UMass Technical Report UM-CS-2014-004

Covert Optical Communication

Boulat A. Bash,[1,2] Andrei H. Gheorghe,[2,3] Monika Patel,[2]
Jonathan L. Habif,[2] Dennis Goeckel,[4] Don Towsley,[1] and Saikat Guha[2]
[1] *School of Computer Science, University of Massachusetts, Amherst, Massachusetts, USA 01003,*
[2] *Quantum Information Processing Group, Raytheon BBN Technologies, Cambridge, Massachusetts, USA 02138,*
[3] *Amherst College, Amherst, Massachusetts, USA 01002,*
[4] *Electrical and Computer Engineering Department,*
University of Massachusetts, Amherst, Massachusetts, USA 01003 *

Encryption prevents unauthorized decoding, but does not ensure stealth—a security demand that a mere presence of a message be undetectable. We characterize the ultimate limit of covert communication that is secure against the most powerful physically-permissible adversary. We show that, although it is impossible over a pure-loss channel, covert communication is attainable in the presence of any excess noise, such as a 300K thermal blackbody. In this case, $\mathcal{O}(\sqrt{n})$ bits can be transmitted reliably and covertly in n optical modes using standard optical communication equipment. The all-powerful adversary may intercept all transmitted photons not received by the intended receiver, and employ arbitrary quantum memory and measurements. Conversely, we show that this square root scaling cannot be outperformed. We corroborate our theory in a proof-of-concept experiment. We believe that our findings will enable practical realizations of covert communication and sensing, both for point-to-point and networked scenarios.

Encryption prevents unauthorized access to transmitted information—a security need critical to modern-day electronic communication. Conventional computationally-secure encryption [1, 2], information-theoretic secrecy [3, 4], and quantum cryptography [5] offer progressively higher levels of security. Quantum key distribution (QKD) allows two distant parties to generate shared secret keys over a lossy-noisy channel that are secure from the most powerful adversary allowed by physics. This shared secret, when subsequently used to encrypt data using the one-time-pad cipher [6], yields the most powerful form of encryption. However, encryption does not mitigate the threat to the users' privacy from the discovery of the very existence of the message itself (e.g., seeking of "meta-data" as detailed in the recent Snowden disclosures [7]), nor does it provide the means to communicate when the adversary forbids it. Thus, low probability of detection (LPD) or *covert* communication systems are desirable that not only protect the message content, but also prevent the detection of the transmission attempt. Here we delineate, and experimentally demonstrate, the ultimate limit of covert communication that is secure against the most powerful adversary physically permissible—the same benchmark of security to which quantum cryptography adheres for encrypted communication.

Covert communication is an ancient discipline [8] revived by the communication revolution of the last century. Modern developments include spread-spectrum radio-frequency (RF) communication [9], where the signal power is suppressed below the noise floor by bandwidth expansion; and steganography [10], where messages are hidden in fixed-size, finite-alphabet covertext objects such as digital images. We recently characterized the information-theoretic limit of classical covert communication on an additive white Gaussian noise (AWGN) channel, the standard model for RF channels [11, 12]. We showed that the sender Alice can reliably transmit $\mathcal{O}(\sqrt{n})$ bits to the intended receiver Bob in n AWGN channel uses with arbitrarily low probability of detection by the adversary Willie. Thus, a non-trivial burst of covert bits can be transmitted when n is large. Our work was generalized to other channel settings [13–16]. Similar square-root laws were also found in steganography[17], where it was shown that Alice can modify $\mathcal{O}(\sqrt{n})$ symbols in a covertext of size n, embedding $\mathcal{O}(\sqrt{n}\log n)$ hidden bits [10, 18–22].

Optical signaling [23, 24] is particularly attractive for covert communication due to its narrow diffraction-limited beam spread in free space [25, 26] and the ease of detecting fiber taps using

*This material is based upon work supported by the National Science Foundation under Grants CNS-1018464 and ECCS-1309573. SG was supported by the aforesaid NSF grant, under subaward number 14-007829 A 00, and the DARPA *Information in a Photon* program under contract number HR0011-10-C-0159. BAB, AHG, JLH and MP would like to acknowledge financial support from Raytheon BBN Technologies.

time-domain reflectometry [27]. Our information-theoretic analysis of covert communication on the AWGN channel also applies to a lossy optical channel with additive Gaussian noise when Alice uses a laser-light transmitter and both Bob and Willie use coherent-detection receivers. However, modern high-sensitivity optical communication components are primarily limited by noise of quantum-mechanical origin. Thus, recent studies on the performance of physical optical communication have focused on this quantum-limited regime [28–30]. Here we establish the quantum limits of covert communication. We demonstrate that covert communication is impossible over a pure-loss channel. However, when the channel has any excess noise (e.g., the unavoidable thermal noise from the blackbody radiation at the operating temperature), Alice can reliably transmit $\mathcal{O}(\sqrt{n})$ covert bits to Bob using n optical modes, even if Willie intercepts all the photons not reaching Bob and employs arbitrary quantum memory and measurements. This is achievable using standard laser-light modulation and homodyne detection (thus the Alice-Bob channel is still an AWGN channel). Thus, noise enables stealth. Indeed, if Willie's detector contributes excess noise (e.g., dark counts in photon-counting detectors), Alice can covertly communicate to Bob, even when the channel itself is pure-loss. We also show that the square-root limit cannot be outperformed. We corroborate our theoretical results with a proof-of-concept experiment, where the excess noise in Willie's detection is emulated by dark counts of his single photon detector. This is the first known implementation of a truly quantum-information-theoretically secure covert communication system that allows communication when all transmissions are prohibited.

INFORMATION-THEORETICALLY COVERT COMMUNICATION

Quantum and classical information-theoretic analyses of covert communication consider the *reliability* and *detectability* of a transmission. We introduce these concepts next.

Reliability—We consider a scenario where Alice attempts to transmit M bits to Bob using n optical modes while Willie attempts to detect her transmission attempt. Each of the 2^M possible M-bit messages maps to an n-mode *codeword*, and their collection forms a *codebook*. Since we consider single-spatial-mode fiber and free-space optical channels, each of the n modes in the codeword corresponds to a signaling interval carrying one modulation symbol. Desirable codebooks ensure that the codewords, when corrupted by the channel, are distinguishable from one another. This provides *reliability*: a guarantee that the probability of Bob's error in decoding Alice's message $\mathbb{P}_e^{(b)} < \delta$ with arbitrarily small $\delta > 0$ for n large enough. In practice, error-correction codes (ECCs) are used to enable reliability.

Detectability— Willie's detector reduces to a binary hypothesis test of Alice's transmission state given his observations of the channel. Denote by $\mathbb{P}_{\mathrm{FA}}$ the probability that Willie raises a false alarm when Alice does not transmit, and by $\mathbb{P}_{\mathrm{MD}}$ the probability that Willie misses the detection of Alice's transmission. Under the assumption of equal prior probabilities on Alice's transmission state (unequal prior probabilities do not affect the asymptotics [11]), Willie's *detection error probability*, $\mathbb{P}_e^{(w)} = (\mathbb{P}_{\mathrm{FA}} + \mathbb{P}_{\mathrm{MD}})/2$. Alice desires a reliable signaling scheme that is *covert*, i.e., ensures $\mathbb{P}_e^{(w)} \geq 1/2 - \epsilon$ for an arbitrarily small $\epsilon > 0$ regardless of Willie's quantum measurement choice (since $\mathbb{P}_e^{(w)} = 1/2$ for a random guess). By decreasing her transmission power, Alice can decrease the effectiveness of Willie's hypothesis test at the expense of the reliability of Bob's decoding. *Information-theoretically secure covert communication* is both reliable and covert. To achieve it, prior to transmission, Alice and Bob share a secret, the cost of which we assume to be substantially less than that of being detected by Willie. Secret-sharing is consistent with other information-hiding systems [10–12, 18–22]; however, as evidenced by the recent results for a restricted class of channels [14, 15], we believe that certain scenarios (e.g., Willie's channel from Alice being worse than Bob's) will allow secretless optical covert communication.

ANALYSIS OF COVERT OPTICAL COMMUNICATION

Here we outline the theoretical development of quantum-information-theoretically secure covert optical communication. Formal theorem statements are deferred to the Methods, with detailed proofs in the Supplementary Information.

Channel model—Consider a single-mode quasi-monochromatic lossy optical channel $\mathcal{E}_{\eta_b}^{\bar{n}_T}$ of

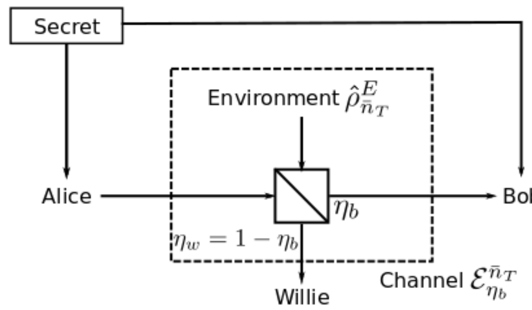

FIG. 1. Channel model. The input-output relationship is captured by a beamsplitter of transmissivity η_b, with the transmitter Alice at one of the input ports and the intended receiver Bob at one of the output ports, and η_b being the fraction of Alice's signaling photons that reach Bob. The other input and output ports of the beamsplitter correspond to the environment and the adversary Willie. Willie collects the entire $\eta_w = 1 - \eta_b$ fraction of Alice's photons that do not reach Bob. This models single-spatial-mode free-space and single-mode fiber optical channels. Alice and Bob share a secret before the transmission.

transmissivity $\eta_b \in (0, 1]$ and thermal noise mean photon number per mode $\bar{n}_T \geq 0$, as depicted in Figure 1. Willie collects the entire $\eta_w = 1 - \eta_b$ fraction of Alice's photons that do not reach Bob but otherwise remains passive, not injecting any light into the channel. Later we argue that being active does not help Willie to detect Alice's transmissions. For a pure loss channel ($\bar{n}_T = 0$), the environment input is in the *vacuum* state $\hat{\rho}^E_0 = |0\rangle\langle 0|^E$, corresponding to the minimum noise the channel must inject to preserve the Heisenberg inequality of quantum mechanics.

Pure loss insufficient for covert communication— Regardless of Alice's strategy, reliable and covert communication over a pure-loss channel to Bob is impossible. Theorem 1 in the Methods demonstrates that Willie can effectively use an ideal *single photon detector* (SPD) on each mode to discriminate between an n-mode vacuum state and any non-vacuum state in Alice's codebook. Willie avoids false alarms since no photons impinge on his SPD when Alice is silent. However, a single *click*—detection of one or more photons—gives away Alice's transmission attempt regardless of the actual quantum state of Alice's signaling photons. Alice is thus constrained to codewords that are nearly indistinguishable from vacuum, rendering unreliable any communication attempt that is designed to be covert. Furthermore, any communication attempt that is designed to be reliable cannot remain covert, as Willie detects it with high probability for large n. This is true even when Alice and Bob have access to an infinitely-large pre-shared secret. Thus, if Willie controlled the environment (as assumed in QKD proofs), by setting it to vacuum, he could deny covert communication between Alice and Bob. However, a positive amount of non-adversarial excess noise—whether from the thermal background or the detector itself—is unavoidable, which enables covert communication.

Channel noise yields the square root law—Now consider the lossy bosonic channel $\mathcal{E}^{\bar{n}_T}_{\eta_b}$, where the environment mode is in a thermal state with mean photon number $\bar{n}_T > 0$. A thermal state is represented by a mixture of coherent states $|\alpha\rangle$— quantum descriptors of ideal laser-light—weighted by a Gaussian distribution over the field amplitude $\alpha \in \mathbb{C}$, $\hat{\rho}^E_{\bar{n}_T} = \frac{1}{\pi \bar{n}_T} \int e^{-|\alpha|^2/\bar{n}_T} |\alpha\rangle\langle\alpha|^E \mathrm{d}^2\alpha$. This thermal noise masks Alice's transmission attempt, enabling covert communication even when Willie has arbitrary resources, such as access to all signaling photons not captured by Bob and any quantum-limited measurement on the light he thus captures. Theorem 2 in the Methods demonstrates that in this scenario Alice can reliably transmit $\mathcal{O}(\sqrt{n})$ covert bits using n optical modes to Bob, who needs only a conventional homodyne-detection receiver. Alice achieves this using mean photon number per mode $\bar{n} = \mathcal{O}(1/\sqrt{n})$. Conversely, Theorem 5 states that if Alice exceeds the limit of $\mathcal{O}(\sqrt{n})$ covert bits in n optical modes, transmission either is detected or unreliable.

Detector noise also enables covert communication— While any $\bar{n}_T > 0$ enables covert communication, the number of covertly-transmitted bits decreases with $\bar{n}_T$. Blackbody radiation is negligible at optical frequencies (e.g., a typical daytime value of $\bar{n}_T \approx 10^{-6}$ photons per mode at the optical telecom wavelength of 1.55μm [31]). However, other sources of excess noise can also hide the transmissions (e.g. detector dark counts and Johnson noise). To illustrate information-hiding capabilities of these noise sources, we consider the (hypothetical) pure-loss channel, for which Willie's optimal receiver is an ideal photon number resolving (PNR) detector on each mode (as discussed in the Supplementary Information). The prevalent form of excess noise afflicting PNR detectors is the *dark counts*—

erroneous detection events stemming from internal spontaneous emission processes. Thus, we consider a pure-loss channel where Willie is equipped with a PNR detector. Theorem 3 in the Methods demonstrates that, using an on-off keying (OOK) coherent state modulation where Alice transmits the *on* symbol $|\alpha\rangle$ with probability $q = \mathcal{O}(1/\sqrt{n})$ and the *off* symbol $|0\rangle$ with probability $1-q$, Alice can reliably transmit $\mathcal{O}(\sqrt{n})$ covert bits using n OOK symbols.

