channels is governed by the *square root law* (SRL): $L_{\rm c}\sqrt{n}$ covert bits are reliably transmissible over n channel uses for a channel-dependent constant $L_{\rm c} > 0$ called covert classical channel capacity. A tutorial [14] and a detailed survey [15] overview these results and their developments.

To date, most of the covert communication effort focused on classical data, i.e., transmission of bits. Motivated by recent advances in quantum communication [16], [17], in this paper we explore covert quantum communication. Specifically, we investigate covert entanglement generation over the lossy thermal-noise bosonic channels, which we call "bosonic channels" for brevity and formally define in Section III-A. They describe quantum-mechanically optical-fiber, free-space-optical (FSO), microwave, and radio-frequency (RF) communication channels. Notably, unlike standard qubits, which are finite-dimensional, the quantum states of light on which these channels act reside in infinite-dimensional Hilbert spaces.

We report an SRL similar to classical communication for covert entanglement generation over the bosonic channels: $L_{\rm EG}\sqrt{n}$ covert entangled bits (ebits) can be generated over n channel uses for a channel-dependent constant $L_{\rm EG} > 0$ called covert entanglement-generation capacity. We characterize the optimal value by providing 1) an upper bound on $L_{\rm EG}$, and 2) a method to generate $L_{\rm EG}\sqrt{n}$ ebits covertly using a classical secret key that is pre-shared between communicating parties. Our approach adapts the covert entanglement generation method for finite-dimensional channels from [18] to infinite-dimensional bosonic channels. This non-trivial derivation builds on the results in [10], [11], [19], [20]. In fact, the covert entanglement generation capacity $L_{\rm EG} = L_{\text{no-EA}}$, the unassisted covert classical capacity derived in [11], which agrees with the finite-dimensional result in [18]. In the process, we also obtain the optimal scheme for ensuring information-theory secrecy of covert classical communication. This prevents leakage of information contained in the transmission in the rare event that it is detected.

Remarkably, our formula $L_{\rm EG}$ has a single-letter form, while only the bounds for the non-covert quantum capacity of a lossy thermal-noise bosonic channel are currently known [21]–[24]. However, we note that the non-covert entanglement generation capacity equals the quantum capacity because entanglement enables teleportation of quantum states with assistance of a classical communication channel. In a covert quantum communication system, classical channel uses must also be covert, necessitating careful analysis that we defer to future work.

We also analyze the performance of sub-optimal but more practical covert entanglement generation methods using single-

This paper was presented in part at the IEEE International Conference on Quantum Computing and Engineering (QCE) in Montréal, QC, Canada, September 2024 [1] and at the IEEE International Symposium on Information Theory (ISIT) in Ann Arbor, MI, USA, June 2025 [2].

This work was supported by the National Science Foundation under Grants No. CCF-2006679 and EEC-1941583, the Israel Science Foundation under Grants No. 939/23 and 2691/23, German-Israeli Project Cooperation (DIP) within the Deutsche Forschungsgemeinschaft (DFG) under Grant No. 2032991, Ollendorff Minerva Center (OMC) of the Technion No. 86160946, and by the Junior Faculty Program for Quantum Science and Technology of Israel Planning and Budgeting Committee of the Council for Higher Education (VATAT) under Grant No. 86636903.

Evan Anderson is with the Wyant College of Optical Sciences, University of Arizona Tucson, AZ, USA (email: ejdanderson@arizona.edu)

Michael Bullock is with Dept. of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, USA (email: bullockm@arizona.edu)

Ohad Kimelfeld is with the Physics Department and Helen Diller Quantum Center, Technion – Israel Institute of Technology, Haifa, Israel (email: ohad.kim@campus.technion.ac.il)

Christopher Eyre is with the Dept. of Mathematics, Brigham Young University Provo, UT, USA

Filip Rozpędek is with the College of Information & Computer Sciences, University of Massachusetts, Amherst, MA, USA (email: frozpedek@cics.umass.edu)

Uzi Pereg is with the Electrical and Computer Engineering Department and Helen Diller Quantum Center, Technion – Israel Institute of Technology, Haifa, Israel (email: uzipereg@technion.ac.il)

Boulat Bash is with Dept. of Electrical and Computer Engineering, and the Wyant College of Optical Sciences, University of Arizona, Tucson, AZ, USA (email: boulat@arizona.edu)

and dual-rail photonic qubits. Unfortunately, the complexity of our optimal entanglement-generation scheme makes its realization using known components extremely challenging. Thus, we explore alternatives using single- and dual-rail photonic qubit modulation. These qubit encodings are used in many quantum information processing tasks, including cluster-state generation [25], entanglement distribution in quantum networks [26]–[31], and quantum key distribution (QKD) [32], [33]. We derive the achievable entanglement generation rates by adapting our covert classical-quantum channel analysis [9]–[13] and report a significant gap from the optimal, motivating future investigation.

The rest of this paper is organized as follows: next we discuss the prior work, and in Section III we state the mathematical preliminaries as well as the system and channel models. In Section IV we provide the ultimate limits of covert entanglement generation. In Section V, we explore using single- and dual-rail photonic qubit encodings. In Section VI we compare performance of our methods. We conclude in Section VII by discussing the implications of our results and areas of future research.

II. Relationship to Prior Work

Initial studies of covert quantum communication focus on covert quantum key distribution (QKD) [34]–[37]. Thus, they do not provide the fundamental limits of covert entanglement generation studied here. Indeed, they do not investigate quantum communication between parties that already possess covert classical resources such as a pre-shared classical secret and a covert classical channel. Our initial work, presented in conferences [1], [2], takes a direct approach in exploring the achievability of quantum covert communication. In [1] we also provide a very loose converse result (which this paper substantially improves by the virtue of equivalence of quantum communication and the combination of entanglement and classical communication). Exploration of covert quantum and classical communication over finite-dimensional quantum channels in [2], [12], [13] inspired the derivation of the fundamental limits for entanglement generation in [18]. Adapting it to infinite-dimensional bosonic channel is not trivial, as illustrated in Section IV.

III. Preliminaries

A. System and Channel Model

Consider the covert communication setting described in Fig. 1. Alice wishes to transmit a quantum state $\hat{\rho}_{A^n}$ to Bob over n uses of the lossy thermal-noise bosonic channel without being detected by an adversarial warden Willie. For covert secret communication, Alice encodes a message m into the A^n subsystems through classical modulation, and seeks to keep this secret from Willie while allowing Bob to decode the message reliably. For covert entanglement generation, she instead creates local entanglement and maps half of the entangled state to the A^n subsystems. Alice seeks to ensure Bob can reliably obtain half of an entangled state, where the other half is maintained locally by her.

The *lossy thermal-noise bosonic channel*, $\mathcal{E}_{A\to BW}^{\eta,\bar{n}_{\mathrm{B}}}$ shown in Figure 2a, is described by a beamsplitter with transmittance $\eta \in [0,1]$, two input modes (Alice and the environment with thermal-state input), and two output modes (Bob and Willie). The thermal state has a photon number basis representation $\hat{\rho}_{\mathrm{th}}(\bar{n}_{\mathrm{B}}) \equiv \sum_{k=0}^{\infty} \frac{\bar{n}_{\mathrm{B}}^k}{(1+\bar{n}_{\mathrm{B}})^{k+1}} |k\rangle\langle k|$. For her input state $\hat{\rho}_{A^n}$, Bob and Willie receive $\hat{\rho}_{B^n} \equiv \mathrm{tr}_{W^n}\left(\mathcal{E}_{A\to BW}^{(\eta,\bar{n}_{\mathrm{B}})\otimes n}(\hat{\rho}_{A^n})\right)$ and $\hat{\rho}_{W^n} \equiv \mathrm{tr}_{B^n}\left(\mathcal{E}_{A\to BW}^{(\eta,\bar{n}_{\mathrm{B}})\otimes n}(\hat{\rho}_{A^n}^{(\psi)})\right)$, respectively. We equivalently represent the bosonic channel via its Stinespring dilation $V_{A\to BWE}^{\eta,\bar{n}_B}$ shown in Figure 2b, where $\hat{\rho}_{B^n} \equiv \mathrm{tr}_{W^nE^n}\left(V_{A\to BWE}^{(\eta,\bar{n}_{\mathrm{B}})\otimes n}(\hat{\rho}_{A^n})\right)$ and $\hat{\rho}_{W^n} \equiv \mathrm{tr}_{B^nE^n}\left(V_{A\to BWE}^{(\eta,\bar{n}_{\mathrm{B}})\otimes n}(\hat{\rho}_{A^n})\right)$. We assume that Alice and Bob pre-share a classical secret, as is standard in covert communications [6]–[14].

B. Hypothesis Testing and Covertness

We assume that Willie has complete knowledge of the system in Fig. 1, except for Alice and Bob's pre-shared secret. Willie must determine from his channel output whether Alice is using the channel (hypothesis H_1) or not (hypothesis H_0). Let $\hat{\rho}_{W^n}^{(0)} \equiv \mathrm{tr}_B\left(\mathcal{E}_{A\to BW}^{(\eta,\bar{n}_{\mathrm{B}})\otimes n}(|0\rangle\langle 0|_A^{\otimes n})\right)$ be the state Willie observes when Alice is quiet, where $|0\rangle\langle 0|_A$ is a vacuum state and is the "innocent" input, similarly, we denote $\hat{\rho}_{W^n}^{(1)}$ as the state Willie receives when Alice transmits. As vacuum is input when Alice is quiet, $\hat{\rho}_{W^n}^{(0)} = \left(\hat{\rho}_W^{(0)}\right)^{\otimes n}$ where $\hat{\rho}_W^{(0)} = \hat{\rho}_{\mathrm{th}}(\eta\bar{n}_{\mathrm{B}})$ is an attenuated thermal state [38].

Willie desires to determine if Alice and Bob are communicating. Therefore, over n channel uses, he attempts to distinguish between $\hat{\rho}_{W^n}^{(1)}$ and $\hat{\rho}_{W^n}^{(0)}$. The null and alternate hypotheses H_0 and H_1 correspond to Alice being quiet and transmitting, respectively. Willie uses arbitrary quantum resources to discriminate between H_0 and H_1, including fault-tolerant quantum computers, perfect quantum measurement, and ideal quantum memories. We additionally allow Willie to collect any photons that do not reach Bob.

Assuming equal priors, i.e., $P(H_1) = P(H_0) = \frac{1}{2}$, Willie's probability of error is $P_W^{(e)} = \frac{P_{\mathrm{FA}}+P_{\mathrm{MD}}}{2}$, with probability of false alarm $P_{\mathrm{FA}} = P(\text{choose } H_1|H_0 \text{ true})$ and probability of missed detection $P_{\mathrm{MD}} = P(\text{choose } H_0|H_1 \text{ true})$. $P_W^{(e)} \leq \frac{1}{2}$ is the trivial upper bound Willie can achieve by using a random decision device. Thus, Alice and Bob try to ensure that Willie's minimum probability of error is close to that of this ineffective device. Formally, they seek $P_W^{(e)} \geq \frac{1}{2} - \delta$, where $\delta > 0$ quantifies the desired level of covertness. Willie's minimum probability of error is bounded by trace distance between his output states under each hypothesis as [16, Sec. 9.1.4]: $P_W^{(e)} \geq \frac{1}{2} - \frac{1}{4}\left\|\hat{\rho}_{W^n} - \hat{\rho}_{W^n}^{(0)}\right\|_1$, where $\|\cdot\|_1$ is the trace norm. The trace distance is often mathematically unwieldy. Conveniently, the quantum relative entropy (QRE) $D\left(\hat{\rho}\|\hat{\sigma}\right) \equiv \mathrm{tr}\left(\hat{\rho}\log\hat{\rho} - \hat{\rho}\log\hat{\sigma}\right)$ is additive over product states, and upper bounds the trace distance via the quantum Pinsker's inequality [16, Th. 11.9.1]: $\frac{1}{4}\left\|\hat{\rho}_{W^n} - \hat{\rho}_{W^n}^{(0)}\right\|_1 \leq \sqrt{\frac{1}{8}D\left(\hat{\rho}_{W^n}\middle\|\hat{\rho}_{W^n}^{(0)}\right)}$. We use QRE as our covertness criterion, as is standard in both

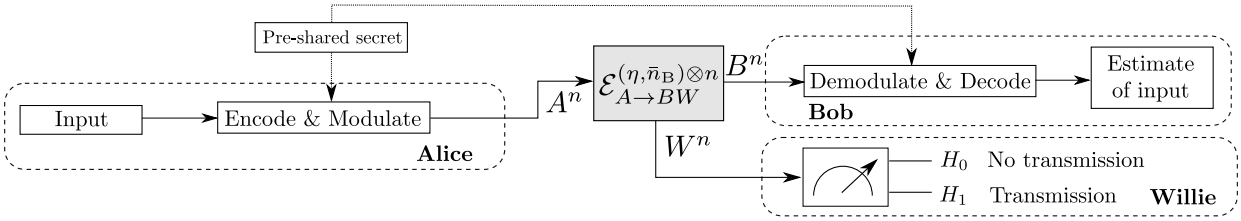


Fig. 1. System model for covert secrecy and covert entanglement generation. Alice either has an input and transmits, or she is quiet. When she has an input, she encodes and modulates a state of system A^n before transmitting it over n uses of the lossy thermal-noise bosonic channel $\mathcal{E}^{\eta,\bar{n}_{\rm B}}_{A\to BW}$ depicted in Fig. 2a. Bob demodulates and decodes to estimate the input. Warden Willie has to decide between hypotheses H_0 and H_1 corresponding to a quiet or transmitting Alice. For covert secrecy, Alice uses position-based coding with a pre-shared secret key k unknown to Willie to encode message m in a QPSK-modulated coherent-state codeword. Bob employs sequential decoding with the pre-shared secret k to estimate m. For entanglement generation, Alice prepares a maximally-entangled state $|\Phi\rangle\langle\Phi|_{RM}$. She encodes the state of system M as a superposition of the codewords from the aforementioned secrecy codebook in system A^n. Bob constructs a coherent version of the sequential decoding scheme used in the secrecy construction to recover a state entangled with the reference system R.

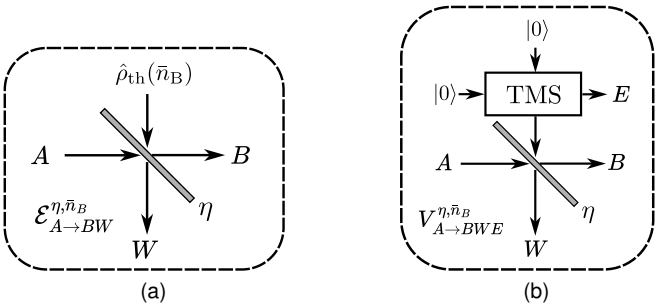


Fig. 2. Bosonic channel model. a) is the lossy thermal-noise bosonic channel $\mathcal{E}^{\eta,\bar{n}_{\rm B}}_{A\to BW}$, with input subsystem at Alice A and output subsystems B and W at Bob and Willie respectively. $\hat{\rho}_{\rm th}(\bar{n}_{\rm B})$ is a thermal state. b) is the Stinespring dilation $V^{\eta,\bar{n}_{\rm B}}_{A\to BWE}$ of the bosonic channel, with an ancillary output subsystem E. TMS is two-mode squeezer with gain $G = 1 + \bar{n}_B$.

classical [7], [8] and quantum [10]–[13] analyses. Formally, we call a communication scheme *covert* if, for $\delta_{\rm C} > 0$, $D\left(\hat{\rho}_{W^n}\middle\|\hat{\rho}^{(0)}_{W^n}\right) \leq \delta_{\rm C}$.

C. Secrecy

Secrecy in communication imposes a bound on the amount of information an adversary gains based on their intercepted state. Given a message $m \in \mathcal{M}$, Alice prepares a quantum state $\hat{\rho}^{(m)}_{A^n}$ and transmits it through n uses of a quantum channel $\mathcal{N}_{A\to BW}$. Willie recovers $\hat{\rho}^{(m)}_{W^n} = {\rm tr}_{B^n}(\mathcal{N}^{\otimes n}_{A\to BW}(\hat{\rho}^{(m)}_{A^n}))$. We call a system *secret* if there exists a constant state $\hat{\hat{\rho}}_{W^n}$ that does not depend on Alice's original message, such that the leakage distance is $\delta_{\rm S}$ small: $\max_{m\in\mathcal{M}} \|\hat{\rho}^{(m)}_{W^n} - \hat{\hat{\rho}}_{W^n}\|_1 \leq \delta_{\rm S}$.

D. Decoding Reliability

For transmission of classical information, we must ensure that Bob is able to decode reliably. A coding scheme $\mathcal{C}$ is reliable if, for any $\epsilon_c > 0$, $\bar{P}_e(\mathcal{C}) \leq \epsilon_c$, where $\bar{P}_e(\mathcal{C})$ denotes the average probability of error over the message and key. We call entanglement generation scheme $\mathcal{G}$ reliable if, for any $\epsilon_g > 0$, $F\left(\hat{\Phi}_{RM}, \hat{\tau}^{(\mathcal{G})}_{R\hat{M}}\right) \geq 1 - \epsilon_g$, where $\hat{\Phi}_{RM}$ is the initial maximally-entangled state at Alice and $\hat{\tau}^{(\mathcal{G})}_{R\hat{M}}$ is the entangled state Bob recovers.

IV. Fundamental Limits on Covert Secrecy and Entanglement Generation

Recall that traditional secrecy rates are defined in bits per channel use, expressed as $R = \frac{\log|\mathcal{M}|}{n}$ for a message set $\mathcal{M}$ and n channel uses. In the covert setting, however, $\log|\mathcal{M}| = O(\sqrt{n})$ yielding a zero-capacity result as $n \to \infty$. Hence, we define a covert secrecy rate as $R_{\rm sec} = \frac{\log|\mathcal{M}|}{\sqrt{n\delta_{\rm C}}}$.

Definition IV.1. (Achievable covert secrecy rate). A covert secret rate is *achievable*, if for large enough uses of the channel n, and for every $\epsilon \in (0,1), \delta_{\rm S} > 0, \delta_{\rm C} > 0$, there exists a $(e^{R_{\rm sec}\sqrt{n\delta_{\rm C}}}, n, \epsilon, \delta_{\rm S}, \delta_{\rm C})$ code for secret and covert classical communication.

Definition IV.2. (Covert secrecy capacity). The covert secrecy capacity $L_{\rm sec}$ is the supremum over all achievable covert secret rates.

Theorem 1. The covert secrecy capacity over a lossy thermal-noise bosonic channel $\mathcal{E}^{(\eta,\bar{n}_{\rm B})}$ is $L_{\rm sec}\left(\mathcal{E}^{(\eta,\bar{n}_{\rm B})}\right) = c_{\rm cov}c_{\rm rel}$ where

$$c_{\rm cov} = \frac{\sqrt{2\eta\bar{n}_{\rm B}(1+\eta\bar{n}_{\rm B})}}{1-\eta} \tag{1}$$

$$c_{\rm rel} = \eta\log\left(1 + \frac{1}{(1-\eta)\bar{n}_{\rm B}}\right). \tag{2}$$

Proof: **Achievability: Construction:** Alice employs position based coding [39] by generating public shared randomness $(X^n)^{\otimes|\mathcal{M}||\mathcal{K}|}$ and distributes copies to Bob and Willie, where X is a uniform random variable over $\{1,\ldots,4\}$ and X^n is an independent and identically-distributed (i.i.d.) random vector. Alice and Bob pre-share a key $k \in \mathcal{K}$ that is kept secret from Willie. Based on message m, Alice subselects codeword $X^n(m,k)$ indexed by (m,k) and encodes it as a product of quadrature phase-shift keyed (QPSK) coherent states: $|\phi(X^n(m,k))\rangle_{A^n} = |\sqrt{\bar{n}_s}e^{j\pi X_1/2}\rangle_{A_1} \otimes \cdots \otimes |\sqrt{\bar{n}_s}e^{j\pi X_n/2}\rangle_{A_n}$ where $\bar{n}_s$ is the mean photon number per mode and $\pi X_n/2$ denotes the phase. She transmits the state over n lossy thermal-noise bosonic channel uses. Bob obtains output systems B^n and employs the sequential decoding POVM $\{\hat{\Omega}^{(m,k)}_{B^n}\}$ [39] based on his knowledge of k. The following lemma yields that this construction is reliable, covert and secret on average for appropriate choice of $(\mathcal{M},\mathcal{K})$:

Lemma 1. Consider the bosonic channel $\mathcal{E}_{A\to BW}^{\eta,\bar{n}_{\mathrm{B}}} = \mathrm{tr}_E\left[\mathcal{V}_{A\to BWE}^{\eta,\bar{n}_b}\right]$ defined in terms of its Stinespring dilation and a partial trace over ancillary system E. Let $\hat{\rho}_{WE}^{\otimes n} = \mathrm{tr}_{B^n}\left[\mathcal{V}_{A\to BWE}^{\eta,\bar{n}_{\mathrm{B}}\otimes n}((\hat{\rho}_A)^{\otimes n})\right]$ and $\hat{\rho}_W^{\otimes n} = \mathrm{tr}_{E^n}\left[\hat{\rho}_{WE}^{\otimes n}\right]$ with $\hat{\rho}_A = \mathrm{Tr}_X[\hat{\rho}_{XA}]$. There exists a random coding scheme defined on QPSK coherent-state codeword input states with mean photon number per mode $\bar{n}_s$ and $\varsigma_n^{(2)} \in o(1)\cap\omega(1/\sqrt{n})$ such that, for $\bar{n}_s \in o(1)$ and for n large enough, such that

$$\log|\mathcal{M}| \geq (1-\varsigma_n)c_{\mathrm{rel}}\bar{n}_s n \tag{3}$$

$$\log|\mathcal{K}| \leq (1+\varsigma_n)\left(g(\eta\bar{n}_b)+g(\bar{n}_b)-2g(\nu-1/2)\right)n \tag{4}$$

and

$$E_{\mathcal{C}}\left[\bar{P}_e(\mathcal{C})\right] \leq e^{-\varsigma_n^{(1)}\sqrt{n}} \tag{5}$$

$$E_{\mathcal{C}}\left[\left|D(\hat{\rho}_{W^n}\|\hat{\rho}_W^{\otimes n}) - D(\hat{\rho}_W^{\otimes n}\|\hat{\rho}_{W^n}^{(0)})\right|\right] \leq e^{-\varsigma_n^{(2)}\sqrt{n}} \tag{6}$$

$$\max_m E_{\mathcal{C}}\left[\left\|\hat{\rho}_{W^nE^n}^m - \hat{\rho}_{WE}^{\otimes n}\right\|_1\right] \leq e^{-\varsigma_n^{(3)}\sqrt{n}} \tag{7}$$

where $\varsigma_n, \varsigma_n^{(1)}, \varsigma_n^{(3)} \in o(1)\cap\omega(1/\sqrt{n})$, $g(x) \triangleq (1+x)\log(1+x) - x\log(x)$ and $\nu = \frac{1}{2}\sqrt{1+2\bar{n}_b(1-\eta)}$.

To prove Lemma 1, we adapt the position-based coding and sequential decoding strategy from [19, Lem. V.1], [39] to ensure secrecy at the cost of a larger key requirement (Lemma 7 in Appendix IV). Applying it to QPSK coherent-state codewords transmitted over n uses of lossy thermal noise bosonic channel yields a random coding scheme that is reliable, covert, and secret on average (Lemma 1). The full proof of Lemma 1 is in Appendix I.

Next, we show that a deterministic coding scheme exists that satisfies the average decoding reliability, covertness, and message average secrecy requirements. We then use the standard expurgation argument to construct bounds on the maximum probability of decoding error and achieve semantic secrecy.

Lemma 2. [18, Lem. 10] There exists a sequence of deterministic coding schemes $\mathcal{C} = \{x^n(m,k)\}$ such that, for n large enough,

$$\log|\mathcal{M}| = (1-\varsigma_n)c_{\mathrm{rel}}\bar{n}_s n \tag{8}$$

$$\log|\mathcal{K}| = (1+\varsigma_n)\left(g(\eta\bar{n}_b)+g(\bar{n}_b)-2g(\nu-1/2)\right)n \tag{9}$$

where

$$\max_{m,k} P_e^{(m,k)}(\mathcal{C}) \leq e^{-\zeta_n^{(1)}\sqrt{n}} \tag{10}$$

$$\left|D(\hat{\rho}_{W^n}\|\hat{\rho}_W^{\otimes n}) - D(\hat{\rho}_W^{\otimes n}\|\hat{\rho}_{W^n}^{(0)})\right| \leq e^{-\zeta_n^{(2)}\sqrt{n}} \tag{11}$$

$$\max_m \left\|\hat{\rho}_{W^nE^n}^m - \hat{\rho}_{WE}^{\otimes n}\right\|_1 \leq e^{-\zeta_n^{(3)}\sqrt{n}} \tag{12}$$

The proof of Lemma 2 is in Appendix II.

Now, we show that the message rate achieved for deterministic coding scheme in Lemma 2 converges to the capacity while maintaining covertness: $\log|\mathcal{M}| \geq (1-\varsigma_n)\eta\log\left(1+\frac{1}{(1-\eta)\bar{n}_b}\right)\bar{n}_S n$. Note that (10) implies decoding reliability, (12) implies secrecy, and (11) implies $D\left(\hat{\bar{\rho}}_{W^n}\|\hat{\rho}_{W^n}^{(0)}\right) \leq nD\left(\hat{\bar{\rho}}_W\|\hat{\rho}_W^{(0)}\right) + e^{-\varsigma_n^{(2)}\sqrt{n}}$, where $\hat{\bar{\rho}}_W$ is Willie's output from single-mode QPSK coherent state constellation input with mean photon number per mode $\bar{n}_s$ and $\varsigma_n^{(2)} \in \omega(1/\sqrt{n})$. Thus, choosing $\bar{n}_s = c_{\mathrm{cov}}\sqrt{\frac{\delta_{\mathrm{c}}}{n}}$ and [13, Th. 2] implies the scheme is covert in the limit $n\to\infty$. Now, using Definition IV.2, we have

$$L_{\mathrm{sec}}\left(\mathcal{E}^{(\eta,\bar{n}_{\mathrm{B}})}\right) \geq \lim_{n\to\infty}\frac{\log|\mathcal{M}|}{\sqrt{\delta_{\mathrm{C}}n}} = c_{\mathrm{cov}}c_{\mathrm{rel}}. \tag{13}$$

Converse: The converse is given by the non-secret covert converse argument provided in [11, Th. 1]. ∎

Remark 1. The key size requirements in (4) and (9) are $\mathcal{O}(n)$ instead of $\mathcal{O}(\sqrt{n})$, as in [18, Th. 2]. These can be strengthened back to $\mathcal{O}(\sqrt{n})$ if we relax the secrecy requirement to not include the ancillary systems E^n. However, in the achievability proof for entanglement generation (Theorem 2) that follows, we use the the larger key in (12). We believe that using the sparse coding (see Section V) may reduce the key size scaling, but leave the analysis to future work.

Similar to that of the covert secrecy rate, entanglement generation rate is defined as $R = \frac{\log(\dim(\mathcal{H}_{\mathcal{M}}))}{n}$ in qubit pairs generated per channel use, where $\dim(\mathcal{H}_{\mathcal{M}})$ is the dimension of entangled state. However, the square root law also governs entanglement generation, and we define the covert entanglement generation rate as $R_{\mathrm{EG}} = \frac{\log(\dim(\mathcal{H}_{\mathcal{M}}))}{\sqrt{n\delta_{\mathrm{C}}}}$.

Definition IV.3. (Achievable covert entanglement-generation rate). A covert entanglement-generation rate is *achievable*, if for large enough uses of the channel n, and for every $\epsilon \in (0,1), \delta_{\mathrm{C}} > 0$, there exists a $(e^{R_{\mathrm{EG}}\sqrt{n\delta_{\mathrm{C}}}}, n, \epsilon, \delta_{\mathrm{C}})$ code for covert entanglement-generation.

Definition IV.4. (Covert entanglement-generation capacity). The covert entanglement generation capacity L_{EG} is the supremum over all *achievable covert* entanglement generation rates.

Theorem 2. Covert entanglement-generation capacity of a lossy thermal-noise bosonic channel $\mathcal{E}^{(\eta,\bar{n}_{\mathrm{B}})}$ is $L_{\mathrm{EG}}\left(\mathcal{E}^{(\eta,\bar{n}_{\mathrm{B}})}\right) = c_{\mathrm{cov}}c_{\mathrm{rel}}$, where c_{rel} and c_{cov} are defined in (2) and (1), respectively.

Proof: **Achievability:** The achievability proof adapts [18, Thm. 2] by using the classical code construction in the proof of Theorem 1, and converting it to a quantum one.