A structured strategy for covert communication— The skewed on-off duty cycle of OOK modulation makes construction of efficient ECCs challenging. Constraining OOK signaling to Q-ary pulse position modulation (PPM) addresses this issue by sacrificing a constant fraction of throughput. Each PPM symbol uses a PPM *frame* to transmit a sequence of Q coherent state pulses, $|0\rangle \ldots |\alpha\rangle \ldots |0\rangle$, encoding message $i \in \{1, 2, \ldots, Q\}$ by transmitting $|\alpha\rangle$ in the i^{th} mode of the PPM frame. Thus, instead of $\mathcal{O}(n)$ bits allowed by OOK, PPM lets $\mathcal{O}\left(\frac{n \log Q}{Q}\right)$ bits be transmitted in n optical modes. However, PPM performs well in the low photon number regime [32] and the symmetry of its symbols enables the use of many efficient ECCs.

To communicate covertly, Alice and Bob use a fraction $\zeta = \mathcal{O}\left(\sqrt{Q/n}\right)$ of n/Q available PPM frames on average, effectively using $\bar{n} = \mathcal{O}(1/\sqrt{n})$ photons per mode. By keeping secret which frames they use, Alice and Bob force Willie to examine all of them, increasing the likelihood of dark counts. An ECC that is known by Willie ensures reliability. However, the transmitted pulse positions are scrambled within the corresponding PPM frames via an operation resembling one-time pad encryption [6], preventing Willie's exploitation of the ECC's structure for detection (rather than protecting the message content). Theorem 4 demonstrates that, using this scheme, Alice reliably transmits $\mathcal{O}\left(\sqrt{\frac{n}{Q}}\log Q\right)$ covert bits at the cost of pre-sharing $\mathcal{O}\left(\sqrt{\frac{n}{Q}}\log n\right)$ secret bits.

EXPERIMENTAL RESULTS

Objective and design—To demonstrate the square-root law of covert optical communication we realized a proof-of-concept test-bed implementation. Alice and Bob engage in an n-mode communication session consisting of n/Q Q-ary PPM frames, $Q = 32$. As described in the Methods, Alice transmits $\zeta n/Q$ PPM symbols on average, using a first order Reed-Solomon (RS) code for error correction. RS codes perform well on channels dominated by *erasures*, which occur in low received-power scenarios, e.g., covert and deep space communication [33]. We defer the specifics of the generation of the transmitted signal to the Methods. We varied n from 3.2×10^6 to 3.2×10^7 in several communication regimes: "careful Alice" ($\zeta = 0.25\sqrt{Q/n}$), "careless Alice" ($\zeta = 0.03\sqrt[4]{Q/n}$), and "dangerously careless Alice" ($\zeta = 0.003$ and $\zeta = 0.008$). For each (n, ζ) pair we conducted 100 experiments and 10^5 Monte-Carlo simulations, measuring Bob's total number of bits received and Willie's detection error probability.

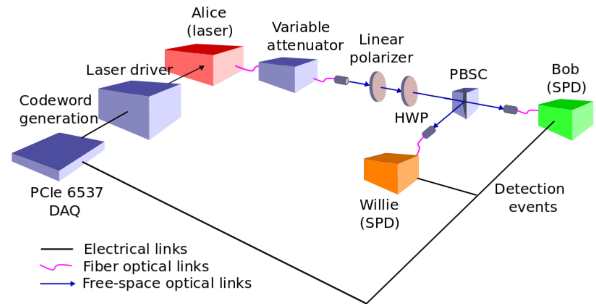

FIG. 2. Experimental setup. A National Instruments PCIe-6537 data acquisition card (DAQ), driven by a 1 MHz clock, controlled the experiment, generating transmissions and reading detection events. Alice generated 1 ns optical pulses using a temperature-stabilized laser diode with center wavelength 1550.2 nm. The pulses were sent into a free-space optical channel, where a half-wave plate (HWP) and polarizing beamsplitter cube (PBSC) sent a fraction η_b of light to Bob, and the remaining light to Willie. Bob and Willie's receivers operated InGaAs Geiger-mode avalanche photodiode SPDs that were gated with 1 ns reverse bias triggered to match the arrival of Alice's pulses.

TABLE I. Optical channel characteristics

	Willie		Bob	
Experimental observations	$p_D^{(w)}$	$\bar{n}_{det}^{(w)}$	$p_D^{(b)}$	$\bar{n}_{det}^{(b)}$
$\zeta = 0.25\sqrt{Q/n}$	9.15×10^{-5}	0.036	2.99×10^{-6}	1.52
$\zeta = 0.03\sqrt[4]{Q/n}$	9.11×10^{-5}	0.032	2.55×10^{-6}	1.14
$\zeta = 0.003$	9.29×10^{-5}	0.032	2.65×10^{-6}	1.07
$\zeta = 0.008$	9.27×10^{-5}	0.028	2.68×10^{-6}	1.05
Target:	9×10^{-5}	0.03	3×10^{-6}	1.4

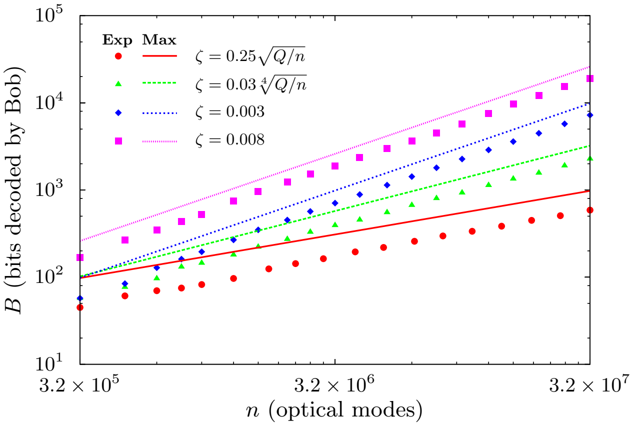

FIG. 3. Bits decoded by Bob. Each data point is an average from 100 experiments, with negligibly small 95% confidence intervals. The symbol error rates are: 1.1×10^{-4} for $\zeta = 0.25\sqrt{Q/n}$, 8.3×10^{-3} for $\zeta = 0.03/\sqrt[4]{Q/n}$, 4.5×10^{-3} for $\zeta = 0.003$, and 1.8×10^{-2} for $\zeta = 0.008$. We also report the maximum throughput $\frac{C_s\zeta n}{Q}$ computed in the Methods using the experimentally-observed values from Table I, where C_s is the per-symbol Shannon capacity [34]. Given the low observed symbol error rate for $\zeta = 0.25\sqrt{Q/n}$, we note that a square root scaling is achievable even using a relatively short RS code; Figure 4 demonstrates that this is achieved covertly.

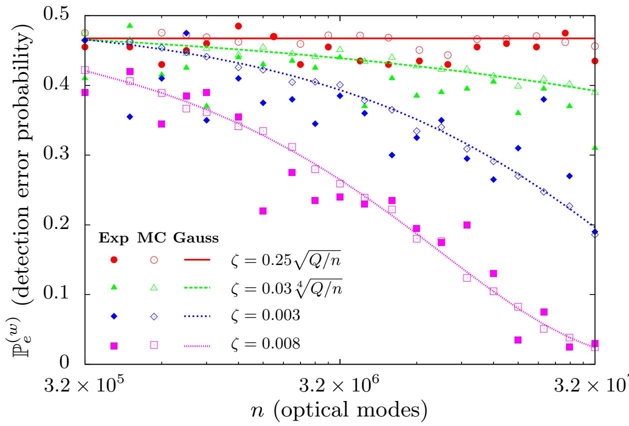

FIG. 4. Willie's error probability. Estimates from 100 experiments have solid fill; estimates from 10^5 Monte-Carlo simulations have clear fill; and Gaussian approximations are lines. The 95% confidence intervals (computed in the Methods) for the experimental estimates are ±0.136; for the Monte-Carlo simulations they are ±0.014. Alice transmits $\zeta n/Q$ PPM symbols on average and Willie's error probability remains constant when Alice obeys the square root law and uses $\zeta = \mathcal{O}(\sqrt{Q/n})$; it drops as n increases if Alice breaks the square root law by using an asymptotically larger ζ.

Implementation—The experiment was conducted using a mixture of fiber-based and free-space optical elements implementing channels from Alice to both Bob and Willie (see Figure 2 for a schematic). Due to the low intensity of Alice's pulses, direct detection using single photon detectors (SPDs), rather than PNR receivers, was sufficient. Several configurations were considered for implementing the background noise at the receivers. We provided noise only during the gating period of the detectors since continuous wave light irradiating Geiger-mode avalanche photodiodes (APDs) suppresses detection efficiency [35]. Instead of providing extraneous optical pulses during the gating window of the APD, we emulated optical noise at the detectors by increasing the detector gate voltage, thus increasing the detector's dark click probability. While the APD dark counts are Poisson-distributed with mean rate $\bar{n}_N$ photons per mode, when $\bar{n}_N \ll 1$, the dark click probability $1-e^{-\bar{n}_N}$ is close to $\frac{\bar{n}_N}{1+\bar{n}_N}$, the probability that an incoherent thermal background with mean photon number per mode $\bar{n}_N$ produces a click. In Table I we report the experimentally-observed and targeted values of dark click probabilities $p_D^{(b)}$ and $p_D^{(w)}$ of Bob's and Willie's detectors, as well as the mean number of photons detected by Bob $\bar{n}_{det}^{(b)} = \eta_b\eta_{\mathrm{QE}}^{(b)}\bar{n}$ and Willie $\bar{n}_{det}^{(w)} = (1-\eta_b)\eta_{\mathrm{QE}}^{(w)}\bar{n}$, where $\bar{n} = 5$ is the mean photon number of Alice's pulses, $\eta_b = 0.97$ is the fraction of light sent to Bob, and $\eta_{\mathrm{QE}}^{(b)}$ and $\eta_{\mathrm{QE}}^{(w)}$ are the quantum efficiencies of Bob's and Willie's detectors, which we do not explicitly calculate. However, quantum efficiency is strongly correlated with the detector's dark click probability [36].

The amount of transmitted information, with other parameters fixed, is proportional to $\bar{n}_{det}^{(b)}/\bar{n}_{det}^{(w)}$. Our choice of $\bar{n}_{det}^{(b)} \gg \bar{n}_{det}^{(w)}$ allowed the experiment to gather a statistically meaningful data sample in a reasonable duration. In an operational free-space laser communication system, a directional transmitter will likely yield just such an asymmetry in coupling between Bob and Willie; however, we note that the only fundamental requirement for implementing information-theoretically secure covert communication is $p_D^{(w)} > 0$, or $\bar{n}_T > 0$.

Results—Alice and Bob use a $(31, 15)$ RS code. Figure 3 reports the number of bits received by Bob with the corresponding symbol error rate in our experiments, and his maximum throughput from Alice (calculated for each regime using the experimentally-observed values from Table I). The details of our analysis are in the Methods. Our relatively short RS code achieves between 45% and 60% of the maximum throughput in the "careful Alice" regime and between 55% and 75% of the maximum in other regimes at reasonable error rates, showing that even a basic code demonstrates our theoretical scaling.

Willie's detection problem can be reduced to a test between two simple hypotheses where the log-likelihood ratio test minimizes $\mathbb{P}_e^{(w)}$ [37]. Figure 4 reports Willie's probability of error estimated from the experiments and the Monte-Carlo study, as well as its analytical Gaussian approximation, with the implementation details deferred to the Methods. Monte-Carlo simulations show that the Gaussian approximation is accurate. More importantly, Figure 4 highlights Alice's safety when she obeys the square root law and her peril when she does not. When $\zeta = \mathcal{O}(1/\sqrt{n})$, $\mathbb{P}_e^{(w)}$ remains constant as n increases. However, for asymptotically larger ζ, $\mathbb{P}_e^{(w)}$ drops at a rate that depends on Alice's carelessness. The drop at $\zeta = 0.008$ vividly demonstrates our converse.

DISCUSSION

We determined that covert communication is achievable provided that the adversary's measurement is subject to non-adversarial excess noise. Excess noise is crucial, as pure loss alone does not allow covert communication, starkly contrasting the QKD scenario. However, the existence of excess noise in practical systems (e.g., blackbody radiation and dark counts) allows covert communication, as demonstrated for the first time in our proof-of-concept optical covert communication system. Even though our results are for an optical channel, they are relevant to RF communication due to the recent advances in quantum-noise-limited microwave-frequency amplifiers and detectors [38]. Finally, our work provides a significant impetus towards the development of covert optical *networks*, eventually scaling privacy to large interconnected systems.

METHODS

Covert Optical Communication Theorems

Here we state our theorems, with proofs deferred to the Supplementary Information. Each theorem can be classified as either an "achievability" or a "converse". Achievability theorems (2, 3, and 4) establish the lower limit on the amount of information that can be covertly transmitted from Alice to Bob, while the converse theorems (1 and 5) demonstrate the upper limit. In essence, the achievability results are obtained by

1. fixing Alice's and Bob's communication system, revealing its construction in entirety (except the shared secret) to Willie;
2. showing that, even with such information, any detector Willie can choose within some natural constraints is ineffective at discriminating Alice's transmission state; and
3. demonstrating that the transmission can be reliably decoded by Bob using the shared secret.

On the other hand, converses are established by

1. fixing Willie's detection scheme (and revealing it to Alice and Bob); and
2. demonstrating that no amount of resources allows Alice to both remain undetected by Willie and exceed the upper limit on the amount of information that is reliably transmitted to Bob.

We start by claiming the inability to instantiate covert communication in the absence of excess noise.

Theorem 1 (Insufficiency of pure-loss for covert communication) *Suppose Willie has a pure-loss channel from Alice and is limited only by the laws of physics in his receiver measurement choice. Then Alice cannot communicate to Bob reliably and covertly even if Alice and Bob have access to a pre-shared secret of unbounded size, an unattenuated observation of the transmission, and a quantum-optimal receiver.*

Next we claim the achievability of the square root law when Willie's channel is subject to excess noise. We first consider a lossy optical channel with additive thermal noise, and claim achievability even when Willie has arbitrary resources such as any quantum-limited measurement on the *isometric extension* of the Alice-to-Bob quantum channel (i.e., access to all signaling photons not captured by Bob).