Code conversion: As in [18, Thm. 2], we convert the classical code from the proof of Theorem 1 into an entanglement generation code. Consider the QPSK codebook $\{x^n(m,k)\}_{m,k}$ used for the result in Lemma 2, where single-mode input state at Alice is a coherent state $|\sqrt{\bar{n}_s}e^{j\pi x/2}\rangle$ for $x \in 1,\ldots,4$. Therefore, a single classical codeword given (m,k) is a product of coherent states: $|x_{\mathrm{coh}}^n(m,k)\rangle_{A^n} \triangleq |\sqrt{\bar{n}_s}e^{j\pi x_1/2}\rangle_{A_1}\otimes\cdots\otimes|\sqrt{\bar{n}_s}e^{j\pi x_n/2}\rangle_{A_n}$. Alice converts this to a quantum codebook $\{|\phi_m\rangle_{A^n} : m\in\mathcal{M}\}$ with $|\phi_m\rangle_{A^n} = \frac{1}{\sqrt{|\mathcal{K}|}}\sum_k e^{jf(m,k)}|x_{\mathrm{coh}}^n(m,k)\rangle_{A^n}$, where $f(m,k)$ is defined later. Alice prepares her encoding by generating a maximally entangled state: $|\Phi\rangle_{RM} = \frac{1}{\sqrt{|\mathcal{M}|}}\sum_m |m\rangle_R\otimes|m\rangle_M$, where subsystems R and M are the resource and message, respectively. She generates a copy of the M subsystem using a CNOT gate as in [18, Eq. (108)] to obtain $|\tau\rangle_{RMM'}$. Alice applies an isometry $\hat{U}_{M'\to A^n}$ that takes $|m\rangle_{M'} \to |\phi_m\rangle_{A^n}$ as

$$|\tau\rangle_{RMA^n} = \left(\hat{I}\otimes\hat{I}\otimes\hat{U}_{M'\to A^n}\right)|\tau\rangle_{RMM'} \tag{14}$$

$$= \frac{1}{\sqrt{|\mathcal{M}|}} \sum_m |m\rangle_R \otimes |m\rangle_M \otimes |\phi_m\rangle_{A^n} . \quad (15)$$

She transmits A^n systems over n copies of $V_{A\to BWE}^{\eta,\bar{n}_b}$. We represent the global state as $|\tau\rangle_{RMB^nW^nE^n} = \frac{1}{\sqrt{|\mathcal{M}|}} \sum_m |m\rangle_R \otimes |m\rangle_M \otimes |\phi_m\rangle_{B^nW^nE^n}$ with $|\phi_m\rangle_{B^nW^nE^n} = V_{A\to BWE}^{\eta,\bar{n}_b\otimes n}(|\phi_m\rangle_{A^n})$ being the channel output given input $|\phi_m\rangle_{A^n}$ and $V_{A\to BWE}^{\eta,\bar{n}_B}$ being the Stinespring representation of the bosonic channel $\mathcal{E}_{A\to BW}^{\eta,\bar{n}_b}$. Recall that Lemma 2 implies that there is a decoding POVM $\{\hat{\Omega}_{B^n}^{(m,k)}\}$ such that given any value of pre-shared key k the classical code achieves $\mathrm{tr}\left[\hat{\Omega}_{B^n}^{(m,k)}\hat{\rho}_{B^n}^{(m,k)}\right] \geq 1 - e^{-\varsigma_n^{(1)}\sqrt{n}}$, for all m, k. We construct a coherent version of this POVM [16, Sec. 5.4] as $\mathcal{D}_{B^n\to B^n\hat{M}\hat{K}} = \sum_{m,k} \sqrt{\hat{\Omega}_{B^n}^{(m,k)}} \otimes |m\rangle_{\hat{M}} \otimes |k\rangle_{\hat{K}}$, which, after its use, yields the global state:

$$|\tau\rangle_{RMB^nW^nE^n\hat{M}\hat{K}} = \left(\hat{I}_{RM} \otimes \mathcal{D}_{B^n\to B^n\hat{M}\hat{K}} \otimes \hat{I}_{W^nE^n}\right) |\tau\rangle_{RMB^nW^nE^n} \quad (16)$$

We now show that this conversion yields an entanglement generation scheme that is reliable and covert.

Lemma 3. Consider covert entanglement generation via lossy thermal noise bosonic channel $\mathcal{E}_{A\to BW}^{\eta,\bar{n}_b}$ with Stinespring dilation $V_{A\to BWE}^{\eta,\bar{n}_b}$. For any $\varsigma_n \in o(1) \cap \omega(1/\sqrt{n})$ there exists $\varsigma_n^{(1)}, \varsigma_n^{(4)} \in o(1) \cap \omega(1/\sqrt{n})$ such that for n sufficiently large

$$\log d_M \geq (1-\varsigma_n) c_{\mathrm{rel}} \bar{n}_s n \quad (17)$$

$$\log |\mathcal{K}| \leq (1+\varsigma_n)\left(g(\eta\bar{n}_b) + g(\bar{n}_b) - 2g(\nu - 1/2)\right) n \quad (18)$$

while

$$F(\hat{\Phi}_{RM}, \hat{\tau}_{R\hat{M}}) \geq 1 - e^{-\varsigma_n^{(1)}\sqrt{n}} \quad (19)$$

$$\left| D\left(\hat{\tilde{\rho}}_{W^n} \middle\| \hat{\rho}_{W^n}^{(0)}\right) - D\left(\hat{\tilde{\rho}}_{W^n}^{\otimes n} \middle\| \hat{\rho}_{W^n}^{(0)}\right) \right| \leq e^{-\varsigma_n^{(4)}\sqrt{n}} \quad (20)$$

where $d_M = \dim(\mathcal{H}_M)$, $\hat{\Phi}_{RM}$ is the maximally entangled state, $\hat{\tau}_{R\hat{M}}$ is Bob's decoded state and $\hat{\tilde{\rho}}_{W^n}$ is Willie's received state from the covert entanglement generation scheme.

The proof of Lemma 3 is adapted from [18] and key steps are provided in Appendix III. The challenge in the present model is to ensure that state approximation holds for non-orthogonal coherent state codewords and to carefully decouple Bob's state from both Willie's subsystems W^n and the ancillary subsystems E^n. Further, showing that covertness is maintained requires use of continuity of entropy results for bosonic systems [40].

We now show that this scheme with rate given by (17) achieves the capacity in the limit $n \to \infty$ while remaining covert. Note that (19) ensures reliability and (20) implies $D\left(\hat{\tilde{\rho}}_{W^n} \| \hat{\rho}_{W^n}^{(0)}\right) \leq nD\left(\hat{\tilde{\rho}}_W \| \hat{\rho}_W^{(0)}\right) + e^{-\varsigma_n^{(4)}\sqrt{n}}$, where $\varsigma_n^{(4)} \in \omega(1/\sqrt{n})$. Thus, choosing $\bar{n}_s = c_{\mathrm{cov}}\sqrt{\frac{\delta}{n}}$ implies the scheme is covert in the limit $n \to \infty$. Now, using Definition IV.2, we have $L_{\mathrm{EG}}\left(\mathcal{E}^{(\eta,\bar{n}_B)}\right) \geq \lim_{n\to\infty} \frac{\log(\dim(\mathcal{H}_{\mathcal{M}}))}{\sqrt{\delta_C n}} = c_{\mathrm{cov}} c_{\mathrm{rel}}$.

Converse: The converse is given by the standard classical covert communication converse argument in [11, Th. 1]. ■

V. Towards Practical Covert Entanglement Generation

While the covert entanglement-generation capacity of the lossy thermal-noise bosonic channel is achievable per Theorem 2, it is unclear how to construct Alice's state in (15) physically. Hence, we investigate entanglement generation using single- and dual-rail photonic qubit encodings.

We require some additional definitions: denote by $[a]^+ = \max(a, 0)$, and the Shannon entropy associated with probability vector $\vec{p}$ by $H(\vec{p}) = -\sum_{p_i\in\vec{p}} p_i \log_2(p_i)$. A single-rail photonic qubit is encoded in a single mode of a photon represented by $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. The dual-rail photonic qubit qubit uses a single photon across two modes: $|\psi\rangle = \alpha|01\rangle + \beta|10\rangle$. We additionally use the χ^2-divergence between two states $\hat{\rho}$ and $\hat{\sigma}$ given by $D_{\chi^2}(\hat{\rho}\|\hat{\sigma}) = \mathrm{tr}[\hat{\rho}^2\hat{\sigma}^{-1}] - 1$.

Lemma 4. Using a single-rail photonic qubit, for n large enough and arbitrary $\vartheta > 0$, $M(n) \geq (1-\vartheta)\sqrt{n}\sqrt{2}c_{\mathrm{cov}}R\sqrt{\delta_C}$ qubit pairs can be generated reliably and covertly over n uses of the lossy thermal-noise bosonic channel acting independently on each channel use, where c_{cov} is in (1), δ_C is the QRE-covertness constraint, and $R \geq [1 - H(\vec{p})]^+$ is the constant achievable rate of reliable qubit pair generation per channel use with $\vec{p} = (p_I, p_X, p_Y, p_Z)$, $p_I = (1 - \frac{3}{4}p_f)q_I$, $p_j = (1 - \frac{3}{4}p_f)q_j + \frac{1}{4}p_f$ for $j = X, Y, Z$ and

$$p_f = \frac{1 + (1-\eta)\bar{n}_B(3 + 2\bar{n}_B - 2\eta(\bar{n}_B + \frac{1}{2}))}{(1 + (1-\eta)\bar{n}_B)^3} \quad (21)$$

$$q_I = \frac{1}{2N(G,\tau)}\left(\frac{2G + \tau - 1}{G^2} + 2\frac{\sqrt{\tau}}{G^{\frac{3}{2}}}\right) \quad (22)$$

$$q_X = \frac{1}{2N(G,\tau)}\left(\frac{G-1}{G^2} + \frac{1-\tau}{G}\right) \quad (23)$$

$$q_Y = \frac{1}{2N(G,\tau)}\left(\frac{G-1}{G^2} + \frac{1-\tau}{G}\right) \quad (24)$$

$$q_Z = \frac{1}{2N(G,\tau)}\left(\frac{2G + \tau - 1}{G^2}G^2 - 2\frac{\sqrt{\tau}}{G^{\frac{3}{2}}}\right) \quad (25)$$

where $G = 1 + (1-\eta)\bar{n}_B$, $\tau = \eta/G$ and $N(G,\tau) = \frac{3G + 2G^{3/2}\sqrt{\tau} + \tau - 2}{2G^2}$.

Proof: **Construction and reliability:** The construction is depicted in Fig. 3. Alice prepares a Bell state $|\Phi\rangle_{RA}$. System R is kept as reference, system A is prepared to be sent through the channel. To ensure covertness, Alice employs the sparse signaling approach from [41]: let $\mathbf{x} \equiv \{x_i, i = 1, \ldots, n\}$ be a binary sequence indicating the selected channel uses for non-innocent transmission, with $x_i = 1$ corresponding to a non-innocent qubit input from Alice on the i^{th} channel use, and $x_i = 0$ to innocent input. Thus, $w(\mathbf{x}) \equiv \sum_{i=1}^n x_i$ is the number of non-innocent inputs for a given $\mathbf{x}$. For an arbitrary $\vartheta > 0$, define $\mathcal{A} \equiv \left\{\mathbf{x} : \left|q - \frac{1}{n}w(\mathbf{x})\right| \leq \vartheta\right\}$ as the set containing length-n binary sequences whose normalized weight is close to $q \in (0,1)$. Let $p_X(x) = \{q \text{ if } x = 1; 1-q \text{ if } x = 0\}$. Denote $p(\mathcal{A}) = \sum_{\mathbf{x}\in\mathcal{A}} \prod_{i=1}^n p_X(x_i)$, and: $p_{\mathbf{X}}(\mathbf{x}) \equiv \left\{\frac{\prod_{i=1}^n p_X(x_i)}{p(\mathcal{A})} \text{ if } \mathbf{x} \in \mathcal{A}; 0 \text{ if } \mathbf{x} \notin \mathcal{A}\right\}$. Alice and Bob choose the channel uses for transmitting qubits by randomly sampling $\mathbf{x} \in \mathcal{A}$ using $p_{\mathbf{X}}$. Their choice $\mathbf{x}$ comprises part of the classical pre-shared secret in Fig. 3.

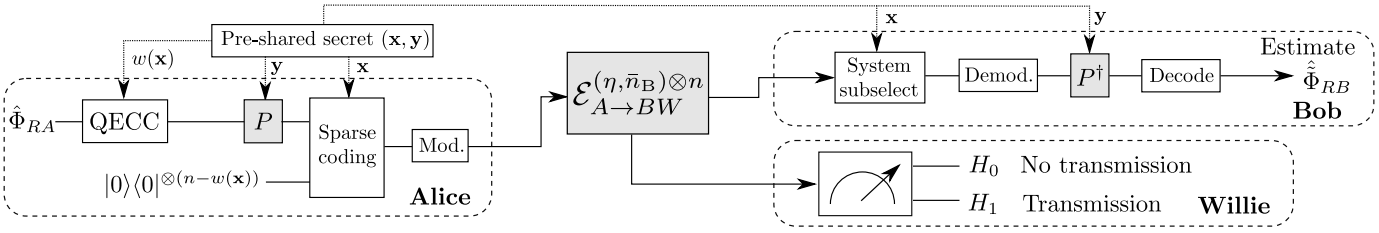


Fig. 3. Construction of achievable covert quantum entanglement generation over the lossy thermal-noise channel $\mathcal{E}^{\eta,\bar{n}_{\mathrm{B}}}_{A\to BW}$. Alice first prepares a Bell state $\hat{\Phi}_{RA}$ entangled with a reference system. She then sends her half of the qubit through the following process: given a pre-shared secret $\mathbf{x}$, she applies a QECC corresponding to the number $w(\mathbf{x})$ of non-innocent output states. This is followed by the application of Pauli gates defined by pre-shared secret sequence $\mathbf{y}$. She then applies sparse coding that "spreads" these $w(\mathbf{x})$ non-innocent states across n channel uses by inserting innocent input states according to locations in $\mathbf{x}$. Alice then transmits the modulated resulting state $\hat{\rho}^{(\psi)}_{A^n}$. Bob sub-selects the systems containing the non-innocent states using $\mathbf{x}$ and demodulates by projecting them onto the qubit basis. He then inverts the Pauli twirling operation using $\mathbf{y}$ and decodes to obtain a state $\hat{\hat{\Phi}}_{RB}$ entangled with the reference system. Willie performs an optimal hypothesis test on whether transmission occurred.

Alice and Bob further generate a set $\{c_i\}_{i\in w(\mathcal{A})}$ of $|w(\mathcal{A})| = \lceil 2\vartheta qn\rceil$ quantum error correction codes (QECCs) following standard techniques [16, Sec. 24.6.3], where $w(\mathcal{A}) \equiv \{w(\mathbf{x}) : \mathbf{x} \in \mathcal{A}\}$. For a given number of non-innocent states available $w(\mathbf{x})$, the QECC $c_{w(\mathbf{x})}$ yields an isometry $\hat{U}_{c_{w(\mathbf{x})}}$ that maps the input state $\hat{U}^{c_{w(\mathbf{x})}}|\psi\rangle_{A^{w(\mathbf{x})}} = |c_{w(\mathbf{x})}(\psi)\rangle_{A^{w(\mathbf{x})}}$. We extend these QECCs to a *covert QECC* c, which maps $\hat{U}^{c(\mathbf{x})} : (\mathbf{x},|\psi\rangle) \to |c(\psi,\mathbf{x})\rangle_{A^n}$. Let the isometry $\hat{U}^{c(\mathbf{x})} = \left(\hat{I}_{A_{x^c}} \otimes \hat{U}^{c_{w(\mathbf{x})}}\right)$ with $|c_{w(\mathbf{x})}(\psi)\rangle_{A^{w(\mathbf{x})}} = \mathrm{tr}_{\mathbf{x}^c}\left(|c(\psi,\mathbf{x})\rangle_{A^n}\right)$ being the QECC mapping for $w(\mathbf{x})$, and $\mathrm{tr}_{\mathbf{x}^c}$ the partial trace over the innocent inputs defined by choice of $\mathbf{x}$. Given $\mathbf{x}$, the covert QECC can be thought of as the sparse encoding of the corresponding non-covert QECC of block length $w(\mathbf{x})$, with innocent states injected in each of the systems $\{A_i : x_i = 0\}$. Hence, the systems occupied by innocent states do not contribute to the code's error-correcting capabilities but are utilized for covertness.

Alice prepares $w(\mathbf{x})$ Bell state copies $|\Phi\rangle^{\otimes w(\mathbf{x})}_{RA}$ and applies the isometry $\hat{U}_{c_{w(\mathbf{x})}}$ on the $A^{w(\mathbf{x})}$ subsystems. Alice and Bob also secretly choose a sequence $\mathbf{y}$ indicating $w(\mathbf{x})$ Pauli gates sampled uniformly at random from $\mathcal{P} \equiv \{\hat{I},\hat{X},\hat{Y},\hat{Z}\}$. The sequence of Pauli gates chosen to spread over n modes is $\hat{P}^{(\mathbf{x},\mathbf{y})}$, where $\hat{I}$ is applied to a mode occupied by an innocent state. Alice applies the Pauli gates as the first stage of the Pauli twirling operation [42]. Thus, Alice transmits the A^n subsystems of entangled encoded state $|\Phi_c\rangle_{R^{w(\mathbf{x})}A^n} = (\hat{I}^{\otimes w(\mathbf{x})}_R \otimes \hat{P}^{(\mathbf{x},\mathbf{y})}\hat{U}^{c(\mathbf{x})})\left(|\mathbf{0}\rangle_{A_{x^c}}|\Phi\rangle^{\otimes w(\mathbf{x})}_{AR}\right)$ given $\mathbf{x},\mathbf{y}$, over the channel $\mathcal{E}^{(\eta,\bar{n}_{\mathrm{B}})}_{A\to BW}$. Bob receives state $\hat{\rho}^{(\psi)}_{B^n}$ and uses $\mathbf{x}$ to subselect the state $\hat{\rho}^{(\psi)}_{B^{w(x)}} = \mathrm{tr}_{\mathbf{x}^c}(\hat{\rho}^{(\psi)}_{B^n})$, which represents the state occupied by output systems of the non-covert QECC.

Denote by $\hat{\rho}^{(\psi)}_{B_i}$ the state of the i^{th} subsystem in $\hat{\rho}^{(\psi)}_{B^{w(x)}}$. $\hat{\rho}^{(\psi)}_{B_i}$ is an arbitrary state equivalent to that in (86) with η replaced by $1-\eta$. The state of each subsystem is demodulated by projecting it onto the qubit basis via application of $\hat{\Pi} = |0\rangle\langle 0| + |1\rangle\langle 1|$. Projection is a probabilistic process with the probability of failure $p_{\mathrm{f}} = \frac{1+(1-\eta)\bar{n}_{\mathrm{B}}(3+2\bar{n}_{\mathrm{B}}-2\eta(\bar{n}_{\mathrm{B}}+1/2))}{(1+(1-\eta)\bar{n}_{\mathrm{B}})^3}$. When Bob fails to project, he replaces the state with the maximally mixed state $\hat{\pi} = \frac{\hat{I}}{2}$. This process yields a state in the qubit basis of $(1-p_{\mathrm{f}})\hat{\Pi}\hat{\rho}^{B_i}_{\psi}\hat{\Pi}^\dagger + p_{\mathrm{f}}\hat{\pi} = (1-p_{\mathrm{f}})\hat{\tau}^{(\psi)}_{B_i} + p_{\mathrm{f}}\hat{\pi}$ where $\hat{\tau}^{(\psi)}_{B_i}$ is the projected state in the i^{th} system. This represents a depolarizing channel for $\hat{\tau}^{(\psi)}_{B_i}$. Bob then uses pre-shared $\mathbf{y}$ to apply the appropriate sequence of Pauli gates, completing the Pauli twirling process. Pauli twirling by Alice and Bob on their qubit states guarantees a Pauli noise channel $\vec{q}_{\mathrm{tw}} = (q_I, q_X, q_Y, q_Z)$ with channel parameters defined in (22)-(25) and derived in Appendix VIII.

Recall a Pauli channel, $\mathcal{E}^{\vec{p}}_{\mathcal{P}}$, maps input state $\hat{\rho}$ as follows: $\mathcal{E}^{\vec{p}}_{\mathcal{P}}(\hat{\rho}) = p_I\hat{I}\hat{\rho}\hat{I} + p_X\hat{X}\hat{\rho}\hat{X} + p_Y\hat{Y}\hat{\rho}\hat{Y} + p_Z\hat{Z}\hat{\rho}\hat{Z}$ for $\vec{p} = (p_I, p_X, p_Y, p_Z)$. A depolarizing channel is given by $\mathcal{E}^{\lambda}_{\mathrm{dep}}(\hat{\rho}) = \mathcal{E}^{\vec{p}_{\mathrm{dep}}(\lambda)}_{\mathcal{P}}(\hat{\rho})$ with depolarizing parameter λ and $\vec{p}_{\mathrm{dep}}(\lambda) = \left(1-\frac{3\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}\right)$. Hence the combination of the depolarizing channel from projection to the qubit basis and Pauli channel generated by twirling yields a combined Pauli channel, $\mathcal{E}^{p_{\mathrm{f}}}_{\mathrm{dep}}\left(\mathcal{E}^{\vec{q}_{\mathrm{tw}}}_{\mathcal{P}}\left(\hat{\tau}^{(\psi)}_{B_i}\right)\right) = \mathcal{E}^{\vec{p}}_{\mathcal{P}}\left(\hat{\tau}^{(\psi)}_{B_i}\right)$ where $\vec{p} = (p_I, p_X, p_Y, p_Z)$ with $p_I = (1-\frac{3}{4}p_{\mathrm{f}})q_I$ and $p_j = (1-\frac{3}{4}p_{\mathrm{f}})q_j + \frac{1}{4}p_{\mathrm{f}}$ for $j = X,Y,Z$. For every $w(\mathbf{x}) \in w(\mathcal{A})$, the hashing bound guarantees the existence of a QECC with an achievable rate of $R = 1 - H(\vec{p})$ [16, Sec. 24.6.3].

Covertness analysis: Willie knows the construction procedure, channel parameters, q, the covert QECC, and the time of transmission. The sequence of random Pauli operations Alice applies to her encoded state is unknown to Willie. Therefore, from Willie's perspective, Alice's average non-innocent input state given $\mathbf{x}$, and tracing out the R systems is $\hat{\rho}_{A^{w(\mathbf{x})}} = \mathcal{E}^{1\otimes w(\mathbf{x})}_{\mathrm{dep}}\left(|c(\psi)\rangle\langle c(\psi)|_{A^{w(\mathbf{x})}}\right) = (\hat{\pi})^{\otimes w(\mathbf{x})}$, where $\mathcal{E}^{1\otimes w(\mathbf{x})}_{\mathrm{dep}}(\cdot)$ is the completely depolarizing channel with $\lambda = 1$ over $w(\mathbf{x})$ non-innocent input states. Thus, Willie observes $\hat{\rho}^{(\psi)}_{W^n} = \hat{\rho}^{(\hat{\pi})}_{W^n} \equiv \sum_{\mathbf{x}\in\mathcal{A}} p_{\mathbf{X}}(\mathbf{x}) \bigotimes_{i=1}^n \hat{\rho}^{\hat{\pi},x_i}_{W_i}$ where $\hat{\rho}^{(\hat{\pi},x_i)}_{W_i} = \{\hat{\rho}^{(\hat{\pi})}_W \text{ if } x_i = 1; \hat{\rho}^{(0)}_W \text{ if } x_i = 0\}$, with $\hat{\rho}^{(\hat{\pi})}_W = \mathcal{E}^{\eta,\bar{n}_{\mathrm{B}}}_{A\to W}(\hat{\pi})$ the non-innocent output state.

We upper-bound the QRE between Willie's output $\hat{\rho}^{W^n}_{\hat{\pi}}$ and the innocent state $\hat{\rho}^{(0)}_{W^n}$ as follows:

$$D\left(\hat{\rho}^{(\hat{\pi})}_{W^n}\middle\|\left(\hat{\rho}^{(0)}_{W^n}\right)\right) = D\left(\left(\hat{\hat{\rho}}^{(\hat{\pi})}_W\right)^{\otimes n}\middle\|\left(\hat{\rho}^{(0)}_{W^n}\right)\right) + o(1) \quad (26)$$

$$= nD\left(\hat{\hat{\rho}}^{(\hat{\pi})}_W\middle\|\hat{\rho}^{(0)}_W\right) + o(1) \quad (27)$$

$$\leq q^2 nD_{\chi^2}\left(\hat{\rho}^{(\hat{\pi})}_W\middle\|\hat{\rho}^{(0)}_W\right), \quad (28)$$

where $\hat{\hat{\rho}}^{(\hat{\pi})}_W = (1-q)\hat{\rho}^{(0)}_W + q\hat{\rho}^{(\hat{\pi})}_W$, (26) is due to the adaptation provided in Appendix VII of the covertness analysis for sparse

signaling approach from [41], (27) is from the additivity of the QRE over product states [16, Ex. 11.8.7], and (28) is by [43, Lemma 1] and [44, Eq. (9)] for large enough n. Thus, the right-hand side of the covertness requirement is bounded by (28). Thus, Alice maintains covertness by choosing $q \leq \sqrt{2}c_{\text{cov}}\sqrt{\frac{\delta_{\text{C}}}{n}}$, where c_{cov} is defined in (1) and derived in Appendix IX. ∎

Remark: The proof of this lemma has been generalized to an arbitrary channel in [2] with proper assumptions about the channel and Willie's output state. These assumptions apply to the bosonic channel and are used directly in Lemma 4 proof.

Lemma 5. Using a dual-rail photonic qubit, for n large enough and arbitrary $\vartheta > 0$, $M(n) \geq (1-\vartheta)\sqrt{n}\frac{c_{\text{cov}}}{\sqrt{2}}R\sqrt{\delta_{\text{C}}}$ e-bits can be transmitted reliably and covertly over n uses of the lossy thermal-noise bosonic channel, where c_{cov} is in (1), and δ_{C} is the covertness constraint. $R \geq [1 - H(\vec{p})]^+$ is the constant achievable rate of reliable e-bit transmission per round, where $\vec{p} = \left[1 - \frac{3p}{4}, \frac{p}{4}, \frac{p}{4}, \frac{p}{4}\right]$, $p = 1 - \frac{\eta}{(1+(1-\eta)\bar{n}_{\text{B}})^4}$.

Proof: The construction of the dual-rail setting is the same as that in Fig. 3 with the exception that Alice's QECC uses dual-rail encoding, and further, $|00\rangle\langle 00|^{\otimes(n-w(\mathbf{x})/2)}$ is used for sparse coding. The proof follows directly from the proof of Lemma 4, where the steps in Appendix IX are performed on the arbitrary dual-rail input state. The additional $\frac{1}{\sqrt{2}}$ term arises from the fact that analysis is performed on a single photon occupying two modes, and hence two channel uses. ∎

Remark: Pauli twirling is often used to approximate noise in non-Pauli channels. Indeed, we use this property to transform an arbitrary quantum channel into a Pauli channel for reliability analysis. However, Pauli twirling has an added benefit in covertness analysis, ensuring that, on average, the non-innocent input state appears as the maximally mixed state from Willie's point of view.

VI. Performance Analysis

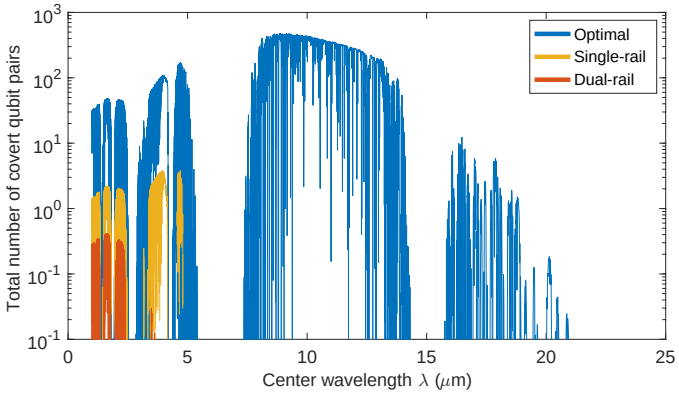


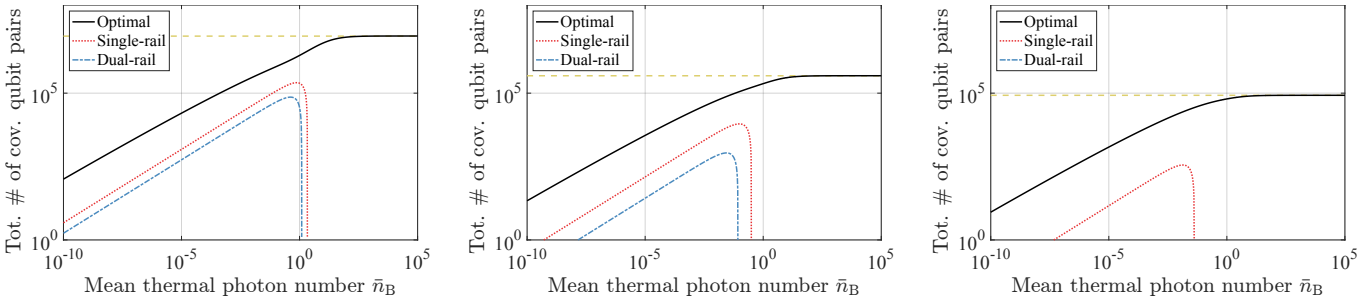
Fig. 4. The total number of covert qubits that can be reliably transmitted with a bandwidth of $W = 10$ GHz and time $T = 60$ seconds vs. transmission center wavelength over a free-space optical link between Alice and Bob described in the beginning of Section VI. The blue line is the optimal bound from Theorem 2. The yellow and red lines represent the single- and dual-rail strategies.

In Fig. 4 we plot covert entanglement-generation rates for a free-space optical (FSO) link between Alice and Bob, who have an $L = 1$ km line-of-sight channel, and aperture radii $r = 10$ cm. We let signaling bandwidth $W = 1$ GHz, thus, over transmission duration of T s, we have $n = TW = 10^9 T$ optical modes. We set $T = 60$s and $\delta = 0.05$. As in [9], we employ a detailed MODTRAN 'Mid-Latitude Summer (MLS)' atmospheric model [45] for an FSO channel at elevation 10 m above the ground level with visibility 23 km in clear weather. We calculate the mean background thermal noise photon number from the total radiance at 60° solar elevation. Unlike [9], we only consider the fundamental transverse electromagnetic (TEM00) spatial mode. This is because higher-order spatial modes do not substantially improve covert classical communication (see [9, Fig. 5]). Since loss increases with higher spatial-mode order, we anticipate the performance gains to be even more modest for covert entanglement generation. Based on the MODTRAN model, the transmittance η and thermal-noise photon number $\bar{n}_{\text{B}}$ depend on the center wavelength. The blue line plots the covert capacity given in Theorem 2 while the red and orange line corresponds to the single- and dual-rail rates. We see only a small range where the single- and dual-rail strategies generate a nonzero number of qubit pairs corresponding to high transmittance.

In Fig. 5, we plot the bounds for the total number of reliably-generated covert ebits (entangled qubit pairs) vs. mean thermal photon number $\bar{n}_{\text{B}}$ for $\delta = 0.05$, $n = 60 \times 10^9$, and various values of transmittance η. The black line is the optimal rate from Theorem 2, while the dotted red and dash-dotted blue lines are the single- and dual-rail strategies. The yellow dashed line is a guide to the eye representing the limit of the optimal rate as $\bar{n}_{\text{B}} \to \infty$. The single- and dual-rail perform poorly compared to the optimal rate, with the gap between them growing as η decreases.