Theorem 2 (Square root law for the thermal noise channel) *Suppose Willie has access to an arbitrarily complex receiver measurement as permitted by the laws of quantum physics and can capture all the photons transmitted by Alice that do not reach Bob. Let*

Willie's channel from Alice be subject to noise from a thermal environment that injects $\bar{n}_T > 0$ photons per optical mode on average, and let Alice and Bob share a secret of sufficient length before communicating. Then Alice can lower-bound Willie's detection error probability $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$ while reliably transmitting $\mathcal{O}(\sqrt{n})$ bits to Bob in n optical modes even if Bob only has access to a (sub-optimal) coherent detection receiver, such as an optical homodyne detector.

In the remaining theorems Willie's detector is a noisy photon number resolving (PNR) receiver. An ideal PNR receiver is an asymptotically optimal detector for Willie in the pure-loss regime (as discussed in the remark following the proof of Theorem 1 in the Supplementary Information). However, any practical implementation of a PNR receiver has a non-zero dark current. Theorems 3 and 4 show that noise from the resulting dark counts enables covert communication even over a pure-loss channel. We model the dark counts per mode in Willie's PNR detector as a Poisson process with average number of dark counts per mode λ_w.

Theorem 3 (Dark counts yield square root law) *Suppose that Willie has a pure-loss channel from Alice, captures all photons transmitted by Alice that do not reach Bob, but is limited to a receiver with a non-zero dark current. Let Alice and Bob share a secret of sufficient length before communicating. Then Alice can lower-bound Willie's detection error probability $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$ while reliably transmitting $\mathcal{O}(\sqrt{n})$ bits to Bob in n optical modes.*

The proof of Theorem 3 demonstrates that $\mathcal{O}(\sqrt{n})$ covert bits can be reliably transmitted using be on-off keying (OOK) coherent state modulation where Alice transmits the *on* symbol $|\alpha\rangle$ with probability $q = \mathcal{O}(1/\sqrt{n})$ and the *off* symbol $|0\rangle$ with probability $1-q$. However, the skewed on-off duty cycle of OOK modulation makes construction of efficient error correction codes (ECCs) challenging. We thus consider *pulse position modulation* (PPM) which constrains the OOK signaling scheme, enabling the use of many efficient ECCs by sacrificing a constant fraction of throughput. Each PPM symbol uses a PPM *frame* to transmit a sequence of Q coherent state pulses, $|0\rangle \ldots |\alpha\rangle \ldots |0\rangle$, encoding message $i \in \{1, 2, \ldots, Q\}$ by transmitting $|\alpha\rangle$ in the i^{th} mode of the PPM frame. Next we claim that the square root scaling is achievable under this structural constraint.

Theorem 4 (Dark counts yield square root law under structured modulation) *Suppose that Willie has a pure-loss channel from Alice, can capture all photons transmitted by Alice that do not reach Bob, but is limited to a PNR receiver with a non-zero dark current. Let Alice and Bob share a secret of sufficient length before communicating. Then Alice can lower-bound Willie's detection error probability $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$ while reliably transmitting $\mathcal{O}(\sqrt{\frac{n}{Q}} \log Q)$ bits to Bob using n optical modes and a Q-ary PPM constellation.*

Finally, we claim the unsurmountability of the square root law. We assume non-zero thermal noise ($\bar{n}_T > 0$) in the channel and non-zero dark count rate ($\lambda_w > 0$) in Willie's detector. Setting $\lambda_w = 0$ yields the converse for Theorem 2, and setting $\bar{n}_T = 0$ yields the converse for Theorems 3 and 4 Setting $\lambda_w = 0$ and $\bar{n}_T = 0$ yields the conditions for Theorem 1. To state the theorem, we use the following asymptotic notation [39]: we say $f(n) = \omega(g(n))$ when $g(n)$ is a lower bound that is not asymptotically tight.

Theorem 5 (Converse of the square root law) *Suppose Alice only uses n-mode codewords with total photon number variance $\sigma_x^2 = \mathcal{O}(n)$. Then, if she attempts to transmit $\omega(\sqrt{n})$ bits in n modes, as $n \to \infty$, she is either detected by Willie with arbitrarily low detection error probability, or Bob cannot decode with arbitrarily low decoding error probability.*

The restriction on the photon number variance of Alice's input states is not onerous since it subsumes all well-known quantum states of a bosonic mode. However, proving this theorem for input states with unbounded photon number variance per mode remains an open problem.

Next we provide details of the experimental methodology.

Experimental methodology

Alice's encoder

Prior to communication, Alice and Bob secretly select a random subset $\mathcal{S}$ of PPM frames to use for transmission: each of the n/Q available PPM frames is selected independently with probability ζ. Alice and Bob then secretly generate a vector $\mathbf{k}$ containing $|\mathcal{S}|$ numbers selected independently uniformly at random from $\{0, 1, \ldots, Q-1\}$, where $|\mathcal{S}|$ denotes the cardinality of $\mathcal{S}$. Alice encodes a message into a codeword of size $|\mathcal{S}|$ using a Reed-Solomon (RS) code. She adds $\mathbf{k}$ modulo Q to this message and transmits it on the PPM frames in $\mathcal{S}$. We note that this is almost identical to the construction of the coding scheme in the proof of Theorem 4 (see the Supplementary Information), with the exception of the use of an RS code for error correction.

Generation of transmitted symbols

Alice generates the length-n binary sequence describing the transmitted signal, with a "1" at a given location indicating a pulse in that mode, and a "0" indicating the absence of a pulse. First, Alice encodes random data, organized into Q-ary symbols, with an RS code and modulo-Q addition of $\mathbf{k}$ to produce a coded sequence of Q-ary symbols. The value of the i^{th} symbol in this sequence indicates which mode in the i^{th} PPM symbol in the set $\mathcal{S}$ contains a pulse, whereas all modes of the PPM frames not in $\mathcal{S}$ remain empty. Mapping occupied modes to "1" and unoccupied modes to "0" results in the desired length-n binary sequence.

To accurately estimate Willie's detection error probability in the face of optical power fluctuations, the length-n binary sequence from above is alternated with a length-n sequence of all "0"s, to produce the final length-$2n$ sequence that is passed to the experimental setup. Willie gets a "clean" look at the channel when Alice is silent using these interleaved "0"s, thus allowing the estimation of both the false alarm and the missed detection probabilities under the same conditions. Bob simply discards the interleaved "0"s.

Bob's decoder

Bob examines only the PPM frames in $\mathcal{S}$. If two or more pulses are detected in a PPM frame, one of them is selected uniformly at random. If no pulses are detected, it is labeled as an *erasure*. After subtracting $\mathbf{k}$ modulo Q from this vector of PPM symbols (subtraction is not performed on erasures), the resultant vector is passed to the RS decoder.

For each experiment we record the total number of bits in the successfully-decoded codewords; the undecoded codewords are discarded. For each pair of parameters (ζ, n) we report the mean of the total number of decoded bits over 100 experiments. The reported symbol error rate is the total number of lost data symbols during all the experiments at the specified communication regime divided by the total number of data symbols transmitted during these experiments. The calculation of the theoretical channel capacity is presented in the Supplementary Information.

Willie's detector

Estimation of $\mathbb{P}_e^{(w)}$—The test statistic for the log-likelihood ratio test is defined as:

$$L = \log \frac{f_1(\mathbf{x}_w)}{f_0(\mathbf{x}_w)} = \sum_{i=1}^{n/Q} \log \left[1 + \zeta p_r^{(w)} \left(\frac{y_i^{(w)}}{Q p_D^{(w)}} - 1 \right) \right], \tag{1}$$

where $f_0(\mathbf{x}_w)$ and $f_1(\mathbf{x}_w)$ are the likelihood functions of the click record $\mathbf{x}_w$ corresponding to Alice being quiet and transmitting, $y_i^{(w)}$ is the number of clicks Willie observes in the i^{th} PPM frame, and $p_r^{(w)} = 1 - e^{-\eta_w \bar{n}}$ is the probability of Willie observing a click stemming from Alice's transmission. Equation (1) is derived in the Supplementary Information. Willie calculates L using equation (1) and compares it to a threshold S, accusing Alice if $L \geq S$. Willie chooses the value of S that minimizes Willie's detection error probability $\mathbb{P}_e^{(w)}$.

For each pair of parameters (n, ζ) as well as Alice's transmission state, we perform m experiments, obtaining a sample vector $\mathbf{y}_w$ from each experiment and calculating the log-likelihood ratio L using (1). We denote by $\mathbf{L}^{(0)} = [L_1^{(0)}, \ldots, L_m^{(0)}]$ and $\mathbf{L}^{(1)} = [L_1^{(1)}, \ldots, L_m^{(1)}]$ the vectors of experimentally observed log-likelihood ratios when Alice does not transmit and transmits, respectively. To estimate Willie's probability of error $\mathbb{P}_e^{(w)}$, we construct empirical distribution functions $\hat{F}_m^{(0)}(x) = \frac{1}{n}\sum_{i=1}^m \mathbf{1}_{L_i^{(0)} \leq x}(x)$ and $\hat{F}_m^{(1)}(x) = \frac{1}{m}\sum_{i=1}^m \mathbf{1}_{L_i^{(1)} \leq x}(x)$, where $\mathbf{1}_{\mathcal{A}}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{A} \\ 0 & \text{if } x \notin \mathcal{A} \end{cases}$ denotes the indicator function. The estimated probability of error is then

$$\hat{\mathbb{P}}_e^{(w)} = \frac{1}{2} \min_S (1 - \hat{F}_m^{(0)}(S) + \hat{F}_m^{(1)}(S)). \tag{2}$$

Monte-Carlo simulation and Gaussian approximation—We perform a Monte-Carlo study using 10^5 simulations per (n, ζ) pair. We generate, encode, and detect the messages as in the physical experiment, and use equation (2) to estimate Willie's probability of error, but simulate the optical channel induced by our choice of a laser-light transmitter and an SPD using its measured characteristics reported in Table I. Similarly, we use the values in Table I for our analytical Gaussian approximation of $\mathbb{P}_e^{(w)}$ described in the Supplementary Information.

Confidence intervals—We compute the confidence intervals for the estimate in equation (2) using Dvoretzky-Keifer-Wolfowitz inequality [40, 41], which relates the distribution function $F_X(x)$ of random variable X to the empirical distribution function $\hat{F}_m(x) = \frac{1}{m}\sum_{i=1}^{m}\mathbf{1}_{X_i\leq x}(x)$ associated with a sequence $\{X_i\}_{i=1}^m$ of m i.i.d. draws of the random variable X as follows:

$$\mathbb{P}(\sup_x |\hat{F}_m(x) - F_X(x)| > \xi) \leq 2e^{-2m\xi^2}, \quad (3)$$

where $\xi > 0$. For x_0, the $(1-\alpha)$ confidence interval for the empirical estimate of $F(x_0)$ is given by $[\max\{\hat{F}_m(x_0)-\xi, 0\}, \min\{\hat{F}_m(x_0)+\xi, 1\}]$ where $\xi = \sqrt{\frac{\log(2/\alpha)}{2m}}$. Thus, $\pm\xi$ is used for reporting the confidence intervals in Figure 4.

[1] Alfred J. Menezes, Scott A. Vanstone, and Paul C. Van Oorschot, *Handbook of Applied Cryptography*, 1st ed. (CRC Press, Inc., Boca Raton, FL, USA, 1996).

[2] J. Talbot and D.J.A. Welsh, *Complexity and Cryptography: An Introduction* (Cambridge University Press, 2006).

[3] A. D. Wyner, "The wiretap channel," Bell System Technical Journal **54**, 1355 (1975).

[4] Imre Csiszár and János Körner, "Broadcast channels with confidential messages," Information Theory, IEEE Transactions on **24**, 339–348 (May 1978).

[5] C. H. Bennett and G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing," in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Press, New York, 1984) pp. 175–179.

[6] Claude E. Shannon, "Communication theory of security," Bell System Technical Journal **28**, 656–715 (1949).

[7] BBC, "Edward Snowden: Leaks that exposed US spy programme," `http://www.bbc.com/news/world-us-canada-23123964` (Jan. 2014).

[8] Herodotus (c. 440 BC) 5.35 and 7.239.

[9] Marvin K. Simon, Jim K. Omura, Robert A. Scholtz, and Barry K. Levitt, *Spread Spectrum Communications Handbook*, revised ed. (McGraw-Hill, 1994).

[10] Jessica Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*, 1st ed. (Cambridge University Press, New York, NY, USA, 2009).

[11] Boulat A. Bash, Dennis Goeckel, and Don Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," IEEE Journal on Selected Areas in Communications **31**, 1921–1930 (2013), arXiv:1202.6423.

[12] Boulat A. Bash, Dennis Goeckel, and Don Towsley, "Square root law for communication with low probability of detection on AWGN channels," in *Proc. of IEEE International Symposium on Information Theory (ISIT)* (Cambridge, MA, USA, 2012).

[13] Boulat A. Bash, Dennis Goeckel, and Don Towsley, "LPD Communication when the Warden Does Not Know When," in *Proc. of IEEE International Symposium on Information Theory (ISIT)* (Honolulu, HI, USA, 2014) arXiv:1403.1013.

[14] Pak Hou Che, Mayank Bakshi, and Sidharth Jaggi, "Reliable deniable communication: Hiding messages in noise," in *Proc. of IEEE International Symposium on Information Theory (ISIT)* (Istanbul, Turkey, 2013) arXiv:1304.6693.

[15] Swanand Kadhe, Sidharth Jaggi, Mayank Bakshi, and Alex Sprintson, "Reliable, deniable, and hidable communication over multipath networks," in *Proc. of IEEE International Symposium on Information Theory (ISIT)* (Honolulu, HI, USA, 2014) arXiv:1401.4451.