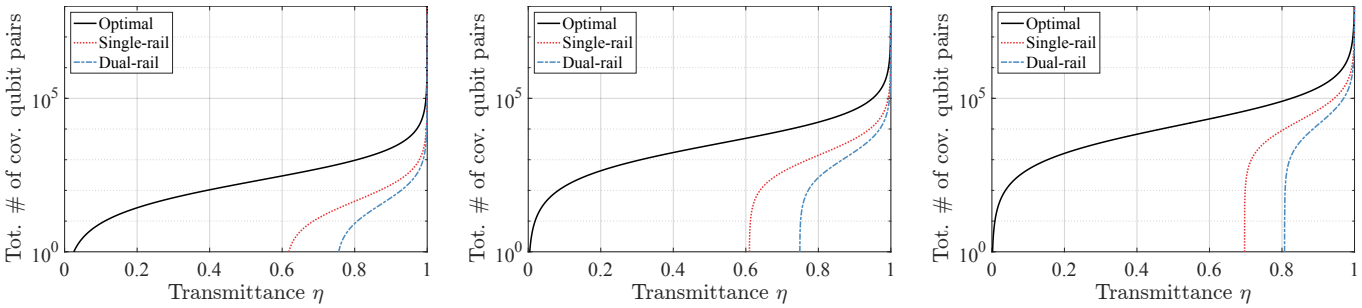
In Fig. 6, we plot the bounds for the total number of reliably-generated covert ebits vs. transmittance η for $\delta = 0.05$, $n = 60 \times 10^9$, and various values of mean thermal photon number $\bar{n}_{\text{B}}$. The black line is the optimal rate from Theorem 2, while the dotted red and dash-dotted blue lines are the single- and dual-rail strategies. For low transmittance, the single- and dual-rail rates can generate zero entangled pairs, while the optimal entanglement rate is non-zero for all values of η. Furthermore, even for high transmittance, a near-order-of-magnitude gap remains.

The results of Fig. 4, Fig. 6, and Fig. 5 demonstrate a substantial gap between the optimal rate and practical strategies. There are many potential reasons for this. The first is that, although the channel is infinite-dimensional, we do not utilize the additional degrees of freedom. It is possible that qudit or bosonic codes could increase the achievable rates. Indeed, Gottesman-Kitaev-Preskill (GKP) qubits have been shown to reach a constant factor gap to the capacity of the pure-loss channel, and perform well in the lossy thermal-noise setting as well [22]. Additionally, the strategies rely on Pauli twirling, which, by the quantum data processing [16, 10.7.2], means the mutual information between Alice and Bob can only decrease after application. Lastly, to ensure that the hashing bound can be used, when projection fails for a given state in a mode, we replace it with a maximally mixed state, discarding information about where the error occurred; this can instead



(a) $\eta = 0.95$ (b) $\eta = 0.8$ (c) $\eta = 0.65$

Fig. 5. Total number of covert entangled qubit pairs as a function of $\bar{n}_{\mathrm{B}}$ for a) $\eta = 0.95$, b) $\eta = 0.8$, and c) $\eta = 0.65$. In each subfigure, $n = 10^8$ and $\delta = 0.05$. The black line is the optimal rate from Theorem 2, while the dotted red and dash-dotted blue lines correspond to the single- and dual-rail rates. The yellow dashed line represents the convergence for the optimal rate as $\bar{n}_{\mathrm{B}} \to \infty$ in each plot.



(a) $\bar{n}_{\mathrm{B}} = 10^{-6}$ (b) $\bar{n}_{\mathrm{B}} = 10^{-3}$ (c) $\bar{n}_{\mathrm{B}} = 10^{-1}$

Fig. 6. Total number of covert entangled qubit pairs as a function of η for a) $\bar{n}_{\mathrm{B}} = 10^{-6}$, b) $\bar{n}_{\mathrm{B}} = 10^{-3}$, and c) $\bar{n}_{\mathrm{B}} = 10^{-1}$, where $n = 60 \times 10^9$ and $\delta = 0.05$ for each. The black line is the optimal rate from Theorem 2, while dotted red and dash-dotted blue lines correspond to the single- and dual-rail rates

be treated as an erasure error, aiding in decoding.

VII. Conclusion

We explore covert entanglement generation over the lossy-thermal noise bosonic channel. We obtain a single-letter expression for the optimal covert entanglement generation capacity L_{EG} and find that it is the same as the classical covert capacity over the bosonic channel derived in [11]. Further, we investigate covert entanglement generation using photonic qubits, and find a substantial room to improve. This and much more is left for future investigation.

The key requirement for the proposed entanglement generation scheme that achieves L_{EG} is $\mathcal{O}(n)$ bits, instead of $\mathcal{O}(\sqrt{n})$ bits as in the finite-dimensional case [18]. However, the key requirement for the single-rail and dual-rail entanglement generation schemes is substantially less: $\mathcal{O}(\sqrt{n}\log n)$ bits and is due to the sparse coding. Adapting Theorem 2 to use sparse coding may yield a reduced key size.

Finally, we comment on the quantum resources required for implementing the proposed single-rail and dual-rail protocols, specifically the random stabilizer codes. Our results apply in the limit of asymptotically large codes. In both proposed protocols, this requires Alice and Bob to store large numbers of physical qubits in quantum memories and apply large Clifford circuits to them. Therefore, covert entanglement generation needs to be investigated under practical constraints on quantum memory and circuit sizes. We anticipate success using high-rate quantum codes, e.g., quantum low-density parity check (LDPC) codes [46].

Acknowledgement

The authors are grateful to Mehrdad Tahmasbi for providing the details on the sparse coding analysis in [41], which formed the basis for Appendix VII. The authors also benefited from discussions with Matthieu R. Bloch, Christos N. Gagatsos, Brian J. Smith, Ryan Camacho, Narayanan Rengaswamy, Kenneth Goodenough, and Saikat Guha.

References

[1] E. J. D. Anderson, C. K. Eyre, I. M. Dailey, F. Rozpędek, and B. A. Bash, "Square root law for covert quantum communication over optical channels," in *Proc. IEEE Int. Conf. Quantum Comput. Eng. (QCE)*, Montréal, QC, Canada, 2024.

[2] E. J. D. Anderson, M. S. Bullock, F. Rozpędek, and B. A. Bash, "Achievability of covert quantum communication," 2025, accepted to the 2025 IEEE International Symposium on Information Theory (ISIT). [Online]. Available: https://arxiv.org/abs/2501.13103

[3] A. J. Menezes, S. A. Vanstone, and P. C. V. Oorschot, *Handbook of Applied Cryptography*. Boca Raton, FL, USA: CRC Press, Inc., 1996.

[4] M. Bloch and J. ao Barros, *Physical-Layer Security*. Cambridge University Press, 2011.

[5] B. A. Bash, D. Goeckel, and D. Towsley, "Square root law for communication with low probability of detection on AWGN channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Cambridge, MA, Jul. 2012.

[6] ——, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.

[7] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.

[8] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3493–3503, Jun. 2016.

[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," *Nat. Commun.*, vol. 6, Oct. 2015.

[10] M. S. Bullock, C. N. Gagatsos, S. Guha, and B. A. Bash, "Fundamental limits of quantum-secure covert communication over bosonic channels," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 3, pp. 471–482, Mar. 2020.

[11] C. N. Gagatsos, M. S. Bullock, and B. A. Bash, "Covert capacity of bosonic channels," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, pp. 555–567, 2020.

[12] A. Sheikholeslami, B. A. Bash, D. Towsley, D. Goeckel, and S. Guha, "Covert communication over classical-quantum channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Barcelona, Spain, Jul. 2016.

[13] M. S. Bullock, A. Sheikholeslami, M. Tahmasbi, R. C. Macdonald, S. Guha, and B. A. Bash, "Fundamental limits of covert communication over classical-quantum channels," 2025, to appear in IEEE Trans. Inf. Theory. [Online]. Available: https://arxiv.org/abs/1601.06826

[14] B. A. Bash, D. Goeckel, S. Guha, and D. Towsley, "Hiding information in noise: Fundamental limits of covert wireless communication," *IEEE Commun. Mag.*, vol. 53, no. 12, 2015.

[15] X. Chen, J. An, Z. Xiong, C. Xing, N. Zhao, F. R. Yu, and A. Nallanathan, "Covert communications: A comprehensive survey," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 2, pp. 1173–1198, 2023.

[16] M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge University Press, 2016.

[17] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: A vision for the road ahead," *Science*, vol. 362, no. 6412, 2018. [Online]. Available: https://science.sciencemag.org/content/362/6412/eaam9288

[18] O. Kimelfeld, B. A. Bash, and U. Pereg, "Covert entanglement generation and secrecy," 2025. [Online]. Available: https://arxiv.org/abs/2503.21002

[19] S.-Y. Wang, T. Erdoğan, and M. Bloch, "Towards a characterization of the covert capacity of bosonic channels under trace distance," in *2022 IEEE International Symposium on Information Theory (ISIT)*. IEEE Press, 2022, p. 318–323.

[20] ——, "Towards a characterization of the covert capacity of bosonic channels under trace distance," https://bloch.ece.gatech.edu/ISIT2022-covert-bosonic.pdf, accessed Apr. 29, 2025, 2022.

[21] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, "Fundamental Limits of Repeaterless Quantum Communications," *Nature Communications*, vol. 8, no. 1, p. 15043, Apr. 2017.

[22] K. Noh, V. V. Albert, and L. Jiang, "Quantum capacity bounds of gaussian thermal loss channels and achievable rates with Gottesman-Kitaev-Preskill codes," *IEEE Trans. Inf. Theory*, vol. 65, no. 4, pp. 2563–2582, 2019.

[23] K. Noh, S. Pirandola, and L. Jiang, "Enhanced energy-constrained quantum communication over bosonic gaussian channels," *Nat. Commun.*, vol. 11, no. 1, p. 457, Jan. 2020.

[24] M. Fanizza, F. Kianvash, and V. Giovannetti, "Estimating quantum and private capacities of gaussian channels via degradable extensions," *Phys. Rev. Lett.*, vol. 127, p. 210501, Nov 2021.

[25] P. Thomas, L. Ruscio, O. Morin, and G. Rempe, "Efficient generation of entangled multiphoton graph states from a single atom," *Nature*, vol. 608, no. 7924, pp. 677–681, Aug. 2022.

[26] S. Takeda, T. Mizuta, M. Fuwa, P. van Loock, and A. Furusawa, "Deterministic quantum teleportation of photonic quantum bits by a hybrid technique," *Nature*, vol. 500, no. 7462, pp. 315–318, Aug. 2013.

[27] S. Guha, H. Krovi, C. A. Fuchs, Z. Dutton, J. A. Slater, C. Simon, and W. Tittel, "Rate-loss analysis of an efficient quantum repeater architecture," *Phys. Rev. A*, vol. 92, p. 022357, Aug 2015.

[28] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres," *Nature*, vol. 526, no. 7575, pp. 682–686, Oct. 2015.

[29] V. Krutyanskiy, M. Galli, V. Krcmarsky, S. Baier, D. A. Fioretto, Y. Pu, A. Mazloom, P. Sekatski, M. Canteri, M. Teller, J. Schupp, J. Bate, M. Meraner, N. Sangouard, B. P. Lanyon, and T. E. Northup, "Entanglement of trapped-ion qubits separated by 230 meters," *Phys. Rev. Lett.*, vol. 130, p. 050803, Feb 2023.

[30] P. Dhara, D. Englund, and S. Guha, "Entangling quantum memories via heralded photonic bell measurement," *Phys. Rev. Res.*, vol. 5, p. 033149, Sep 2023.

[31] K. Azuma, S. E. Economou, D. Elkouss, P. Hilaire, L. Jiang, H.-K. Lo, and I. Tzitrin, "Quantum repeaters: From quantum networks to the quantum internet," *Rev. Mod. Phys.*, vol. 95, p. 045006, Dec 2023.

[32] T. Honjo, S. W. Nam, H. Takesue, Q. Zhang, H. Kamada, Y. Nishida, O. Tadanaga, M. Asobe, B. Baek, R. Hadfield, S. Miki, M. Fujiwara, M. Sasaki, Z. Wang, K. Inoue, and Y. Yamamoto, "Long-distance entanglement-based quantum key distribution over optical fiber," *Opt. Express*, vol. 16, no. 23, pp. 19 118–19 126, Nov 2008.

[33] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Rev. Mod. Phys.*, vol. 81, pp. 1301–1350, Sep. 2009.

[34] J. M. Arrazola and V. Scarani, "Covert Quantum Communication," *Phys. Rev. Lett.*, vol. 117, no. 25, p. 250503, Dec. 2016.

[35] M. Tahmasbi and M. R. Bloch, "Framework for covert and secret key expansion over classical-quantum channels," *Phys. Rev. A*, vol. 99, p. 052329, May 2019.

[36] ——, "Toward undetectable quantum key distribution over bosonic channels," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, no. 2, pp. 585–598, 2020.

[37] ——, "Covert and secret key expansion over quantum channels under collective attacks," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 7113–7131, Nov. 2020.

[38] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Rev. Mod. Phys.*, vol. 84, pp. 621–669, May 2012.

[39] M. M. Wilde, "Position-based coding and convex splitting for private communication over quantum channels," *Quantum Inf. Process.*, vol. 16, no. 10, p. 264, Sep. 2017.

[40] A. Winter, "Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints," *Commun. Math. Physics*, pp. 1–23, 2016.

[41] M. Tahmasbi, B. A. Bash, S. Guha, and M. Bloch, "Signaling for covert quantum sensing," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, 2021, pp. 1041–1045.

[42] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, "Symmetrized characterization of noisy quantum processes," *Science (New York, N.Y.)*, vol. 317, no. 5846, pp. 1893–1896, Sep. 2007.

[43] M. S. Bullock, A. Sheikholeslami, M. Tahmasbi, R. C. Macdonald, S. Guha, and B. A. Bash, "Covert Communication over Classical-Quantum Channels," arXiv:1601.06826v7 [quant-ph], Jul. 2023.

[44] M. Ruskai and F. H. Stillinger, "Convexity inequalities for estimating free energy and relative entropy," *J. Phys. A*, vol. 23, no. 12, pp. 2421–2437, Jun. 1990.

[45] A. Berk, G. P. Anderson, P. K. Acharya, L. S. Bernstein, L. Muratov, J. Lee, M. Fox, S. M. Adler-Golden, J. H. Chetwynd, Jr., M. L. Hoke, R. B. Lockwood, J. A. Gardner, T. W. Cooley, C. C. Borel, P. E. Lewis, and E. P. Shettle, "MODTRAN5: 2006 update," vol. 6233, 2006, pp. 62 331F–62 331F–8.

[46] N. P. Breuckmann and J. N. Eberhardt, "Quantum low-density parity-check codes," *PRX Quantum*, vol. 2, no. 4, p. 040101, 2021.

[47] M. M. Wilde, S. Khatri, E. Kaur, and S. Guha, "Second-order coding rates for key distillation in quantum key distribution," 2019.