[16] Jie Hou and Gerhard Kramer, "Effective secrecy: Reliability, confusion and stealth," in *Proc. of IEEE International Symposium on Information Theory (ISIT)* (Honolulu, HI, USA, 2014) arXiv:1311.1411.

[17] The $\log n$ improvement in steganographic application versus covert communication over a noisy channel is attributable to the noiseless Alice-to-Bob channel, and the similarity in the square root laws is due to the mathematics of classical [37] and quantum [42] statistical hypothesis testing.

[18] Andrew D. Ker, "Batch steganography and pooled steganalysis," (Springer Berlin Heidelberg, 2007) pp. 265–281.

[19] Tomáš Filler, Andrew D. Ker, and Jessica Fridrich, "The square root law of steganographic capacity for markov covers," Media Forensics and Security **7254** (2009).

[20] Andrew D. Ker, "The square root law requires a linear key," in *Proceedings of the 11th ACM workshop on Multimedia and Security*, MM&Sec '09 (Princeton, NJ, USA, 2009) pp. 85–92.

[21] Andrew D. Ker, "The square root law does not require a linear key," in *Proceedings of the 12th ACM*

workshop on Multimedia and Security, MM&Sec '10 (Rome, Italy, 2010) pp. 213–224.

[22] Bilal A. Shaw and Todd A. Brun, "Quantum steganography with noisy quantum channels," Phys. Rev. A **83**, 022310 (Feb 2011).

[23] O. Bouchet, H. Sizun, C. Boisrobert, F. de Fornel, and P.N. Favennec, *Free-Space Optics: Propagation and Communication* (Wiley, 2010).

[24] John Senior, *Optical Fiber Communications*, 3rd ed. (Pearson Education, 2009).

[25] R.M. Gagliardi and S. Karp, *Optical Communications*, 2nd ed. (Wiley, 1995).

[26] J.W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts & Company, 2005).

[27] Duwayne R. Anderson, Larry M. Johnson, and Florian G. Bell, *Troubleshooting Optical Fiber Networks: Understanding and Using Optical Time-Domain Reflectometers*, 2nd ed. (Elsevier Academic Press, 2004).

[28] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, "Classical capacity of the lossy bosonic channel: The exact solution," Phys. Rev. Lett. **92**, 027902 (Jan 2004).

[29] Michael M. Wolf, David Pérez-García, and Geza Giedke, "Quantum capacities of bosonic channels," Phys. Rev. Lett. **98**, 130501 (Mar 2007).

[30] Mark M. Wilde, Patrick Hayden, and Saikat Guha, "Information trade-offs for optical quantum communication," Phys. Rev. Lett. **108**, 140501 (Apr 2012).

[31] N.S. Kopeika and J. Bordogna, "Background noise in optical communication systems," Proc. of the IEEE **58**, 1571–1577 (Oct. 1970).

[32] Ligong Wang and Gregory W Wornell, "Refined analysis of the poisson channel in the high-photon-efficiency regime," in *Proc. of IEEE Information Theory Workshop (ITW)* (Lausanne, Switzerland, 2012) pp. 582–586.

[33] Bruce Moision, Jon Hamkins, and Michael Cheng, "Design of a coded modulation for deep space optical communications," in *Information Theory and its Applications (ITA) Workshop* (University of California San Diego, 2006).

[34] Claude E. Shannon, "A mathematical theory of communication," Bell System Technical Journal **27** (1948).

[35] Vadim Makarov, "Controlling passively quenched single photon detectors by bright light," New J. Phys. **11**, 065003 (2009).

[36] Gregoire Ribordy, Jean-Daniel Gautier, Hugo Zbinden, and Nicolas Gisin, "Performance of InGaAs/InP avalanche photodiodes as gated-mode photon counters," App. Opt. **37**, 2272–2277 (1998).

[37] Erich Lehmann and Joseph Romano, *Testing Statistical Hypotheses*, 3rd ed. (Springer, New York, NY, USA, 2005).

[38] Baleegh Abdo, Katrina Sliwa, S. Shankar, Michael Hatridge, Luigi Frunzio, Robert Schoelkopf, and Michel Devoret, "Josephson directional amplifier for quantum measurement of superconducting circuits," Phys. Rev. Lett. **112**, 167701 (Apr 2014).

[39] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, *Introduction to Algorithms*, 2nd ed. (MIT Press, Cambridge, Massachusetts, 2001).

[40] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator," Ann. Math. Statist. **27**, 642–669 (1956).

[41] P. Massart, "The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality," The Annals of Probability **18**, 1269–1283 (07 1990).

[42] Carl W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, Inc., New York, NY, USA, 1976).

[43] Robert G. Gallager, *Information Theory and Reliable Communication* (John Wiley and Sons, Inc., New York, 1968).

[44] Richard A. Campos, Bahaa E. A. Saleh, and Malvin C. Teich, "Quantum-mechanical lossless beam splitter: SU(2) symmetry and photon statistics," Phys. Rev. A **40**, 1371–1384 (Aug 1989).

[45] M.M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).

[46] Stefano Pirandola and Seth Lloyd, "Computable bounds for the discrimination of gaussian states," Phys. Rev. A **78**, 012331 (Jul 2008).

[47] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, 2nd ed. (John Wiley & Sons, Hoboken, NJ, USA, 2002).

[48] We use [43, Theorem 5.6.2], setting parameter $s = 1$.

[49] A. S. Holevo, "The capacity of the quantum channel with general signal states," IEEE Trans. Inf. Theory **44**, 269–273 (1998).

SUPPLEMENTARY INFORMATION

I. COVERT OPTICAL COMMUNICATION THEOREMS

Here we re-state the theorems from the main paper and provide their proofs.

Theorem 1 (Insufficiency of pure-loss for covert communication) *Suppose Willie has a pure-loss channel from Alice and is limited only by the laws of physics in his receiver measurement choice. Then Alice cannot communicate to Bob reliably and covertly even if Alice and Bob have access to a pre-shared secret of unbounded size, an unattenuated observation of the transmission, and a quantum-optimal receiver.*

In the proof of this theorem we denote a tensor product of n Fock (or photon number) states by $|\mathbf{u}\rangle \equiv |u_1\rangle \otimes |u_2\rangle \otimes \cdots \otimes |u_n\rangle$, where vector $\mathbf{u} \in \mathbb{N}_0^n$ and $\mathbb{N}_0$ is the set of non-negative integers. Specifically, $|\mathbf{0}\rangle \equiv |0\rangle^{\otimes n}$. Before proceeding with the proof, we prove the following lemma:

Lemma 2 *Given the input of n-mode vacuum state $|\mathbf{0}\rangle^{E^n}$ on the "environment" port and an n-mode entangled state $|\psi\rangle^{A^n} = \sum_{\mathbf{k}} a_{\mathbf{k}} |\mathbf{k}\rangle^{A^n}$ on the "Alice" port of a beamsplitter with transmissivity $\eta_b = 1 - \eta_w$, the diagonal elements of the output state ρ^{W^n} on the "Willie" port can be expressed in the n-fold Fock state basis as follows:*

$$ {}^{W^n}\langle \mathbf{s}| \hat{\rho}^{W^n} |\mathbf{s}\rangle^{W^n} = \sum_{\mathbf{k}\in\mathbb{N}_0^n} |a_{\mathbf{k}}|^2 \prod_{i=1}^{n} \binom{k_i}{s_i} (1-\eta_w)^{k_i - s_i} \eta_w^{s_i}. \tag{S1} $$

Proof. A beamsplitter can be described as a unitary transformation U_{BS} from the two input modes (Alice's and the environment's ports) to the two output modes (Bob's and Willie's ports). Given a Fock state input $|t\rangle^A$ on Alice's port and vacuum input $|0\rangle^E$ on the environment's port, the output at Bob's and Willie's ports is described as follows [44, Section IV.D]:

$$ U_{BS} |t\rangle^A |0\rangle^E = \sum_{m=0}^{t} \sqrt{\binom{t}{m} \eta_w^m (1-\eta_w)^{t-m}}\, |m\rangle^W |t-m\rangle^B . $$

Thus,

$$ U_{BS}^{\otimes n} |\mathbf{t}\rangle^{A^n} |\mathbf{0}\rangle^{E^n} = \bigotimes_{i=1}^{n} \sum_{m_i=0}^{t_i} \sqrt{\binom{t_i}{m_i} \eta_w^{m_i} (1-\eta_w)^{t_i-m_i}}\, |m_i\rangle^{W_i} |t_i - m_i\rangle^{B_i} , $$

which implies

$$ U_{BS}^{\otimes n} |\psi\rangle^{A^n} |\mathbf{0}\rangle^{E^n} = \sum_{\mathbf{t}\in\mathbb{N}_0^n} a_{\mathbf{t}} \bigotimes_{i=1}^{n} \sum_{m_i=0}^{t_i} \sqrt{\binom{t_i}{m_i} \eta_w^{m_i} (1-\eta_w)^{t_i-m_i}}\, |m_i\rangle^{W_i} |t_i - m_i\rangle^{B_i} \equiv |\phi\rangle^{W^n B^n} . $$

Now, the partial trace of the output state $\rho^{BW} = |\phi\rangle^{W^n B^n}$ over Bob's system reveals Willie's output state:

$$ \begin{aligned} \rho^{W^n} &= \mathrm{Tr}_{B^n}\left[|\phi\rangle^{W^n B^n}\, {}^{W^n B^n}\langle\phi| \right] \\ &= \sum_{\mathbf{x}\in\mathbb{N}_0^n} \left| {}^{B^n}\langle \mathbf{x}|\phi\rangle^{W^n B^n} \right|^2 , \end{aligned} $$

with

$$ \begin{aligned} {}^{B^n}\langle \mathbf{x}|\phi\rangle^{W^n B^n} &= \sum_{\mathbf{t}\in\mathbb{N}_0^n} a_{\mathbf{t}} \bigotimes_{i=1}^{n} \sum_{m_i=0}^{t_i} \sqrt{\binom{t_i}{m_i} \eta_w^{m_i} (1-\eta_w)^{t_i-m_i}}\, |m_i\rangle^{W_i}\, {}^{B_i}\langle x_i | t_i - m_i\rangle^{B_i} \\ &= \sum_{\mathbf{t}\in\mathbb{N}_0^n} a_{\mathbf{t}} \bigotimes_{i=1}^{n} \sqrt{\binom{t_i}{x_i} \eta_w^{t_i - x_i} (1-\eta_w)^{x_i}}\, |t_i - x_i\rangle^{W_i} , \end{aligned} \tag{S2} $$

where equation (S2) is due to the orthogonality of the Fock states. Thus,

$$ {}^{W^n}\!\langle \mathbf{s}|\, \hat{\rho}^{W^n}\, |\mathbf{s}\rangle^{W^n} = \sum_{\mathbf{x}\in\mathbb{N}_0^n} \left| {}^{W^n}\!\langle \mathbf{s}|\; {}^{B^n}\!\langle \mathbf{x}|\phi\rangle^{W^n B^n} \right|^2, \tag{S3} $$

where

$$ \begin{aligned} {}^{W^n}\!\langle \mathbf{s}|\; {}^{B^n}\!\langle \mathbf{x}|\phi\rangle^{W^n B^n} &= \sum_{\mathbf{t}\in\mathbb{N}_0^n} a_{\mathbf{t}} \prod_{i=1}^n \sqrt{\binom{t_i}{x_i} \eta_w^{t_i - x_i} (1-\eta_w)^{x_i}}\, \delta_{s_i, t_i - x_i} \\ &= a_{\mathbf{x}+\mathbf{s}} \prod_{i=1}^n \sqrt{\binom{x_i + s_i}{x_i} \eta_w^{s_i} (1-\eta_w)^{x_i}}, \end{aligned} \tag{S4} $$

with $\delta_{a,b} = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}$. Substituting $\mathbf{k} = \mathbf{x} + \mathbf{s}$ into equation (S4) and substituting the right-hand side (RHS) of (S4) into equation (S3) yields

$$ \begin{aligned} {}^{W^n}\!\langle \mathbf{s}|\, \hat{\rho}^{W^n}\, |\mathbf{s}\rangle^{W^n} &= \sum_{\mathbf{k}\in\mathbb{N}_0^n} \left| a_{\mathbf{k}} \prod_{i=1}^n \sqrt{\binom{k_i}{s_i} \eta_w^{s_i} (1-\eta_w)^{k_i - s_i}} \right|^2 \\ &= \sum_{\mathbf{k}\in\mathbb{N}_0^n} |a_{\mathbf{k}}|^2 \prod_{i=1}^n \binom{k_i}{s_i} \eta_w^{s_i} (1-\eta_w)^{k_i - s_i} \end{aligned} \tag{S5} $$

where equation (S5) is due to $\eta_w \in [0, 1)$. ■

Proof. *(Theorem 1)* Alice sends one of 2^M (equally likely) M-bit messages by choosing an element from an arbitrary codebook $\{\hat{\rho}_x^{A^n}, x = 1, \ldots, 2^M\}$, where a state $\hat{\rho}_x^{A^n} = |\psi_x\rangle^{A^n A^n}\!\langle \psi_x|$ encodes an M-bit message W_x. $|\psi_x\rangle^{A^n} = \sum_{\mathbf{k}\in\mathbb{N}_0^n} a_{\mathbf{k}}(x) |\mathbf{k}\rangle^{A^n}$ is a general n-mode pure state, where $|\mathbf{k}\rangle \equiv |k_1\rangle \otimes |k_2\rangle \otimes \cdots \otimes |k_n\rangle$ is a tensor product of n Fock states. We limit our analysis to pure input states since, by convexity, using mixed states as inputs can only degrade the performance (since that is equivalent to transmitting a randomly chosen pure state from an ensemble and discarding the knowledge of that choice).