[48] M. E. Shirokov, "Entropy characteristics of subsets of states." *Izvestiya: Mathematics*, vol. 70, no. 6, p. 1265–1292, Dec. 2006.

[49] M. Orszag, *Quantum Optics*, 3rd ed. Berlin, Germany: Springer, 2016.

[50] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. Elsevier Academic Press, 2007.

APPENDIX I

Proof: Lemma 1. Alice and Bob employ the construction in the proof of Theorem 1. Thus, combination of Lemma 7 in Appendix IV, [47, Cor. 11] and [11, Eq. (12)] gives us $E_{\mathcal{C}}\left[\bar{P}_e(\mathcal{C})\right] \leq \epsilon$ and $E_{\mathcal{C}}\left[\frac{1}{2}\left\|\hat{\rho}^{\hat{\gamma}n}_{W^nE^n} - \hat{\rho}^{\hat{\gamma}\otimes n}_{WE}\right\|\right] \leq \kappa - \gamma_2$, for every m, if

$$\begin{aligned}\log|\mathcal{M}| &\geq nD(\hat{\rho}_{XB}\|\hat{\rho}_X\otimes\hat{\rho}_B) \\ &\quad + \sqrt{nV(\hat{\rho}_{XB}\|\hat{\rho}_X\otimes\hat{\rho}_B)}\Phi^{-1}(\epsilon - C_n^{(1)}) \\ \log|\mathcal{K}| &\leq nD(\hat{\rho}_{XWE}\|\hat{\rho}_X\otimes\hat{\rho}_{WE})\end{aligned} \tag{29}$$

$$- \sqrt{nV(\hat{\rho}_{XWE}\|\hat{\rho}_X \otimes \hat{\rho}_{WE})}\Phi^{-1}\left(\frac{\kappa^2}{10} - C_n^{(2)}\right) \tag{30}$$

where $C_n^{(1)} = \gamma_1 + \frac{C_{\text{BE}}T(\hat{\rho}_{XB}\|\hat{\rho}_X\otimes\hat{\rho}_B)}{\sqrt{nV(\hat{\rho}_{XB}\|\hat{\rho}_X\otimes\hat{\rho}_B)}}$ and $C_n^{(2)} = \gamma_2 + \gamma_3 + \frac{C_{\text{BE}}T(\hat{\rho}_{XWE}\|\hat{\rho}_X\otimes\hat{\rho}_{WE})}{\sqrt{nV(\hat{\rho}_{XWE}\|\hat{\rho}_X\otimes\hat{\rho}_{WE})}}$ and C_{BE} is the Berry-Esseen constant. Now, setting $\epsilon = e^{-\varsigma_n^{(1)}\sqrt{n}}$ for $\varsigma_n^{(1)} \in \omega\left(\frac{1}{\sqrt{n}}\right) \cap o(1)$, $\kappa = e^{-\varsigma_n^{(3)}\sqrt{n}}$ for $\varsigma_n^{(3)} \in o(1) \cap \omega\left(\frac{1}{\sqrt{n}}\right)$, and $\gamma_1 \in o\left(e^{-\varsigma_n^{(1)}\sqrt{n}}\right)$ and $\gamma_2 = \gamma_3 \in o\left(e^{-\varsigma_n^{(3)}\sqrt{n}}\right)$, we have by [11, Lem. 1, Lem. 2, Thm. 1], [19, Eq. (9)], $\varsigma_n \in o(1)$, and for n large enough

$$\log|\mathcal{M}| \geq (1-\varsigma_n)nD(\hat{\rho}_{XB}\|\hat{\rho}_X\otimes\hat{\rho}_B) \tag{31}$$
$$= (1-\varsigma_n)n\eta\log\left(1+\frac{1}{(1-\eta)\bar{n}_b}\right)\bar{n}_s \tag{32}$$
$$\log|\mathcal{K}| \leq (1+\varsigma_n)nD(\hat{\rho}_{XWE}\|\hat{\rho}_X\otimes\hat{\rho}_{WE}) \tag{33}$$
$$\leq (1+\varsigma_n)n\left(g(\eta\bar{n}_b) + g(\bar{n}_b) - 2g(\nu - 1/2)\right) \tag{34}$$

while $E_{\mathcal{C}}\left[\bar{P}_e(\mathcal{C})\right] \leq e^{-\varsigma_n^{(1)}\sqrt{n}}$ and

$$\max_m E_{\mathcal{C}}\left[\frac{1}{2}\left\|\hat{\bar{\rho}}^m_{W^nE^n} - \hat{\bar{\rho}}^{\otimes n}_{WE}\right\|\right] \leq e^{-\varsigma_n^{(3)}\sqrt{n}}, \tag{35}$$

where $g(x) \triangleq (1+x)\log(1+x) - x\log(x)$ and $\nu = \frac{1}{2}\sqrt{1+2\bar{n}_b(1-\eta)}$. The bound on $D(\hat{\rho}_{XWE}\|\hat{\rho}_X\otimes\hat{\rho}_{WE})$ in (34) is derived in Appendix V for completeness and follows from the assumption $\bar{n}_s \in o(1)$. Now, we show that (35) yields (6). By convexity, we have

$$E_{\mathcal{C}}\left[\frac{1}{2}\left\|\hat{\bar{\rho}}_{W^n} - \hat{\bar{\rho}}^{\otimes n}_W\right\|\right] \leq E_{\mathcal{C}}\left[\frac{1}{2}\left\|\hat{\bar{\rho}}^m_{W^nE^n} - \hat{\bar{\rho}}^{\otimes n}_{WE}\right\|\right] \tag{36}$$
$$\leq e^{-\varsigma_n^{(3)}\sqrt{n}}. \tag{37}$$

Now, note that

$$E_{\mathcal{C}}\left[\left|D(\hat{\bar{\rho}}_{W^n}\|\hat{\bar{\rho}}^{\otimes n}_W) - D(\hat{\bar{\rho}}^{\otimes n}_W\|\hat{\rho}^{(0)}_{W^n})\right|\right]$$
$$= E_{\mathcal{C}}\left[\left|\left(S\left(\hat{\bar{\rho}}^{\otimes n}_{W^n}\right) - S\left(\hat{\bar{\rho}}_{W^n}\right)\right)\right.\right. \tag{38}$$
$$\left.\left. + \operatorname{tr}\left(\left(\hat{\bar{\rho}}^{\otimes n}_{W^n} - \hat{\bar{\rho}}_{W^n}\right)\log\left(\hat{\rho}^{(0)}_{W^n}\right)\right)\right|\right]. \tag{39}$$

Let $E = \max(n\bar{n}_s, n\bar{n}_b) \in \mathcal{O}(n)$. By [40, Lemma 15],

$$\left|S\left(\hat{\bar{\rho}}^{\otimes n}_W\right) - S\left(\hat{\bar{\rho}}_{W^n}\right)\right|$$
$$\leq \left\|\hat{\bar{\rho}}_{W^n} - \hat{\bar{\rho}}^{\otimes n}_W\right\| nS_{\max}\left(\frac{2E}{n\left\|\hat{\bar{\rho}}_{W^n} - \hat{\bar{\rho}}^{\otimes n}_W\right\|}\right)$$
$$+ h\left(\frac{1}{2}\left\|\hat{\bar{\rho}}_{W^n} - \hat{\bar{\rho}}^{\otimes n}_W\right\|\right) \tag{40}$$
$$\in \mathcal{O}\left(\left(\left\|\hat{\bar{\rho}}_{W^n} - \hat{\bar{\rho}}^{\otimes n}_W\right\| n\right)^2 \log\left(\frac{1}{\left\|\hat{\bar{\rho}}_{W^n} - \hat{\bar{\rho}}^{\otimes n}_W\right\|}\right)\right) \tag{41}$$

where $S_{\max}(E) < \infty$ is the maximum entropy under energy constraint $E < \infty$ and (41) follows from [40, Rem. 13], [48, Prop. 1(ii)]. Thus, (37) implies $E_{\mathcal{C}}\left[\left|S\left(\hat{\bar{\rho}}^{\otimes n}_W\right) - S\left(\hat{\bar{\rho}}_{W^n}\right)\right|\right] \leq e^{-\gamma_n^{(2)}\sqrt{n}}$ for some $\gamma_n^{(2)} \in o(1)\cap\omega\left(\frac{1}{\sqrt{n}}\right)$. Finally, we examine the second term in the expectation on the right-hand side of (39):

$$\operatorname{tr}\left(\left(\hat{\bar{\rho}}^{\otimes n}_W - \hat{\bar{\rho}}_{W^n}\right)\log\left(\hat{\rho}^{(0)}_{W^n}\right)\right)$$
$$= \operatorname{tr}\left(\left(\hat{\bar{\rho}}^{\otimes n}_W - \hat{\bar{\rho}}_{W^n}\right)\sum_{i=1}^n \log\left(\frac{e^{-\beta(\eta\bar{n}_b)\hat{N}_{W_i}}}{Z(\beta(\eta\bar{n}_b))}\right)\right) \tag{42}$$
$$= \operatorname{tr}\left(\left(\hat{\bar{\rho}}^{\otimes n}_W - \hat{\bar{\rho}}_{W^n}\right)\right) n\log\left(\frac{1}{Z(\beta(\eta\bar{n}_b))}\right)$$
$$- \beta(\eta\bar{n}_b)\operatorname{tr}\left[\sum_{i=1}^n\left(\hat{\bar{\rho}}^{\otimes n}_W - \hat{\bar{\rho}}_{W^n}\right)\hat{N}_{W_i}\right] \tag{43}$$
$$= 0. \tag{44}$$

where (44) follows from the fact that $\hat{\bar{\rho}}_{W^n}$ is a QPSK modulated coherent state codeword with mean photon number per mode equal to that of $\hat{\bar{\rho}}^{\otimes n}_W$. □

APPENDIX II

Proof: Lemma 2. We follow the standard derandomization procedure. Let us define the events $A_1 \triangleq \left\{\bar{P}_e \leq e^{-\zeta_n^{(1)}\sqrt{n}}\right\}$, $A_2 \triangleq \left\{\left|D(\hat{\bar{\rho}}_{W^n}\|\hat{\bar{\rho}}^{\otimes n}_W) - D(\hat{\bar{\rho}}^{\otimes n}_W\|\hat{\rho}^{(0)}_{W^n})\right| \leq e^{-\zeta_n^{(2)}\sqrt{n}}\right\}$, and $A_3 \triangleq \left\{\frac{1}{|\mathcal{M}|}\sum_m \frac{1}{2}\left\|\hat{\bar{\rho}}^m_{W^nE^n} - \hat{\bar{\rho}}^{\otimes n}_{WE}\right\| \leq e^{-\zeta_n^{(3)}\sqrt{n}}\right\}$. Now, Lemma 1, the union bound and Markov's inequality implies

$$\Pr\left(\bigcap_{i=1}^3 A_i\right) \geq 1 - \sum_{i=1}^3 Pr(A_i) \geq 1 - \sum_{i=1}^3 e^{-\left(\varsigma_n^{(i)} - \zeta_n^{(i)}\right)\sqrt{n}} \tag{45}$$

Thus for some $\epsilon \in (0,1)$, setting $\zeta_n^{(k)} = (1-\epsilon)\varsigma_n^{(k)}$ for $k = \{1,2,3\}$ yields the right hand side of (45) converges to unity in the limit of $n \to \infty$. Therefore, we conclude that there exists at least one coding scheme for n sufficiently large that satisfies

$$\bar{P}_e(\mathcal{C}) \leq e^{-\zeta_n^{(1)}\sqrt{n}} \tag{46}$$
$$\left|D(\hat{\bar{\rho}}_{W^n}\|\hat{\bar{\rho}}^{\otimes n}_W) - D(\hat{\bar{\rho}}^{\otimes n}_W\|\hat{\rho}^{(0)}_{W^n})\right| \leq e^{-\zeta_n^{(2)}\sqrt{n}} \tag{47}$$
$$\frac{1}{|\mathcal{M}|}\sum_m \left\|\hat{\bar{\rho}}^m_{W^nE^n} - \hat{\bar{\rho}}^{\otimes n}_{WE}\right\| \leq e^{-\zeta_n^{(3)}\sqrt{n}} \tag{48}$$

Lemma 2 follows by applying (46)-(48) to the proof of [18, Lem. 10]. Since our key size scales in n instead of $\sqrt{n}$, expurgation does not affect covertness. □

APPENDIX III

Proof: Lemma 3. We employ the strategy described in the proof of [18, Thm. 2]. The complete analysis of the strategy for the lossy-thermal noise bosonic channel is presented below for completeness. Consider the global state in (16). We now approximate the state following [18, Thm. 2].

State Approximation: We now state a Lemma that allows us to approximate our global output state:

Lemma 6. [18, Lem. 14] For $|\tau\rangle_{RMB^nW^nE^n\hat{M}}$ defined in (16), there exist phases $\{g(m,k)\}$ such that $|\tau\rangle_{RMB^nW^nE^n\hat{M}}$ can be approximated by

$$|\chi\rangle_{RMB^nW^nE^n\hat{M}}$$

$$= \frac{1}{\sqrt{|\mathcal{M}||\mathcal{K}|}} \sum_{m,k,m'} |m\rangle_M \otimes |m\rangle_R \otimes e^{ig(m,k)} |x^n(m,k)\rangle_{B^n W^n E^n} \otimes |m'\rangle_{\hat{M}} \tag{49}$$

while $\| |\tau\rangle\langle\tau|_{RMB^n W^n E^n \hat{M}} - |\chi\rangle\langle\chi|_{RMB^n W^n E^n \hat{M}} \| \leq 2\sqrt{2} e^{-\varsigma_n^{(1)}\sqrt{n}}$, where $|x^n(m,k)\rangle_{B^n W^n E^n} = V_{A\to BWE}^{\eta,\bar{n}_b \otimes n}(|x^n_{\text{coh}}(m,k)\rangle_{A^n})$.

The proof of [18, Lem. 14] found in [18, Appendix A] must be modified since our codewords are coherent states which are not orthogonal. Orthogonality is used in an intermediate step in the proof of [16, Lem. A.0.3]. However, this proof, in fact, holds for well-behaved non-orthogonal states, as shown for completeness in Appendix VI.

Decoupling: Recall that $\hat{\hat{\rho}}^m_{W^n E^n}$ is Willie's average received state coupled with the environment output systems E^n given message m in the coding scheme analyzed in Lemma 2. The decoupling steps follow [18, Sec. VI.B.3] by replacing W^n with $W^n E^n$ and substituting appropriate secrecy bounds (11).

Obtaining Bipartite Entanglement Without Assistance: As in [18], we convert the GHZ state into bipartite entanglement by using a classical link and then show that the classical link is not necessary. We follow the same procedure as that of [18, Sec. VI.B.4, Sec. VI.B.5] with appropriate substitution of reliability and secrecy bounds.