Let Willie use an ideal SPD on all n modes, given by positive operator-valued measure (POVM) $\left\{ |0\rangle\langle 0|, \sum_{j=1}^{\infty} |j\rangle\langle j| \right\}^{\otimes n}$. When W_u is transmitted, Willie's hypothesis test reduces to discriminating between the states

$$ \hat{\rho}_0^{W^n} = |\mathbf{0}\rangle^{W^n W^n}\!\langle \mathbf{0}| \text{ and} \tag{S6} $$

$$ \hat{\rho}_1^{W^n} = \hat{\rho}_u^{W^n}, \tag{S7} $$

where $\hat{\rho}_u^{W^n}$ is the output state of a pure-loss channel with transmissivity η_w corresponding to an input state $\hat{\rho}_u^{A^n}$. Thus, Willie's average error probability is:

$$ \mathbb{P}_e^{(w)} = \frac{1}{2^{M+1}} \sum_{u=1}^{2^M} {}^{W^n}\!\langle \mathbf{0}|\, \hat{\rho}_u^{W^n}\, |\mathbf{0}\rangle^{W^n}, \tag{S8} $$

since messages are sent equiprobably. Note that the error is entirely due to missed codeword detections, as Willie's receiver never raises a false alarm. By Lemma 2,

$$ \begin{aligned} {}^{W^n}\!\langle \mathbf{0}|\, \hat{\rho}_u^{W}\, |\mathbf{0}\rangle^{W^n} &= \sum_{\mathbf{k}\in\mathbb{N}_0^n} |a_{\mathbf{k}}(u)|^2 (1-\eta_w)^{\sum_{i=1}^n k_i} \\ &\leq |a_{\mathbf{0}}(u)|^2 + (1 - |a_{\mathbf{0}}(u)|^2)(1-\eta_w) \\ &= 1 - \eta_w \left(1 - |a_{\mathbf{0}}(u)|^2\right). \end{aligned} \tag{S9} $$

Substituting equation (S9) into equation (S8) yields:

$$\mathbb{P}_e^{(w)} \leq \frac{1}{2} - \frac{\eta_w}{2}\left(1 - \frac{1}{2^M}\sum_{u=1}^{2^M} |a_{\mathbf{0}}(u)|^2\right).$$

Thus, to ensure $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$, Alice must use a codebook with the probability of transmitting zero photons:

$$\frac{1}{2^M}\sum_{u=1}^{2^M} |a_{\mathbf{0}}(u)|^2 \geq 1 - \frac{2\epsilon}{\eta_w}. \tag{S10}$$

Equation (S10) can be restated as an upper bound on the probability of transmitting one or more photons:

$$\frac{1}{2^M}\sum_{u=1}^{2^M} \left(1 - |a_{\mathbf{0}}(u)|^2\right) \leq \frac{2\epsilon}{\eta_w}. \tag{S11}$$

Now we show that there exists an interval $(0, \epsilon_0]$, $\epsilon_0 > 0$ such that if $\epsilon \in (0, \epsilon_0]$, Bob's average decoding error probability $\mathbb{P}_e^{(b)} \geq \delta_0$ where $\delta_0 > 0$, thus making covert communication over a pure-loss channel unreliable.

Denote by $E_{u\to v}$ the event that the transmitted message W_u is decoded by Bob as $W_v \neq W_u$. Given that W_u is transmitted, the decoding error probability is the probability of the union of events $\cup_{v=0, v\neq u}^{2^M} E_{u\to v}$. Let Bob choose a POVM $\{\Lambda_j^*\}$ that minimizes the average probability of error over n optical channel modes:

$$\mathbb{P}_e^{(b)} = \inf_{\{\Lambda_j\}} \frac{1}{2^M}\sum_{u=1}^{2^M} \mathbb{P}\left(\cup_{v=0, v\neq u}^{2^M} E_{u\to v}\right). \tag{S12}$$

Now consider a codebook that meets the necessary condition for covert communication given in equation (S11). Define the subset of this codebook $\left\{\hat{\rho}_u^{A^n}, u \in \mathcal{A}\right\}$ where $\mathcal{A} = \left\{u : 1 - |a_{\mathbf{0}}(u)|^2 \leq \frac{4\epsilon}{\eta_w}\right\}$. We lower-bound (S12) as follows:

$$\begin{aligned}
\mathbb{P}_e^{(b)} &= \frac{1}{2^M}\sum_{u\in\bar{\mathcal{A}}} \mathbb{P}\left(\cup_{v=0, v\neq u}^{2^M} E_{u\to v}\right) + \frac{1}{2^M}\sum_{u\in\mathcal{A}} \mathbb{P}\left(\cup_{v=0, v\neq u}^{2^M} E_{u\to v}\right) && \text{(S13)}\\
&\geq \frac{1}{2^M}\sum_{u\in\mathcal{A}} \mathbb{P}\left(\cup_{v=0, v\neq u}^{2^M} E_{u\to v}\right), && \text{(S14)}
\end{aligned}$$

where the probabilities in equation (S13) are with respect to the POVM $\{\Lambda_j^*\}$ that minimizes equation (S12) over the entire codebook. Without loss of generality, let's assume that $|\mathcal{A}|$ is even, and split $\mathcal{A}$ into two equal-sized subsets $\mathcal{A}^{(\text{left})}$ and $\mathcal{A}^{(\text{right})}$ (formally, $\mathcal{A}^{(\text{left})} \cup \mathcal{A}^{(\text{right})} = \mathcal{A}$, $\mathcal{A}^{(\text{left})} \cap \mathcal{A}^{(\text{right})} = \emptyset$, and $|\mathcal{A}^{(\text{left})}| = |\mathcal{A}^{(\text{right})}|$). Let $g : \mathcal{A}^{(\text{left})} \to \mathcal{A}^{(\text{right})}$ be a bijection. We can thus re-write (S14):

$$\begin{aligned}
\mathbb{P}_e^{(b)} &\geq \frac{1}{2^M}\sum_{u\in\mathcal{A}^{(\text{left})}} 2\left(\frac{\mathbb{P}\left(\cup_{v=0, v\neq u}^{2^M} E_{u\to v}\right)}{2} + \frac{\mathbb{P}\left(\cup_{v=0, v\neq g(u)}^{2^M} E_{g(u)\to v}\right)}{2}\right)\\
&\geq \frac{1}{2^M}\sum_{u\in\mathcal{A}^{(\text{left})}} 2\left(\frac{\mathbb{P}\left(E_{u\to g(u)}\right)}{2} + \frac{\mathbb{P}\left(E_{g(u)\to u}\right)}{2}\right), && \text{(S15)}
\end{aligned}$$

where the second lower bound is due to the events $E_{u\to g(u)}$ and $E_{g(u)\to u}$ being contained in the unions $\cup_{v=0, v\neq u}^{2^M} E_{u\to v}$ and $\cup_{v=0, v\neq g(u)}^{2^M} E_{g(u)\to v}$, respectively. The summation term in equation (S15),

$$\mathbb{P}_e(u) \equiv \frac{\mathbb{P}\left(E_{u\to g(u)}\right)}{2} + \frac{\mathbb{P}\left(E_{g(u)\to u}\right)}{2}, \tag{S16}$$

is Bob's average probability of error when Alice only sends messages W_u and $W_{g(u)}$ equiprobably. We thus reduce the analytically intractable problem of discriminating between many states in equation (S12) to a quantum binary hypothesis test.

The lower bound on the probability of error in discriminating two received codewords is obtained by lower-bounding the probability of error in discriminating two codewords *before* they are sent (this is equivalent to Bob having an unattenuated unity-transmissivity channel from Alice). Recalling that $\hat{\rho}_u^{A^n} = |\psi_u\rangle^{A^n A^n}\langle\psi_u|$ and $\hat{\rho}_{g(u)}^{A^n} = |\psi_{g(u)}\rangle^{A^n A^n}\langle\psi_{g(u)}|$ are pure states, the lower bound on the probability of error in discriminating between $|\psi_u^{A^n}\rangle$ and $|\psi_{g(u)}^{A^n}\rangle$ is [42, Chapter IV.2 (c), Equation (2.34)]:

$$\mathbb{P}_e(u) \geq \left[1 - \sqrt{1 - F\left(|\psi_u\rangle^{A^n}, |\psi_{g(u)}\rangle^{A^n}\right)}\right] \Big/ 2\,, \tag{S17}$$

where $F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$ is the fidelity between the pure states $|\psi\rangle$ and $|\phi\rangle$. Lower-bounding $F\left(|\psi_u\rangle^{A^n}, |\psi_{g(u)}\rangle^{A^n}\right)$ lower-bounds the RHS of equation (S17). For pure states $|\psi\rangle$ and $|\phi\rangle$, $F(|\psi\rangle, |\phi\rangle) = 1 - \left(\frac{1}{2}\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_1\right)^2$, where $\|\rho - \sigma\|_1$ is the trace distance [45, Equation (9.134)]. Thus,

$$\begin{aligned} F\left(|\psi_u\rangle^{A^n}, |\psi_{g(u)}\rangle^{A^n}\right) &= 1 - \left(\frac{1}{2}\|\hat{\rho}_u^{A^n} - \hat{\rho}_{g(u)}^{A^n}\|_1\right)^2 \\ &\geq 1 - \left(\frac{\|\hat{\rho}_u^{A^n} - |\mathbf{0}\rangle^{A^n A^n}\langle\mathbf{0}|\,\|_1}{2} + \frac{\|\hat{\rho}_{g(u)}^{A^n} - |\mathbf{0}\rangle^{A^n A^n}\langle\mathbf{0}|\,\|_1}{2}\right)^2 \\ &= 1 - \left(\sqrt{1 - \left|{}^{A^n}\langle\mathbf{0}|\psi_u\rangle^{A^n}\right|^2} + \sqrt{1 - \left|{}^{A^n}\langle\mathbf{0}|\psi_{g(u)}\rangle^{A^n}\right|^2}\right)^2, \end{aligned} \tag{S18}$$

where the inequality is due to the triangle inequality for trace distance. Substituting (S18) into (S17) yields:

$$\mathbb{P}_e(u) \geq \left[1 - \sqrt{1 - \left|{}^{A^n}\langle\mathbf{0}|\psi_u\rangle^{A^n}\right|^2} - \sqrt{1 - \left|{}^{A^n}\langle\mathbf{0}|\psi_{g(u)}\rangle^{A^n}\right|^2}\right] \Big/ 2\,. \tag{S19}$$

Since $\left|{}^{A^n}\langle\mathbf{0}|\psi_u\rangle^{A^n}\right|^2 = |a_{\mathbf{0}}(u)|^2$ and, by the construction of $\mathcal{A}$, $1 - |a_{\mathbf{0}}(u)|^2 \leq \frac{4\epsilon}{\eta_w}$ and $1 - |a_{\mathbf{0}}(g(u))|^2 \leq \frac{4\epsilon}{\eta_w}$, we have:

$$\mathbb{P}_e(u) \geq \frac{1}{2} - 2\sqrt{\frac{\epsilon}{\eta_w}}. \tag{S20}$$

Recalling the definition of $\mathbb{P}_e(u)$ in equation (S16), we substitute (S20) into (S15) to obtain:

$$\mathbb{P}_e^{(b)} \geq \frac{|\mathcal{A}|}{2^M}\left(\frac{1}{2} - 2\sqrt{\frac{\epsilon}{\eta_w}}\right), \tag{S21}$$

Now, re-stating the condition for covert communication (S11) yields:

$$\begin{aligned} \frac{2\epsilon}{\eta_w} &\geq \frac{1}{2^M}\sum_{u\in\overline{\mathcal{A}}}\left(1 - |a_{\mathbf{0}}(u)|^2\right) \\ &\geq \frac{\left(2^M - |\mathcal{A}|\right)}{2^M}\frac{4\epsilon}{\eta_w} \end{aligned} \tag{S22}$$

with equality (S22) due to $1 - |a_{\mathbf{0}}(u)|^2 > \frac{4\epsilon}{\eta_w}$ for all codewords in $\overline{\mathcal{A}}$ by the construction of $\mathcal{A}$. Solving inequality in (S22) for $\frac{|\mathcal{A}|}{2^M}$ yields the lower bound on the fraction of the codewords in $\mathcal{A}$,

$$\frac{|\mathcal{A}|}{2^M} \geq \frac{1}{2}. \tag{S23}$$

Combining equations (S21) and (S23) results in a positive lower bound on Bob's probability of decoding error $\mathbb{P}_e^{(b)} \geq \frac{1}{4} - \sqrt{\frac{\epsilon}{\eta_w}}$ for $\epsilon \in \left(0, \frac{\eta_w}{16}\right]$ and any n, and demonstrates that reliable covert communication over a pure-loss channel is impossible. ■

Remark—The minimum probability of discrimination error between the states given by equations (S6) and (S7) satisfies [46, Section III]:

$$\frac{1 - \sqrt{1 - {}^{W^n}\langle\mathbf{0}|\, \hat{\rho}_u^{W^n}\, |\mathbf{0}\rangle^{W^n}}}{2} \leq \min \mathbb{P}_e^{(w)} \leq \frac{1}{2} {}^{W^n}\langle\mathbf{0}|\, \hat{\rho}_u^{W^n}\, |\mathbf{0}\rangle^{W^n}.$$

Since $\frac{{}^{W^n}\langle\mathbf{0}|\hat{\rho}_u^{W^n}|\mathbf{0}\rangle^{W^n}}{4} \leq \frac{1-\sqrt{1-{}^{W^n}\langle\mathbf{0}|\hat{\rho}_u^{W^n}|\mathbf{0}\rangle^{W^n}}}{2}$, the error probability for the SPD is at most twice that of an optimal discriminator. Thus, the SPD is an asymptotically optimal detector when the channel from Alice is pure-loss. Since the photon number resolving (PNR) receiver, given by the POVM elements $\{|0\rangle\langle 0|, |1\rangle\langle 1|, |2\rangle\langle 2|, \ldots\}^{\otimes n}$, could be used to mimic the SPD with the detection event threshold set at one photon, the PNR receiver is also asymptotically optimal in this scenario.