Covertness: We must adapt [18, Sec. VI.B.6] to hold for infinite-dimensional output states at Willie. Willie's received state is given by $\hat{\hat{\rho}}_{W^n} = \text{tr}_{RMB^n E^n \hat{M}\hat{K}} \left[|\tau\rangle\langle\tau|_{RMB^n W^n \hat{M}\hat{K}} \right]$ which we may approximate by the state obtained in Lemma 6 with appropriate subsystems traced out: $\text{tr}_{RMB^n E^n \hat{M}\hat{K}} \left[|\chi\rangle\langle\chi|_{RMB^n W^n \hat{M}\hat{K}} \right] = \hat{\hat{\rho}}_{W^n}$. By the triangle inequality and the monotonicity of trace, we have

$$\left\| \hat{\hat{\rho}}_{W^n} - \hat{\hat{\rho}}^{\otimes n}_W \right\| \leq \left\| \hat{\hat{\rho}}_{W^n} - \hat{\hat{\rho}}_{W^n} \right\| + \left\| \hat{\hat{\rho}}_{W^n} - \hat{\hat{\rho}}^{\otimes n}_W \right\| \tag{50}$$

$$\leq 2\sqrt{2} e^{-\frac{1}{2}\varsigma_n^{(1)}\sqrt{n}} + e^{-\varsigma_n^{(3)}\sqrt{n}} \tag{51}$$

Thus, following similar steps as those from (39)-(44) for appropriate substitution of states, we have $\left| D\left(\hat{\hat{\rho}}_{W^n} \middle\| \hat{\rho}^{(0)}_{W^n}\right) - D\left(\hat{\hat{\rho}}^{\otimes n}_{W^n} \middle\| \hat{\rho}^{(0)}_{W^n}\right) \right| \leq e^{-\varsigma_n^{(4)}\sqrt{n}}$ for some $\varsigma_n^{(4)} \in o(1) \cap \omega(1/\sqrt{n})$ that depends on $\varsigma_n^{(1)}, \varsigma_n^{(2)}$. □

Appendix IV

Lemma 7. [19, Lem. V.1] Consider bipartate classical-quantum state $\hat{\rho}_{XA}$ and a channel $\mathcal{N}_{A\to BW}$. Then, for constants $\epsilon \in (0,1)$, $\kappa \in (0, \delta/2)$, $\gamma_1 \in \left(0, \frac{\epsilon^2}{10}\right)$, $\gamma_2 \in \left(0, \frac{\kappa}{2}\right)$, and $\gamma_3 \in \left(0, \frac{\kappa}{2} - \gamma_2\right)$, there exists a coding scheme such that

$$\log|\mathcal{M}| \geq D_H^{\epsilon^2/10 - \gamma_2}(\hat{\rho}_{XB} \| \hat{\rho}_X \otimes \hat{\rho}_B) - \log \frac{4\epsilon^2}{10\gamma_1^2}, \tag{52}$$

$$\log|\mathcal{K}| \leq D_{\max}^{\kappa/2 - \gamma_2 - \gamma_3}(\hat{\rho}_{XW} \| \hat{\rho}_X \otimes \hat{\rho}_W) + 2\log \frac{2\sqrt{2}}{\gamma_2\gamma_3} \tag{53}$$

that satisfies: $E_{\mathcal{C},K}\left[\bar{P}_e^K(\mathcal{C})\right] \leq \frac{\epsilon^2}{10}$ and $\max_m E_{\mathcal{C}}\left[\frac{1}{2}\left\|\hat{\hat{\rho}}^m_W - \hat{\hat{\rho}}_W\right\|\right] \leq \kappa - \gamma_2$, where $\hat{\hat{\rho}}_W = \text{tr}_B\left[\mathcal{N}_{A\to BW}(\hat{\rho}_A)\right]$.

Proof. The proof follows from construction and analysis in [20, App. B]. Here, we have increased the key length in (53) compared to [19, Eq. (6)], allowing for the strengthened secrecy bound compared to that of [19, Lemma V.1]. □

Appendix V

Here, we derive the bound used in (34). We have

$$D(\hat{\rho}_{XWE} \| \hat{\rho}_X \otimes \hat{\rho}_{WE}) = S(\hat{\hat{\rho}}_{WE}) - S(\hat{\rho}^x_{WE}) \tag{54}$$

$$\leq S(\hat{\hat{\rho}}_W) + S(\hat{\hat{\rho}}_E) - S(\hat{\rho}^x_{WE}). \tag{55}$$

where (54) is because displacement is unitary and (55) is due to the subadditivity of von Neumann entropy. Notice that $\hat{\hat{\rho}}_E = \hat{\rho}_{\text{th}}(\bar{n}_b)$, which implies $S(\hat{\hat{\rho}}_E) = g(\bar{n}_b)$ for $g(x) \triangleq (1+x)\log_2(1+x) - x\log_2(x)$. Now, we take a Taylor series expansion of $S(\hat{\hat{\rho}}_W)$ about the displacement term $u = \sqrt{\bar{n}_s}$. Setting $u = 0$, we have $S(\hat{\hat{\rho}}_W)\big|_{u=0} = S(\hat{\rho}_{\text{th}}(\eta\bar{n}_b)) = g(\eta\bar{n}_b)$. The remaining terms in the Taylor series expansion are computed in [11, App. A] as $\frac{dS(\hat{\hat{\rho}}_W)}{du}\Big|_{u=0} = 0$ and $\frac{d^2 S(\hat{\hat{\rho}}_W)}{du^2}\Big|_{u=0} = 2\log\left(1 + \frac{1}{\eta\bar{n}_b}\right)$, and thus, $S(\hat{\hat{\rho}}_W) = g(\eta\bar{n}_b) + \mathcal{O}(\bar{n}_s)$. We compute $S(\hat{\rho}^x_{WE})$ by noting that it is a Gaussian state and by performing symplectic transformations on the input covariance matrix (CM) of the channel. Conditioned on x, the input CM is $\frac{\hat{I}_{6\times 6}}{2}$. For $\hat{M}_{\text{TMS}}, \hat{M}_{\text{BS}}$ the symplectic transformation matrices corresponding to two-mode squeezing and a beamsplitter respectively, the resulting output CM $\hat{V}_{EWB} = \frac{1}{2}\hat{M}_{\text{BS}}\hat{M}_{\text{TMS}}\hat{M}^{\text{T}}_{\text{TMS}}\hat{M}^{\text{T}}_{\text{BS}}$. Tracing out Bob's system yields a reduced CM $\hat{V}_{WE}$, which has symplectic eigenvalues: $\nu_1 = \nu_2 = \frac{1}{2}\sqrt{1 + 2\bar{n}_{\text{B}}(1-\eta)} \triangleq \nu$. Thus, $S(\hat{\rho}^x_{WE}) = 2g(\nu - \frac{1}{2})$.

Appendix VI

Here, we modify the proof of [16, Lemma A.0.3] to show that it holds for non-orthogonal collections of states $\{|\zeta_i\rangle\}$ and $\{|\chi_i\rangle\}$. It suffices to show that the first equality in [16, Eq. (A.19)] holds for these non-orthogonal collections: $\frac{1}{N}\sum_s \langle\hat{\chi}_s|\hat{\zeta}_s\rangle = \frac{1}{N^2}\sum_{s,k,l} e^{\frac{j2\pi(l-k)s}{N}} \langle\chi_k|\zeta_l\rangle = \frac{1}{N}\sum_i \langle\chi_i|\zeta_i\rangle$.

Appendix VII
Covertness Analysis for Sparse Signaling

Here we adapt the approach from [41] to show validity of (26). By the definition of QRE, we have:

$$\left| D\left(\hat{\rho}^{(\hat{\pi})}_{W^n} \middle\| \left(\hat{\rho}^{(0)}_W\right)^{\otimes n}\right) - D\left(\left(\hat{\hat{\rho}}^{(\hat{\pi})}_W\right)^{\otimes n} \middle\| \left(\hat{\rho}^{(0)}_W\right)^{\otimes n}\right) \right|$$

$$= \left| \left(S\left(\left(\hat{\hat{\rho}}^{(\hat{\pi})}_W\right)^{\otimes n}\right) - S\left(\hat{\rho}^{(\hat{\pi})}_{W^n}\right)\right) \right. \tag{56}$$

$$\left. + \text{tr}\left(\left(\left(\hat{\hat{\rho}}^{(\hat{\pi})}_W\right)^{\otimes n} - \hat{\rho}^{(\hat{\pi})}_{W^n}\right)\log\left(\hat{\rho}^{(0)}_{W^n}\right)\right)\right|. \tag{57}$$

where $S(\hat{\rho}) = -\text{tr}(\hat{\rho}\log\hat{\rho})$ is the von Neumann entropy of quantum state $\hat{\rho}$. We upper-bound the last two terms of (57). First, denote $\epsilon \equiv \frac{1}{2}\left\|\left(\hat{\hat{\rho}}^{(\hat{\pi})}_W\right)^{\otimes n} - \hat{\rho}^{(\hat{\pi})}_{W^n}\right\|_1$, and note

$$\frac{1}{2}\left\|\left(\hat{\hat{\rho}}^{(\hat{\pi})}_W\right)^{\otimes n} - \hat{\rho}^{(\hat{\pi})}_{W^n}\right\|_1 \leq \frac{1}{2}\sum_{\mathbf{x}}\left|\prod_{i=1}^n p_X(x_i) - p_{\mathbf{X}}(\mathbf{x})\right| \tag{58}$$

$$= p\left(\bar{\mathcal{A}}\right) \leq 2e^{-\frac{1}{3}qn\vartheta^2}, \tag{59}$$

where (58) is by the data processing inequality with classical statistical (total variation) distance, $p\left(\bar{\mathcal{A}}\right) = 1 - p\left(\mathcal{A}\right)$, and the inequality in (59) is the Chernoff bound. For our setting of $q \propto 1/\sqrt{n}$, ϵ decays to zero exponentially in $\sqrt{n}$.

Let $\hat{H}_{W_i} = \hbar\omega(\hat{n}_{W_i} + \frac{1}{2})$ be the Hamiltonian for the i^{th} system at Willie where $\bar{n}_{W_i}$ is the number operator, and let $E \equiv \max\left(\sum_{i=1}^{n} \operatorname{tr}\left(\hat{\bar{\rho}}_W^{(\hat{\pi})} \hat{H}^{W_i}\right), \sum_{i=1}^{n} \operatorname{tr}\left(\hat{\rho}_{W^n}^{(\hat{\pi})} \hat{H}^{W_i}\right)\right) = O(n)$, where $f(n) = O(g(n))$ means that $f(n)$ grows no faster asymptotically than $g(n)$: $\limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty$. By [40, Lemma 15],

$$S\left(\left(\hat{\bar{\rho}}_W^{(\hat{\pi})}\right)^{\otimes n}\right) - S\left(\hat{\rho}_{W^n}^{(\hat{\pi})}\right) \leq 2\epsilon n S_{\max}\left(\frac{E}{\epsilon n}\right) + h(\epsilon), \tag{60}$$

where $S_{\max}(E) < \infty$ is the maximum entropy under energy constraint $E < \infty$. [40, Rem. 13], [48, Prop. 1(ii)] and the fact that $\epsilon n \to 0$ implies that (60) vanishes as $n \to \infty$.

To bound the last term in (57), we decompose $\left(\hat{\bar{\rho}}_W^{(\hat{\pi})}\right)^{\otimes n} = p\left(\mathcal{A}\right)\hat{\rho}_{W^n}^{(\hat{\pi})} + p\left(\bar{\mathcal{A}}\right)\hat{\sigma}_{W^n}^{(\hat{\pi})}$, where $\hat{\sigma}_{W^n}^{(\hat{\pi})}$ is a density operator with $E_{\hat{\sigma}} \equiv \sum_{i=1}^{n} \operatorname{tr}\left(\hat{\sigma}_{W^n}^{(\hat{\pi})} \hat{H}_{W_i}\right) = O(n)$. Let $\hat{\Delta}_{W^n}^{(\hat{\pi})} \equiv \hat{\rho}_{W^n}^{(\hat{\pi})} - \hat{\sigma}_{W^n}^{(\hat{\pi})}$ and note that $\operatorname{tr}\left(\hat{\Delta}_{W^n}^{(\hat{\pi})}\right) = 0$ and $E_{\hat{\Delta}} \equiv \sum_{i=1}^{n} \operatorname{tr}\left(\hat{\Delta}_{W^n}^{(\hat{\pi})} \hat{H}_{W_i}\right) = O(n)$. Then,

$$\begin{aligned}
&\operatorname{tr}\left(\left(\left(\hat{\bar{\rho}}_W^{(\hat{\pi})}\right)^{\otimes n} - \hat{\rho}_{W^n}^{(\hat{\pi})}\right) \log\left(\left(\hat{\rho}_{W^n}^{(0)}\right)\right)\right) \\
&= p\left(\bar{\mathcal{A}}\right) \operatorname{tr}\left(\hat{\Delta}_{W^n}^{(\hat{\pi})} \log\left(\left(\hat{\rho}_{W^n}^{(0)}\right)\right)\right) &(61)\\
&= p\left(\bar{\mathcal{A}}\right) \operatorname{tr}\left(\hat{\Delta}_{W^n}^{(\hat{\pi})} \sum_{i=1}^{n} \log\left[\frac{1}{Z(\beta(E_0))} e^{-\beta(E_0)\hat{H}_{W_i}}\right]\right) &(62)\\
&= p\left(\bar{\mathcal{A}}\right) n \log\left(\frac{1}{Z(\beta(E_0))}\right) \operatorname{tr}\left(\hat{\Delta}_{W^n}^{(\hat{\pi})}\right) - p\left(\bar{\mathcal{A}}\right)\beta(E_0) E_{\hat{\Delta}} \\
&= o(1), &(63)
\end{aligned}$$

where (62) is due to operator exponentiation and that $\hat{\rho}_W^{(0)} = \frac{1}{Z(\beta(E_0))} e^{-\beta(E_0)\hat{H}_W}$ with $\hat{H}_W = \hat{n}_w$ is a thermal state, and (63) follows from (59).

Appendix VIII
Pauli Twirling Channel Parameters

In the Alice-to-Bob scenario, Pauli twirling works in the following way: Alice chooses a random Pauli operator and applies it to the state here before transmission. The state is transmitted through the channel, and Bob applies the same Pauli to "undo" the twirling, effectively yielding the transformation for a single qubit: $\hat{\rho}_B = \frac{1}{4}\sum_{i=0}^{3} P_i \mathcal{N}(P_i \hat{\rho}_A P_i) P_i$, where $P_i \in \{I, X, Y, Z\}$ indexed from 0 to 3 respectively.

Given a channel with Kraus operators K_j indexed by j, the Choi state of the channel is written as $\hat{\rho}^{\text{Choi}} = \sum_{j=0}^{\infty} (I \otimes K_j) \left|\Phi^+\right\rangle \left\langle\Phi^+\right| (I \otimes K_j)$, where $\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is the first Bell state. Then the Pauli channel parameters generated by Pauli twirling are given in terms of the Choi state:

$$p_I = \left\langle\Phi^+\right| \hat{\rho}^{\text{Choi}} \left|\Phi^+\right\rangle, p_X = \left\langle\Psi^+\right| \hat{\rho}^{\text{Choi}} \left|\Psi^+\right\rangle, \tag{64}$$

$$p_Y = \left\langle\Psi^-\right| \hat{\rho}^{\text{Choi}} \left|\Psi^-\right\rangle, p_Z = \left\langle\Phi^-\right| \hat{\rho}^{\text{Choi}} \left|\Phi^-\right\rangle, \tag{65}$$

with $\left|\Phi^+\right\rangle, \left|\Psi^+\right\rangle, \left|\Phi^-\right\rangle, \left|\Psi^-\right\rangle$ being the standard Bell states.