Theorem 3 (Square root law for the thermal noise channel) *Suppose Willie has access to an arbitrarily complex receiver measurement as permitted by the laws of quantum physics and can capture all the photons transmitted by Alice that do not reach Bob. Let Willie's channel from Alice be subject to noise from a thermal environment that injects $\bar{n}_T > 0$ photons per optical mode on average, and let Alice and Bob share a secret of sufficient length before communicating. Then Alice can lower-bound Willie's detection error probability $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$ while reliably transmitting $\mathcal{O}(\sqrt{n})$ bits to Bob in n optical modes even if Bob only has access to a (sub-optimal) coherent detection receiver, such as an optical homodyne detector.*

First, we define quantum relative entropy and prove a lemma:

Definition 4 Quantum relative entropy *between states $\hat{\rho}_0$ and $\hat{\rho}_1$ is $D(\hat{\rho}_0\|\hat{\rho}_1) \equiv \mathrm{Tr}\{\hat{\rho}_0(\ln\hat{\rho}_0 - \ln\hat{\rho}_1)\}$.*

Lemma 5 (Quantum relative entropy between two thermal states) *If $\hat{\rho}_0 = \sum_{n=0}^{\infty} \frac{\bar{n}_0^n}{(1+\bar{n}_0)^{1+n}} |n\rangle\langle n|$ and $\hat{\rho}_1 = \sum_{n=0}^{\infty} \frac{\bar{n}_1^n}{(1+\bar{n}_1)^{1+n}} |n\rangle\langle n|$, then $D(\hat{\rho}_0\|\hat{\rho}_1) = \bar{n}_0 \ln\frac{\bar{n}_0(1+\bar{n}_1)}{\bar{n}_1(1+\bar{n}_0)} + \ln\frac{1+\bar{n}_1}{1+\bar{n}_0}$*

Proof. Express $D(\hat{\rho}_0\|\hat{\rho}_1) = -\mathrm{Tr}\{\hat{\rho}_0(\ln\hat{\rho}_1)\} - S(\hat{\rho}_0)$, where $S(\hat{\rho}_0) \equiv -\mathrm{Tr}[\hat{\rho}_0 \ln\hat{\rho}_0]$ is the von Neumann entropy of the state $\hat{\rho}_0$:

$$S(\hat{\rho}_0) = \ln(1+\bar{n}_0) + \bar{n}_0 \ln\left(1 + \frac{1}{\bar{n}_0}\right). \tag{S24}$$

Now,

$$\begin{aligned}
\mathrm{Tr}[\hat{\rho}_0 \ln\hat{\rho}_1] &= \mathrm{Tr}\left[\left(\sum_{n=0}^{\infty} \frac{\bar{n}_0^n}{(1+\bar{n}_0)^{1+n}} |n\rangle\langle n|\right)\left(\sum_{n=0}^{\infty} \ln\frac{\bar{n}_1^n}{(1+\bar{n}_1)^{1+n}} |n\rangle\langle n|\right)\right] \\
&= \sum_{n=0}^{\infty} \frac{\bar{n}_0^n}{(1+\bar{n}_0)^{1+n}} \ln\frac{\bar{n}_1^n}{(1+\bar{n}_1)^{1+n}} \\
&= \frac{1}{1+\bar{n}_0} \ln\frac{1}{1+\bar{n}_1} \sum_{n=0}^{\infty}\left(\frac{\bar{n}_0}{1+\bar{n}_0}\right)^n + \ln\frac{\bar{n}_1}{1+\bar{n}_1} \sum_{n=0}^{\infty} \frac{n}{1+\bar{n}_0} \cdot \left(\frac{\bar{n}_0}{1+\bar{n}_0}\right)^n \\
&= \ln\frac{1}{1+\bar{n}_1} + \bar{n}_0 \ln\frac{\bar{n}_1}{1+\bar{n}_1}
\end{aligned} \tag{S25}$$

where (S25) is due to the geometric series $\sum_{n=0}^{\infty}\left(\frac{\bar{n}_0}{1+\bar{n}_0}\right)^n = \left(1-\frac{\bar{n}_0}{1+\bar{n}_0}\right)^{-1}$ and $\sum_{n=0}^{\infty}\frac{n}{1+\bar{n}_0}\left(\frac{\bar{n}_0}{1+\bar{n}_0}\right)^n = \bar{n}_0$ being the expression for the mean of geometrically-distributed random variable $X \sim \text{Geom}\left(\frac{1}{1+\bar{n}_0}\right)$. Combining (S24) and (S25) yields the lemma. ∎

Proof. *(Theorem 3) Construction:* Let Alice use a zero-mean isotropic Gaussian-distributed coherent state input $\{p(\alpha), |\alpha\rangle\}$, where $\alpha \in \mathbb{C}$, $p(\alpha) = e^{-|\alpha|^2/\bar{n}}/\pi\bar{n}$ with mean photon number per symbol $\bar{n} = \int_{\mathbb{C}} |\alpha|^2 p(\alpha)\mathrm{d}^2\alpha$. Alice encodes M-bit blocks of input into codewords of length n symbols by generating 2^M codewords $\{\bigotimes_{i=1}^{n}|\alpha_i\rangle_k\}_{k=1}^{2^M}$, each according to $p(\bigotimes_{i=1}^{n}|\alpha_i\rangle) = \prod_{i=1}^{n} p(\alpha_i)$, where $\bigotimes_{i=1}^{n}|\alpha_i\rangle = |\alpha_1 \ldots \alpha_n\rangle$ is an n-mode tensor-product coherent state. The codebook is used only once to send a single message and is kept secret from Willie, though he knows how it is constructed.

Analysis (Willie): Since Willie does not have access to Alice's codebook, Willie has to discriminate between the following n-copy quantum states:

$$\hat{\rho}_0^{\otimes n} = \left(\sum_{i=0}^{\infty}\frac{(\eta_b\bar{n}_T)^i}{(1+\eta_b\bar{n}_T)^{1+i}}|i\rangle\langle i|\right)^{\otimes n}, \text{ and}$$

$$\hat{\rho}_1^{\otimes n} = \left(\sum_{i=0}^{\infty}\frac{(\eta_w\bar{n}+\eta_b\bar{n}_T)^i}{(1+\eta_w\bar{n}+\eta^{(n)}\bar{n}_T)^{1+i}}|i\rangle\langle i|\right)^{\otimes n}.$$

Willie's average probability of error in discriminating between $\hat{\rho}_0^{\otimes n}$ and $\hat{\rho}_1^{\otimes n}$ is [45, Section 9.1.4]:

$$\mathbb{P}_e^{(w)} \geq \frac{1}{2}\left[1 - \frac{1}{2}\|\hat{\rho}_1^{\otimes n} - \hat{\rho}_0^{\otimes n}\|_1\right],$$

where the minimum in this case is attained by a PNR detection. The trace distance $\|\hat{\rho}_0 - \hat{\rho}_1\|_1$ between states $\hat{\rho}_1$ and $\hat{\rho}_1$ is upper-bounded the quantum relative entropy (QRE) using quantum Pinsker's Inequality [45, Theorem 11.9.2] as follows:

$$\|\hat{\rho}_0 - \hat{\rho}_1\|_1 \leq \sqrt{2D(\hat{\rho}_0\|\hat{\rho}_1)},$$

Thus,

$$\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \sqrt{\frac{1}{8}D(\hat{\rho}_0^{\otimes n}\|\hat{\rho}_1^{\otimes n})}. \tag{S26}$$

QRE is additive for tensor product states:

$$D(\hat{\rho}_0^{\otimes n}\|\hat{\rho}_1^{\otimes n}) = nD(\hat{\rho}_0\|\hat{\rho}_1). \tag{S27}$$

By Lemma 5,

$$D(\hat{\rho}_0\|\hat{\rho}_1) = \eta_b\bar{n}_T \ln\frac{(1+\eta_w\bar{n}+\eta_b\bar{n}_T)\eta_b\bar{n}_T}{(\eta_w\bar{n}+\eta_b\bar{n}_T)(1+\eta_b\bar{n}_T)} + \ln\frac{1+\eta_w\bar{n}+\eta_b\bar{n}_T}{1+\eta_b\bar{n}_T}. \tag{S28}$$

The first two terms of the Taylor series expansion of the RHS of (S28) with respect to $\bar{n}$ at $\bar{n} = 0$ are zero and the fourth term is negative. Thus, using Taylor's Theorem with the remainder, we can upper-bound equation (S28) by the third term as follows:

$$D(\hat{\rho}_0\|\hat{\rho}_1) \leq \frac{\eta_w^2\bar{n}^2}{2\eta_b\bar{n}_T(1+\eta_b\bar{n}_T)}. \tag{S29}$$

Combining equations (S26), (S27), and (S29) yields:

$$\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \frac{\eta_w\bar{n}\sqrt{n}}{4\sqrt{\eta_b\bar{n}_T(1+\eta_b\bar{n}_T)}} \tag{S30}$$

Therefore, setting

$$\bar{n} = \frac{4\epsilon\sqrt{\eta_b \bar{n}_T(1+\eta_b\bar{n}_T)}}{\sqrt{n}\eta_w} \tag{S31}$$

ensures that Willie's error probability is lower-bounded by $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ over n optical modes.

Analysis (Bob): Suppose Bob uses a coherent detection receiver. A homodyne receiver, which is more efficient than a heterodyne receiver in the low photon number regime [28], induces an AWGN channel with noise power $\sigma_b^2 = \frac{2(1-\eta_b)\bar{n}_T+1}{4\eta_b}$. Since Alice uses Gaussian modulation with symbol power $\bar{n}$ defined in equation (S31), we can upper-bound $\mathbb{P}_e^{(b)}$ by [11, Equation (9)]:

$$\mathbb{P}_e^{(b)} \leq 2^{B-\frac{n}{2}\log_2\left(1+\bar{n}/2\sigma_b^2\right)}, \tag{S32}$$

where B is the number of transmitted bits. Substitution of $\bar{n}$ from (S31) into (S32) shows that $\mathcal{O}(\sqrt{n})$ bits can be covertly transmitted from Alice to Bob with $\mathbb{P}_e^{(b)} < \delta$ for arbitrary $\delta > 0$ given large enough n. ■

Before proving Theorems 7 and 8, we state a lemma that is used in their proofs.

Lemma 6 (Classical relative entropy bound on $\mathbb{P}_e$ of binary hypothesis test) *Denote by $\mathbb{P}_0$ and $\mathbb{P}_1$ the respective probability distributions of observations when H_0 and H_1 is true. Assuming equal prior probabilities for each hypothesis, the probability of discrimination error is $\mathbb{P}_e \leq \frac{1}{2} - \sqrt{\frac{1}{8}D(\mathbb{P}_0\|\mathbb{P}_1)}$, where $D(\mathbb{P}_0\|\mathbb{P}_1) = -\sum_x p_0(x)\ln\frac{p_1(x)}{p_0(x)}$ is the classical relative entropy between $\mathbb{P}_0$ and $\mathbb{P}_1$ and $p_0(x)$ and $p_1(x)$ are the respective probability mass functions of $\mathbb{P}_0$ and $\mathbb{P}_1$.*

Proof. The minimum probability of discrimination error between H_0 and H_1 is characterized by [37, Theorem 13.1.1]:

$$\min \mathbb{P}_e = \frac{1}{2} - \frac{1}{4}\|p_0(x) - p_1(x)\|_1,$$

where $\|a-b\|_1$ is the $\mathcal{L}_1$ norm. By classical Pinsker's inequality [47, Lemma 11.6.1],

$$\|p_0(x) - p_1(x)\|_1 \leq \sqrt{2D(\mathbb{P}_0\|\mathbb{P}_1)},$$

and the lemma follows. ■

Theorem 7 (Dark counts yield square root law) *Suppose that Willie has a pure-loss channel from Alice, captures all photons transmitted by Alice that do not reach Bob, but is limited to a receiver with a non-zero dark current. Let Alice and Bob share a secret of sufficient length before communicating. Then Alice can lower-bound Willie's detection error probability $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$ while reliably transmitting $\mathcal{O}(\sqrt{n})$ bits to Bob in n optical modes.*

Proof. *Construction:* Let Alice use a coherent state on-off keying (OOK) modulation $\{\pi_i, |\psi_i\rangle\langle\psi_i|\}$, $i = 1, 2$, where $\pi_1 = 1-q$, $\pi_2 = q$, $|\psi_1\rangle = |0\rangle$, $|\psi_2\rangle = |\alpha\rangle$. Alice and Bob generate a random codebook with each codeword symbol chosen i.i.d. from the above binary OOK constellation.

Analysis (Willie): Willie records vector $\mathbf{y}_w = [y_1, \ldots, y_n]$, where y_i is the number of photons observed in the i^{th} mode. Denote by $\mathbb{P}_0$ the distribution of $\mathbf{y}_w$ when Alice does not transmit and by $\mathbb{P}_1$ the distribution when she transmits. When Alice does not transmit, Willie's receiver observes a Poisson dark count process with rate λ_w photons per mode. Thus, $\{y_i\}$ is independent and identically distributed (i.i.d.) sequence of Poisson random variables with rate λ_w, and $\mathbb{P}_0 = \mathbb{P}_w^n$ where $\mathbb{P}_w = \text{Poisson}(\lambda_w)$. When Alice transmits, although Willie captures all of her transmitted energy that does not reach Bob, he does not have access to Alice's and Bob's codebook. Since the dark counts are independent of the transmitted pulses, each observation is a mixture of two independent Poisson random variables. Thus, each $y_i \sim \mathbb{P}_s$ is i.i.d., with

$\mathbb{P}_s = (1-q)\text{Poisson}(\lambda_w) + q\text{Poisson}(\lambda_w + \eta_w|\alpha|^2)$ and $\mathbb{P}_1 = \mathbb{P}_s^n$. By Lemma 6, $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \sqrt{\frac{1}{8}D(\mathbb{P}_0\|\mathbb{P}_1)}$. Since the classical relative entropy is additive for product distributions, $D(\mathbb{P}_0\|\mathbb{P}_1) = nD(\mathbb{P}_w\|\mathbb{P}_s)$. Now,

$$\begin{aligned} D(\mathbb{P}_w\|\mathbb{P}_s) &= -\sum_{y=0}^{\infty} \frac{\lambda_w^y e^{-\lambda_w}}{y!} \log\left[1-q+q\left(1+\frac{\eta_w|\alpha|^2}{\lambda_w}\right)e^{-\eta_w|\alpha|^2}\right] \\ &\leq \frac{q^2\left(e^{(\eta_w|\alpha|^2)^2/\lambda_w}-1\right)}{2} \end{aligned} \tag{S33}$$

where the inequality is due to the Taylor's Theorem with the remainder applied to the Taylor series expansion of equation (S33) with respect to q at $q=0$. Thus,

$$\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \frac{q}{4}\sqrt{n\left(e^{(\eta_w|\alpha|^2)^2/\lambda_w}-1\right)}. \tag{S34}$$

Therefore, to ensure that $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$, Alice sets

$$q = \frac{4\epsilon}{\sqrt{n\left(e^{(\eta_w|\alpha|^2)^2/\lambda_w}-1\right)}}. \tag{S35}$$

Analysis (Bob): Suppose Bob uses a practical single photon detector (SPD) receiver with probability of a dark click per mode $p_D^{(b)}$. This induces a binary asymmetric channel between Alice and Bob, where the click probabilities, conditional on the input, are $\mathbb{P}(\text{click} \mid \text{input } |0\rangle) = p_D^{(b)}$ and $\mathbb{P}(\text{click} \mid \text{input } |\alpha\rangle) = 1 - e^{-\eta_b|\alpha|^2}(1-p_D^{(b)})$, with the corresponding no-click probabilities $\mathbb{P}(\text{no-click} \mid \text{input } |0\rangle) = 1 - p_D^{(b)}$ and $\mathbb{P}(\text{no-click} \mid \text{input } |\alpha\rangle) = e^{-\eta_b|\alpha|^2}(1-p_D^{(b)})$. At each mode, a click corresponds to "1" and no-click to "0". Let Bob use a maximum likelihood decoder on this sequence. Then the standard upper bound on Bob's average decoding error probability is [48] $\mathbb{P}_e^{(b)} \leq e^{B-nE_0}$, where B is the number of transmitted bits, and E_0 is:

$$E_0 = -\ln\left[\left(1-p_d^{(b)}\right)\left(1-q\left(1-e^{-\frac{\eta|\alpha|^2}{2}}\right)\right)^2 + \left((1-q)\sqrt{p_D^{(b)}} + q\sqrt{1-\left(1-p_D^{(b)}\right)e^{-\eta|\alpha|^2}}\right)^2\right]$$

The Taylor series expansion of E_0 with respect to q at $q=0$ yields $E_0 = qC + \mathcal{O}(q^2)$, where

$$C = 2e^{-\eta_n|\alpha|^2/2}\left(e^{\eta_n|\alpha|^2/2} - 1 + p_D^{(b)} - \sqrt{p_D^{(b)}\left(e^{\eta_n|\alpha|^2/2}-1+p_D^{(b)}\right)}\right)$$

is a positive constant. Since $q = \mathcal{O}(1/\sqrt{n})$, this demonstrates that $\mathcal{O}(\sqrt{n})$ bits can be covertly transmitted from Alice to Bob with $\mathbb{P}_e^{(b)} < \delta$ for arbitrary $\delta > 0$ given large enough n. ■

Theorem 8 (Dark counts yield square root law under structured modulation) *Suppose that Willie has a pure-loss channel from Alice, can capture all photons transmitted by Alice that do not reach Bob, but is limited to a PNR receiver with a non-zero dark current. Let Alice and Bob share a secret of sufficient length before communicating. Then Alice can lower-bound Willie's detection error probability $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$ while reliably transmitting $\mathcal{O}(\sqrt{\frac{n}{Q}}\log Q)$ bits to Bob using n optical modes and a Q-ary PPM constellation.*

Proof. *Construction:* Prior to communication, Alice and Bob secretly choose a random subset $\mathcal{S}$ of PPM frames to use for transmission by selecting each of n/Q available PPM frames independently with probability ζ. Alice and Bob then secretly generate a vector $\mathbf{k}$ containing $|\mathcal{S}|$ numbers selected independently uniformly at random from $\{0, 1, \ldots, Q-1\}$, where $|\mathcal{S}|$ denotes the cardinality of $\mathcal{S}$. Alice encodes a message into a codeword of size $|\mathcal{S}|$ using an ECC that may be known to Willie. She adds $\mathbf{k}$ modulo Q to this message and transmits it on the PPM frames in $\mathcal{S}$.

Analysis (Willie): Willie detects each PPM frame received from Alice, recording the photon counts in $\mathbf{y}_w = [\mathbf{y}_1^{(w)}, \dots, \mathbf{y}_n^{(w)}]$ where $\mathbf{y}_i^{(w)} = [y_{i,1}^{(w)}, \dots, y_{i,Q}^{(w)}]$ and $y_{i,j}^{(w)}$ it the number of photons observed in the j^{th} mode of the i^{th} PPM frame. Denote by $\mathbb{P}_0$ the distribution of $\mathbf{y}_w$ when Alice does not transmit and by $\mathbb{P}_1$ the distribution when she transmits. When Alice does not transmit, Willie's receiver observes a Poisson dark count process with rate λ_w photons per mode, implying that $\mathbf{y}_w$ is a vector of nQ i.i.d. Poisson(λ_w) random variables. Therefore, $\{\mathbf{y}_i^{(w)}\}$ is i.i.d. with $\mathbf{y}_i^{(w)} \sim \mathbb{P}_w$ and $\mathbb{P}_0 = \mathbb{P}_w^n$, where $\mathbb{P}_w$ is the distribution of Q i.i.d. Poisson(λ_w) random variables with p.m.f.:

$$p_0(\mathbf{y}_i^{(w)}) = \prod_{j=1}^{Q} \frac{\lambda_w^{y_{i,j}^{(w)}} e^{-\lambda_w}}{y_{i,j}^{(w)}!}. \tag{S36}$$

When Alice transmits, by construction, each PPM frame is randomly selected for transmission with probability ζ. In each selected PPM frame, a pulse is transmitted using one of Q modes chosen equiprobably. Therefore, in this case $\{\mathbf{y}_i^{(w)}\}$ is also i.i.d. with $\mathbf{y}_i^{(w)} \sim \mathbb{P}_s$ and $\mathbb{P}_1 = \mathbb{P}_s^n$, where the p.m.f. of $\mathbb{P}_s$ is:

$$p_1(\mathbf{y}_i^{(w)}) = (1-\zeta) \prod_{j=1}^{Q} \frac{\lambda_w^{y_{i,j}^{(w)}} e^{-\lambda_w}}{y_{i,j}^{(w)}!} + \frac{\zeta}{Q} \sum_{m=1}^{Q} \frac{(\eta_w |\alpha|^2 + \lambda_w)^{y_{i,m}^{(w)}} e^{-\eta_w|\alpha|^2 - \lambda_w}}{y_{i,m}^{(w)}!} \prod_{\substack{j=1 \\ j \neq m}}^{Q} \frac{\lambda_w^{y_{i,j}^{(w)}} e^{-\lambda_w}}{y_{i,j}^{(w)}!}. \tag{S37}$$

Since the classical relative entropy is additive for product distributions, $D(\mathbb{P}_0 \| \mathbb{P}_1) = \frac{n}{Q} D(\mathbb{P}_w \| \mathbb{P}_s)$. Now, denoting by $\mathbf{x} = [x_1, \cdots, x_Q]$ where $x_j \in \mathbb{N}_0$, we have:

$$\begin{aligned} D(\mathbb{P}_w \| \mathbb{P}_s) &= -\sum_{\mathbf{x} \in \mathbb{N}_0^Q} \prod_{j=1}^{Q} \frac{\lambda_w^{x_j} e^{-\lambda_w}}{x_j!} \log \left[1 - \zeta + \frac{\zeta}{Q} \sum_{m=1}^{Q} \left(1 + \frac{\eta_w |\alpha|^2}{\lambda_w} \right)^{x_m} e^{-\eta_w |\alpha|^2} \right] \\ &\leq \frac{\zeta^2 \left(e^{(\eta_w|\alpha|^2)^2/\lambda_w} - 1 \right)}{2Q} \end{aligned} \tag{S38}$$

where the inequality is due to the Taylor's Theorem with the remainder applied to the Taylor series expansion of equation (S38) with respect to ζ at $\zeta = 0$. By Lemma 6, $\zeta = \frac{4\epsilon Q}{\sqrt{n\left(e^{(\eta_w|\alpha|^2)^2/\lambda_w} - 1\right)}}$ ensures that Willie's error probability is lower-bounded by $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$.

Analysis (Bob): As in the proof of Theorem 7, Bob uses a practical SPD receiver with probability of a dark click $p_D^{(b)}$. Bob examines only the PPM frames in $\mathcal{S}$. If two or more clicks are detected in a PPM frame, a PPM symbol is assigned by selecting one of the clicks uniformly at random. If no clicks are detected, the PPM frame is labeled as an *erasure*. After subtracting $\mathbf{k}$ modulo Q from this vector of PPM symbols (subtraction is not performed on erasures), the resultant vector is passed to the decoder. A random coding argument [43, Theorem 5.6.2] yields reliable transmission of $\mathcal{O}\left(\sqrt{\frac{n}{Q}} \log Q\right)$ covert bits. ■

Theorem 9 (Converse of the square root law) *Suppose Alice only uses n-mode codewords with total photon number variance $\sigma_x^2 = \mathcal{O}(n)$. Then, if she attempts to transmit $\omega(\sqrt{n})$ bits in n modes, as $n \to \infty$, she is either detected by Willie with arbitrarily low detection error probability, or Bob cannot decode with arbitrarily low decoding error probability.*

Proof. As in the proof of Theorem 1, Alice sends one of 2^M (equally likely) M-bit messages by choosing an element from an arbitrary codebook $\{\hat{\rho}_x^{A^n}, x = 1, \dots, 2^M\}$, where a state $\hat{\rho}_x^{A^n} = |\psi_x\rangle^{A^n A^n}\langle\psi_x|$ encodes an M-bit message W_x. $|\psi_x\rangle^{A^n} = \sum_{\mathbf{k} \in \mathbb{N}_0^n} a_{\mathbf{k}}(x) |\mathbf{k}\rangle$ is a general n-mode pure state, where $|\mathbf{k}\rangle \equiv |k_1\rangle \otimes$

$|k_2\rangle \otimes \cdots \otimes |k_n\rangle$ is a tensor product of n Fock states. The mean photon number of a codeword $\hat{\rho}_x^{A^n}$ is $\bar{n}_x = \sum_{\mathbf{k}\in\mathbb{N}_0^n}(\sum_{i=1}^n k_i)|a_{\mathbf{k}}(x)|^2$, and the photon number variance is $\sigma_x^2 = \sum_{\mathbf{k}\in\mathbb{N}_0^n}(\sum_{i=1}^n k_i)^2|a_{\mathbf{k}}(x)|^2 - \bar{n}_x^2 = \mathcal{O}(n)$. We limit our analysis to pure input states since, by convexity, using mixed states as inputs can only deteriorate the performance (since that is equivalent to transmitting a randomly chosen pure state from an ensemble and discarding the knowledge of that choice).

Willie uses a noisy PNR receiver to observe his channel from Alice, and records the total photon count X_{tot} over n modes. For some threshold S that we discuss later, Willie declares that Alice transmitted when $X_{tot} \geq S$, and did not transmit when $X_{tot} < S$. When Alice does not transmit, Willie observes noise: $X_{tot}^{(0)} = X_D + X_T$, where X_D is the number of dark counts due to the spontaneous emission process at the detector, and X_T is the number of photons observed due to the thermal background. Since the dark counts are modeled by a Poisson process with rate λ_w photons per mode, both the mean and variance of the observed dark counts per mode is λ_w. The mean of the number of photons observed per mode from the thermal background with mean photon number per mode $\bar{n}_T$ is $(1-\eta_w)\bar{n}_T$ and the variance is $(1-\eta_w)^2(\bar{n}_T+\bar{n}_T^2)$. Thus, the mean of the total number of noise photons observed per mode is $\mu_N = \lambda_w + (1-\eta_w)\bar{n}_T$, and, due to the statistical independence of the noise processes, the variance is $\sigma_N^2 = \lambda_w + (1-\eta_w)^2(\bar{n}_T+\bar{n}_T^2)$. We upper-bound the false alarm probability using Chebyshev's inequality:

$$\begin{aligned}\mathbb{P}_{\mathrm{FA}} &= \mathbb{P}(X_{tot}^{(0)} \geq S)\\ &\leq \frac{n\sigma_N^2}{(S-n\mu_N)^2},\end{aligned} \tag{S39}$$

where equation (S39) is due to the memorylessness of the noise processes. Thus, to obtain the desired $\mathbb{P}_{\mathrm{FA}}^*$, Willie sets threshold $S = n\mu_N + \sqrt{n\sigma_N^2/\mathbb{P}_{\mathrm{FA}}^*}$.

When Alice transmits codeword $\hat{\rho}_u^{A^n}$ corresponding to message W_u, Willie observes $X_{tot}^{(1)} = X_u + X_D + X_T$, where X_u is the count due to Alice's transmission. We upper-bound the missed detection probability using Chebyshev's inequality:

$$\begin{aligned}\mathbb{P}_{\mathrm{MD}} &= \mathbb{P}(X_{tot}^{(1)} < S)\\ &\leq \mathbb{P}\left(|X_{tot}^{(1)} - \eta_w\bar{n}_u - n\mu_N| \geq \eta_w\bar{n}_u - \sqrt{\frac{n\sigma_N^2}{\mathbb{P}_{\mathrm{FA}}^*}}\right)\\ &\leq \frac{n\sigma_N^2 + \eta_w^2\sigma_u^2}{(\eta_w\bar{n}_u - \sqrt{n\sigma_N^2/\mathbb{P}_{\mathrm{FA}}^*})^2},\end{aligned} \tag{S40}$$

where equation (S40) is due to the independence between the noise and Alice's codeword. Since $\sigma_u^2 = \mathcal{O}(n)$, if $\bar{n}_u = \omega(\sqrt{n})$, then $\lim_{n\to\infty}\mathbb{P}_{\mathrm{MD}} = 0$. Thus, given large enough n, Willie can detect Alice's codewords that have mean photon number $\bar{n}_u = \omega(\sqrt{n})$ with probability of error $\mathbb{P}_e^{(w)} \leq \epsilon$ for any $\epsilon > 0$.