Let us now apply this to the lossy thermal-noise channel followed by projection. The lossy thermal-noise channel can be decomposed into a pure-loss channel of transmissivity $\tau = \eta/G$ followed by a quantum-limited amplifier with gain coefficient $G = 1 + (1-\eta)\bar{n}_{\text{B}}$. As we are primarily concerned with the Choi matrix, the decomposition has the following Kraus operator representation with input state $\hat{\rho} = \left|\Phi^+\right\rangle\left\langle\Phi^+\right|$: $\mathcal{E}_{A\to B}^{\eta,\bar{n}_{\text{B}}}(\hat{\rho}) = \sum_{k,l} \mathbf{B}_k \mathbf{A}_l \hat{\rho} \mathbf{A}_l^\dagger \mathbf{B}_k^\dagger$ with $\mathbf{A}_l = \mathbf{I} \otimes \sqrt{\frac{(1-\tau)^l}{l!}} \tau^{\frac{\hat{n}}{2}} \hat{a}^l$, and $\mathbf{B}_k = \mathbf{I} \otimes \sqrt{\frac{1}{k!}\frac{1}{G}\left(\frac{G-1}{G}\right)^k} \hat{a}^{\dagger k} G^{-\frac{\hat{n}}{2}}$ where $\hat{a}, \hat{a}^\dagger$ are the annihilation and creation operators respectively and $\hat{n} = \hat{a}^\dagger\hat{a}$ is the number operator.

Furthermore, Bob projects onto the computational basis on receiving the state from the physical channel output, where the projection operator is $\hat{\mathbf{\Pi}} = I \otimes |0\rangle\langle 0| + I \otimes |1\rangle\langle 1| = \left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \left|\Phi^-\right\rangle\left\langle\Phi^-\right| + \left|\Psi^+\right\rangle\left\langle\Psi^+\right| + \left|\Psi^-\right\rangle\left\langle\Psi^+\right|$. The process of projection is probabilistic with probability of failure $p = 1 - \frac{1+(1-\eta)\bar{n}_{\text{B}}\left(3+2\bar{n}_{\text{B}}-2\eta\left(\bar{n}_{\text{B}}+\frac{1}{2}\right)\right)}{\left(1+(1-\eta)\bar{n}_{\text{B}}\right)^3}$. When Bob fails to project, he replaces the state with the maximally mixed state $\frac{\hat{\pi}}{4}$. The total action of the physical channel and projection yields:

$$\hat{\rho}_B = (1-p)\hat{\mathbf{\Pi}}\left(\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \mathbf{B}_k \mathbf{A}_l \hat{\rho} \mathbf{A}_l^\dagger \mathbf{B}_k^\dagger\right)\hat{\mathbf{\Pi}} + p\frac{\hat{\pi}}{4}. \tag{66}$$

By definition, $\hat{\rho}_B$ in (66) defines the Choi state for our channel with projection when $\hat{\rho} = \left|\Phi^+\right\rangle\left\langle\Phi^+\right|$. Let us further simplify (66). The system's initial state being in the Bell basis limits $l = 0, 1$ as the annihilation operator applied more than once will guarantee the null state. Furthermore, the projection operator $\hat{\mathbf{\Pi}}$ only selects states within the Bell basis, indicating that the creation operator from $\mathbf{B}_k$ can only support $k = 0, 1$. Hence,

$$\begin{aligned}
&\sum_{k,l\in\{0,1\}} \mathbf{B}_k \mathbf{A}_l \hat{\rho} \mathbf{A}_l^\dagger \mathbf{B}_k^\dagger &(67)\\
&= \frac{1}{2}\left[\frac{1}{G}|00\rangle\langle 00| + \frac{G-1}{G^2}|01\rangle\langle 01| + \frac{1-\tau}{G}|10\rangle\langle 10|\right. \\
&\quad + \frac{G+\tau-1}{G^2}|11\rangle\langle 11| + \frac{\tau}{G}|12\rangle\langle 12| \\
&\quad + \frac{\sqrt{\tau}}{G^{\frac{3}{2}}}(|00\rangle\langle 11| + |11\rangle\langle 00|) \\
&\quad \left. + \frac{\sqrt{\tau}(G-1)}{G^{\frac{3}{2}}}(|01\rangle\langle 12| + |12\rangle\langle 01|)\right] &(68)
\end{aligned}$$

Note that the last term contains $|12\rangle$ and $\langle 12|$, which is the only one that will not be selected by the projection.

Next, we must find the Pauli channel parameters generated by twirling. Notably, we do not need to perform the inverse twirling on the maximally mixed state portion of (66) since it will also produce a maximally mixed state. Furthermore, up to this point, we have ignored the normalization; therefore, after projection,

$$\begin{aligned}
\hat{\rho}_B^{\text{Choi}} = \frac{1}{2N(G,\tau)}&\left[\frac{1}{G}|00\rangle\langle 00| + \frac{G-1}{G^2}|01\rangle\langle 01|\right. \\
&\left. + \frac{1-\tau}{G}|10\rangle\langle 10| + \frac{G+\tau-1}{G^2}|11\rangle\langle 11|\right.
\end{aligned}$$

$$+\frac{\sqrt{\tau}}{G^{\frac{3}{2}}}\left(|00\rangle\langle 11|+|11\rangle\langle 00|\right)\Big] \tag{69}$$

with normalization constant $N(G,\tau)=\frac{3G+2G^{3/2}\sqrt{\tau}+\tau-2}{2G^2}$. Evaluating (64) and (65) for (69) yields (22)-(25).

Appendix IX
χ^2-divergence derivation for single-rail qubits

We first derive the output state at Willie to calculate the χ^2-divergence between Willie's observed innocent state and when Alice transmits. In the single-rail case, Alice inputs an arbitrary qubit into the channel of the form $\hat{\rho}_A=\begin{pmatrix}|\alpha|^2 & \gamma\\ \gamma^* & |\beta|^2\end{pmatrix}$.

Utilizing anti-normally ordered characteristic functions,

$$\chi_A^{\hat{\rho}_W}(\zeta)=\chi_A^{\hat{\rho}_A}\left(\sqrt{1-\eta}\zeta\right)\chi_A^{\hat{\rho}_E}\left(-\sqrt{\eta}\zeta\right), \tag{70}$$

where $\hat{\rho}_W$ is the state at Willie, $\hat{\rho}_E$ is the environment's thermal state input. For a thermal state with mean thermal photon number $\bar{n}_{\mathrm{B}}$, the characteristic function is known [49, Sec. 7.4.3.2]:

$$\chi_A^{\hat{\rho}_E}(-\sqrt{\eta}\zeta)=e^{-(1+\bar{n}_{\mathrm{B}})\eta|\zeta|^2}, \tag{71}$$

The next step is finding the characteristic function for $\hat{\rho}_A$.

$$\chi_A^{\hat{\rho}_A}\left(\sqrt{1-\eta}\zeta\right)$$
$$=e^{-(1-\eta)|\zeta|^2}\operatorname{tr}\left(\hat{\rho}_A e^{\zeta\sqrt{1-\eta}\hat{a}^\dagger}e^{-\zeta^*\sqrt{1-\eta}\hat{a}}\right) \tag{72}$$
$$=e^{-(1-\eta)|\zeta|^2}\sum_{n=0}^{\infty}\langle n|\left(|\alpha|^2|0\rangle\langle 0|+\gamma|0\rangle\langle 1|+\gamma^*|1\rangle\langle 0|\right.$$
$$\left.+|\beta|^2|1\rangle\langle 1|\right)e^{\zeta\sqrt{1-\eta}\hat{a}^\dagger}e^{-\zeta^*\sqrt{1-\eta}\hat{a}}|n\rangle \tag{73}$$

We can evaluate (73) term by term, noting that the infinite sums drop due to orthogonality of Fock states. Let us first evaluate the $|\alpha|^2$ term:

$$|\alpha|^2\langle 0|e^{\zeta\sqrt{1-\eta}\hat{a}^\dagger}e^{-\zeta^*\sqrt{1-\eta}\hat{a}}|0\rangle \tag{74}$$
$$=|\alpha|^2\langle 0|\left(\sum_{k'=0}^{n}(\hat{a}^\dagger)^{k'}\frac{(\sqrt{1-\eta}\zeta)^{k'}}{k'!}\right)$$
$$\left(\sum_{k=0}^{n}\frac{\left(-\sqrt{1-\eta}\zeta^*\right)^k}{k!}(\hat{a})^k\right)|0\rangle \tag{75}$$
$$=|\alpha|^2 \tag{76}$$

where (76) comes from the fact that the only non-zero term is when no ladder operators are applied and $k'=k=0$.

Let us now consider the $|\beta|^2$ term:

$$|\beta|^2\langle 1|e^{\zeta\sqrt{1-\eta}\hat{a}^\dagger}e^{-\zeta^*\sqrt{1-\eta}\hat{a}}|1\rangle \tag{77}$$
$$=|\beta|^2\left(1-|\zeta|^2(1-\eta)\right) \tag{78}$$

which follows the same as the α term. Furthermore,

$$\gamma\langle 1|e^{\zeta\sqrt{1-\eta}\hat{a}^\dagger}e^{-\zeta^*\sqrt{1-\eta}\hat{a}}|0\rangle=\gamma\zeta\sqrt{1-\eta} \tag{79}$$
$$-\gamma^*\langle 0|e^{\zeta\sqrt{1-\eta}\hat{a}^\dagger}e^{-\zeta^*\sqrt{1-\eta}\hat{a}}|1\rangle=-\gamma^*\zeta^*\sqrt{1-\eta} \tag{80}$$

Putting it all together $\chi_A^{\hat{\rho}_A}$ and (73) evaluates to:

$$\chi_A^{\hat{\rho}_A}\left(\sqrt{1-\eta}\zeta\right)$$
$$=e^{-(1-\eta)|\zeta|^2}\left(|\alpha|^2+|\beta|^2\left(1-|\zeta|^2(1-\eta)\right)\right.$$
$$\left.+\gamma\zeta\sqrt{1-\eta}-\gamma^*\zeta^*\sqrt{1-\eta}\right). \tag{81}$$

A quantum state $\hat{\rho}_W$ and its characteristic function $\chi_A^{\hat{\rho}_W}(\cdot)$ are related via the operator Fourier transform [38]:

$$\hat{\rho}_W=\int\frac{d^2\zeta}{\pi}\chi_A^{\hat{\rho}_W}(\zeta)e^{\zeta\hat{w}^\dagger}e^{-\zeta^*\hat{w}}, \tag{82}$$

where the integral is over the complex plane for ζ. However, we can convert such an integral to polar coordinates using $|\zeta|^2=r^2$, $\zeta=r(\cos(\theta)+j\sin(\theta))$, $\zeta^*=r(\cos(\theta)-j\sin(\theta))$:

$$\hat{\rho}_W=\int_{r=0}^{\infty}rdr\int_{\theta=0}^{2\pi}\frac{d\theta}{\pi}\chi_A^{\hat{\rho}_W}(r,\theta)e^{r(\cos(\theta)+j\sin(\theta))\hat{w}^\dagger}$$
$$\times e^{-r(\cos(\theta)-j\sin(\theta))\hat{w}}, \tag{83}$$

Likewise, $\chi_A^{\hat{\rho}_W}$ in polar coordinates is:

$$\chi_A^{\hat{\rho}_W}(r,\theta)=e^{-(1+\eta\bar{n}_{\mathrm{B}})\eta r^2}$$
$$\times\left(1-|\beta|^2r^2(1-\eta)\right.$$
$$+r\sqrt{1-\eta}(\gamma(\cos(\theta)+j\sin(\theta)))$$
$$\left.-r\sqrt{1-\eta}(\gamma^*(\cos(\theta)-j\sin(\theta)))\right) \tag{84}$$

where the α-dependence is gone due to $|\alpha|^2+|\beta|^2=1$. Furthermore, a generalized single-mode state in the Fock basis is written as

$$\hat{\rho}_W=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\langle m|\hat{\rho}_W|n\rangle|m\rangle\langle n|. \tag{85}$$

where we replace the right hand side $\hat{\rho}_W$ with that of (82) and evaluate the integrals to get the following state for $\hat{\rho}_W$:

$$\hat{\rho}_W=\sum_{m=0}^{\infty}\left[\frac{(\eta\bar{n}_{\mathrm{B}})^m}{(1+\eta\bar{n}_{\mathrm{B}})^{m+1}}\right.$$
$$\left.+|\beta|^2(1-\eta)\frac{(\eta\bar{n}_{\mathrm{B}})^{m-1}(m-\eta\bar{n}_{\mathrm{B}})}{(1+\eta\bar{n}_{\mathrm{B}})^{m+2}}\right]|m\rangle\langle m|$$
$$-\gamma\sqrt{(1-\eta)(m+1)}\frac{(\eta\bar{n}_{\mathrm{B}})^m}{(1+\eta\bar{n}_{\mathrm{B}})^{m+2}}|m\rangle\langle m+1|$$
$$-\gamma^*\sqrt{(1-\eta)m}\frac{(\eta\bar{n}_{\mathrm{B}})^{m-1}}{(1+\eta\bar{n}_{\mathrm{B}})^{m+1}}|m\rangle\langle m-1|. \tag{86}$$

As Alice applies Pauli twirling, or inputs 1/2 of a Bell pair into the channel, from Willie's perspective, on average, it is the maximally mixed state with $\alpha=\beta=1/\sqrt{2}$ and $\gamma=\gamma^*=0$. In this case (86) becomes

$$\hat{\rho}_W=\sum_{m=0}^{\infty}\left[\frac{(\eta\bar{n}_{\mathrm{B}})^m}{(1+\eta\bar{n}_{\mathrm{B}})^{m+1}}\right.$$
$$\left.+\frac{(1-\eta)}{2}\frac{(\eta\bar{n}_{\mathrm{B}})^{m-1}(m-\eta\bar{n}_{\mathrm{B}})}{(1+\eta\bar{n}_{\mathrm{B}})^{m+2}}\right]|m\rangle\langle m| \tag{87}$$

While Willie's output state when Alice is quiet is an attenuated thermal state $\hat{\rho}_W^{(0)}=\hat{\rho}_{\eta\bar{n}_{\mathrm{B}}}$ defined in III-A.

Since both $\hat{\rho}_W$ and $\hat{\rho}_W^{(0)}$ are diagonal in the Fock basis, calculation of $D_{\chi^2}\left(\hat{\rho}_W\middle\|\hat{\rho}_W^{(0)}\right)=\operatorname{tr}[(\hat{\rho}_W)^2(\hat{\rho}_W^{(0)})^{-1}]-1$ is straightforward. We square the coefficients of $\hat{\rho}_W$ and take the reciprocal of $\hat{\rho}_W^{(0)}$'s coefficients, and multiply them term by term. Taking the infinite sum of the elements, and using known identities from [50], yields $D_{\chi^2}\left(\hat{\rho}_W\middle\|\hat{\rho}_W^{(0)}\right)=\frac{(1-\eta)^2}{4\eta\bar{n}_{\mathrm{B}}(1+\eta\bar{n}_{\mathrm{B}})}$.