Now, if Alice wants to lower-bound $\mathbb{P}_e^{(w)}$, her codebook must contain a positive fraction of codewords with mean photon number upper-bounded by $\bar{n}_{\mathcal{U}} = \mathcal{O}(\sqrt{n})$. Formally, there must exist a subset of the codebook $\left\{\hat{\rho}_u^{A^n}, u \in \mathcal{U}\right\}$, where $\mathcal{U} = \{u : \bar{n}_u \leq \bar{n}_{\mathcal{U}}\}$, with $\frac{|\mathcal{U}|}{2^M} \geq \kappa$ and $\kappa > 0$. Suppose Bob has an unattenuated pure-loss channel from Alice ($\eta_b = 0$ and $\bar{n}_T = 0$) and access to any receiver allowed by quantum mechanics. The decoding error probability $\mathbb{P}_e^{(b)}$ in such scenario clearly lower-bounds the decoding error probability in a practical scenario where the channel from Alice is lossy and either the channel or the receiver are noisy. Denote by $E_{a\to k}$ the event that a transmitted message W_a is decoded as $W_k \neq W_a$. Since the messages are equiprobable, the average probability of error for the codebook containing only the codewords in $\mathcal{U}$ is:

$$\mathbb{P}_e^{(b)}(\mathcal{U}) = \frac{1}{|\mathcal{U}|}\sum_{a\in\mathcal{U}}\mathbb{P}\left(\cup_{k\in\mathcal{U}\setminus\{a\}}E_{a\to k}\right). \tag{S41}$$

Since the probability that a message is sent from $\mathcal{U}$ is κ,

$$\mathbb{P}_e^{(b)} \geq \kappa\mathbb{P}_e^{(b)}(\mathcal{U}). \tag{S42}$$

Equality holds only when Bob receives messages that are not in $\mathcal{U}$ error-free and knows when the messages from $\mathcal{U}$ are sent (in other words, equality holds when the set of messages on which decoder can err is reduced to $\mathcal{U}$). Denote by W_a, $a \in \mathcal{U}$, the message transmitted by Alice, and by $\hat{W}_a$ Bob's decoding of W_a. Then, since each message is equiprobable and $|\mathcal{U}| = \kappa 2^M$,

$$\begin{aligned}\log_2 \kappa + M &= H(W_a) && \text{(S43)}\\ &= I(W_a;\hat{W}_a) + H(W_a|\hat{W}_a) && \text{(S44)}\\ &\leq I(W_a;\hat{W}_a) + 1 + (\log_2 \kappa + M)\mathbb{P}_e^{(b)}(\mathcal{U}) && \text{(S45)}\\ &\leq \chi\left(\left\{\frac{1}{|\mathcal{U}|}, \hat{\rho}_u^{A^n}\right\}\right) + 1 + (\log_2 \kappa + M)\mathbb{P}_e^{(b)}(\mathcal{U}) && \text{(S46)}\end{aligned}$$

where (S44) is from the definition of mutual information, (S45) is due to classical Fano's inequality [47, Equation (9.37)], and (S46) is the Holevo bound $I(X;Y) \leq \chi(\{p_X(x), \hat{\rho}_x\})$ [49]. The mutual information $I(X;Y)$ is between a classical input X and a classical output Y, which is a function of the prior probability distribution $p_X(x)$, and the conditional probability distribution $p_{Y|X}(y|x)$, with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The classical input x maps to a quantum state $\hat{\rho}_x$. A *specific choice* of a quantum measurement, described by POVM elements $\{\Pi_y, y \in \mathcal{Y}\}$, induces the conditional probability distribution $p_{Y|X}(y|x) = \mathrm{Tr}\left[\Pi_y \hat{\rho}_x\right]$. The Holevo information, $\chi(\{p_x, \hat{\rho}_x\}) = S\left(\sum_{x\in\mathcal{X}} p_x \hat{\rho}_x\right) - \sum_{x\in\mathcal{X}} p_x S(\hat{\rho}_x)$, where $S(\hat{\rho}) \equiv -\mathrm{Tr}[\hat{\rho}\ln\hat{\rho}]$ is the von Neumann entropy of the state $\hat{\rho}$, is not a function of the quantum measurement. Since $\hat{\rho}_u^{A^n} = |\psi_u\rangle^{A^n A^n}\langle\psi_u|$ is a pure state, $\chi\left(\{\frac{1}{|\mathcal{U}|}, \hat{\rho}_u^{A^n}\}\right) = S\left(\frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}} |\psi_u\rangle^{A^n A^n}\langle\psi_u|\right)$. Denote the "average codeword" in $\mathcal{U}$ by $\bar{\rho}^{A^n} = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}} |\psi_u\rangle^{A^n A^n}\langle\psi_u|$, and the state of the j^{th} mode of $\bar{\rho}^{A^n}$ by $\bar{\rho}_j^{A^n}$. We obtain $\bar{\rho}_j^{A^n}$ by tracing out all the other modes in $\bar{\rho}^{A^n}$ and denote its mean photon number by $\bar{n}_j$ (i.e. $\bar{n}_j$ is the mean photon number of the j^{th} mode of $\bar{\rho}^{A^n}$). Finally, denote a coherent state ensemble with a zero-mean circularly-symmetric Gaussian distribution by $\hat{\rho}_{\bar{n}}^T = \frac{1}{\pi\bar{n}}\int e^{-|\alpha|^2/\bar{n}}|\alpha\rangle\langle\alpha|\mathrm{d}^2\alpha$. The von Neumann entropy of $\hat{\rho}_{\bar{n}}^T$, $S\left(\hat{\rho}_{\bar{n}}^T\right) = \log_2(1+\bar{n}) + \bar{n}\log_2\left(1+\frac{1}{\bar{n}}\right)$. Now,

$$\begin{aligned}S\left(\bar{\rho}^{A^n}\right) &\leq \sum_{j=1}^{n} S\left(\bar{\rho}_j^{A^n}\right) && \text{(S47)}\\ &\leq \sum_{i=1}^{n} \log_2(1+\bar{n}_j) + \bar{n}_j \log_2\left(1+\frac{1}{\bar{n}_j}\right) && \text{(S48)}\\ &\leq n\left(\log_2\left(1+\frac{\bar{n}_{\mathcal{U}}}{n}\right) + \frac{\bar{n}_{\mathcal{U}}}{n}\log_2\left(1+\frac{n}{\bar{n}_{\mathcal{U}}}\right)\right), && \text{(S49)}\end{aligned}$$

where (S47) follows from the sub-additivity of the von Neumann entropy and (S48) is due to $\hat{\rho}_{\bar{n}}^T$ maximizing the von Neumann entropy of a single-mode state with mean photon number constraint $\bar{n}$ [28]. Now, $S\left(\hat{\rho}_{\bar{n}}^T\right)$ is concave and increasing for $\bar{n} > 0$, and, since $\sum_{j=1}^n \bar{n}_j \leq \bar{n}_{\mathcal{U}}$ by construction of $\mathcal{U}$, the application of Jensen's inequality yields (S49). Combining (S46) and (S49) and solving for $\mathbb{P}_e^{(b)}(\mathcal{U})$ yields:

$$\mathbb{P}_e^{(b)}(\mathcal{U}) \geq 1 - \frac{\log_2\left(1+\frac{\bar{n}_{\mathcal{U}}}{n}\right) + \frac{\bar{n}_{\mathcal{U}}}{n}\log_2\left(1+\frac{n}{\bar{n}_{\mathcal{U}}}\right) + \frac{1}{n}}{\frac{\log_2\kappa}{n} + \frac{M}{n}}. \tag{S50}$$

Substituting (S50) into (S42) yields the following lower bound on Bob's decoding error probability:

$$\mathbb{P}_e^{(b)} \geq \kappa\left[1 - \frac{\log_2\left(1+\frac{\bar{n}_{\mathcal{U}}}{n}\right) + \frac{\bar{n}_{\mathcal{U}}}{n}\log_2\left(1+\frac{n}{\bar{n}_{\mathcal{U}}}\right) + \frac{1}{n}}{\frac{\log_2\kappa}{n} + \frac{M}{n}}\right] \tag{S51}$$

Since Alice transmits $\omega(\sqrt{n})$ bits in n modes, $M/n = \omega(1/\sqrt{n})$ bits/symbol. However, since $\bar{n}_{\mathcal{U}} = \mathcal{O}(\sqrt{n})$, as $n \to \infty$, $\mathbb{P}_e^{(b)}$ is bounded away from zero for any $\kappa > 0$. Thus, Alice cannot transmit $\omega(\sqrt{n})$ bits in n optical modes both covertly and reliably. ■

II. DETAILS OF THE EXPERIMENTAL METHODS

Here we provide the mathematical details of the experimental methods.

Calculation of Bob's maximum throughput

The Q-ary PPM signaling combined with Bob's device for assigning symbols to received PPM frames induces a discrete memoryless channel described by a conditional distribution $\mathbb{P}(Y|X)$, where $X \in \{1,\ldots,Q\}$ is Alice's input symbol and $Y \in \{1,\ldots,Q,\mathcal{E}\}$ is Bob's output symbol with $\mathcal{E}$ indicating an erasure. Since Bob observes Alice's pulse with probability $1-e^{-\bar{n}_{det}^{(b)}}$, $\mathbb{P}(Y|X)$ is characterized as follows:

$$\mathbb{P}(Y=x|X=x) = \left(1-e^{-\bar{n}_{det}^{(b)}}\right)\sum_{i=0}^{Q-1}\frac{1}{i+1}\left(p_D^{(b)}\right)^i\left(1-p_D^{(b)}\right)^{Q-1-i} + e^{-\bar{n}_{det}^{(b)}}\sum_{i=1}^{Q}\frac{1}{i}\left(p_D^{(b)}\right)^i\left(1-p_D^{(b)}\right)^{Q-i}$$

$$\mathbb{P}(Y=\mathcal{E}|X=x) = e^{-\bar{n}_{det}^{(b)}}\left(1-p_D^{(b)}\right)^Q$$

$$\mathbb{P}(Y=y, y\notin\{x,\mathcal{E}\}|X=x) = \frac{1-\mathbb{P}(Y=x|x)-\mathbb{P}(Y=\mathcal{E}|x)}{Q-1}$$

The symmetry of this channel yields the Shannon capacity [34] $C_s = I(X;Y)$, where $\mathbb{P}(X=x)=\frac{1}{Q}$ for $x=1,\ldots,Q$ and $I(X;Y)$ is the mutual information between X and Y. We use the experimentally-observed values from Table I to compute C_s for each regime, and plot $\frac{C_s\zeta n}{Q}$ since $\frac{\zeta n}{Q}$ is the expected number of PPM frames selected for transmission.

Derivation of the log-likelihood ratio test statistic

The log-likelihood ratio test statistic is given by $L=\frac{f_1(\mathbf{x}_w)}{f_0(\mathbf{x}_w)}$, where $f_0(\mathbf{x}_w)$ and $f_1(\mathbf{x}_w)$ are the likelihood functions of the click record $\mathbf{x}_w$ corresponding to Alice being quiet and transmitting. Click record $\mathbf{x}_w$ contains Willie's observations of each PPM frame on his channel from Alice $\mathbf{x}_w = [\mathbf{x}_1^{(w)},\ldots,\mathbf{x}_{n/Q}^{(w)}]$. $\mathbf{x}_i^{(w)} = [x_{i,1}^{(w)},\ldots,x_{i,Q}^{(w)}]$ contains the observation of the i^{th} PPM frame with $x_{i,j}^{(w)}\in\{0,1\}$, where "0" and "1" indicate the absence and the presence of a click, respectively. When Alice does not transmit, Willie only observes dark clicks. Thus, each $\mathbf{x}_w$ is a vector of i.i.d. Bernoulli $\left(p_D^{(w)}\right)$ random variables. The likelihood function of $\mathbf{x}_w$ under H_0 is then:

$$f_0(\mathbf{x}_w) = \prod_{i=1}^{n/Q}\left(p_D^{(w)}\right)^{\sum_{j=1}^{Q}x_{i,j}^{(w)}}\left(1-p_D^{(w)}\right)^{Q-\sum_{j=1}^{Q}x_{i,j}^{(w)}}.$$

Now consider the scenario when Alice transmits. The secret shared between Alice and Bob identifies the random subset $\mathcal{S}$ of the PPM frames used for transmission, and a random vector $\mathbf{k}$ which is modulo-added to the codeword. Modulo addition of $\mathbf{k}$ effectively selects a random pulse location within each PPM frame. Note that, while both the construction in the proof of Theorem 4 and Alice's encoder described in the Methods generate $\mathcal{S}$ first and then $\mathbf{k}$, the order of these operations can be reversed: we can first fix a random location of a pulse in each of n/Q PPM frames, and then select a random subset of these frames. Consider Willie's observation of the i^{th} PPM frame, and suppose that the l^{th} mode is used if the frame is selected for transmission. Denote the probability of Willie's detector observing Alice's pulse by $p_r^{(w)} = 1-e^{-\bar{n}_{det}^{(w)}}$. By construction, frames are selected for transmission independent of each other with probability ζ. Willie's detector registers a click on this mode when one of the following disjoint events occurs:

- PPM frame is selected and pulse is detected by Willie (probability $\zeta p_r^{(w)}$);

- PPM frame is selected, but Willie, instead of detecting it, records a dark click (probability $\zeta\left(1-p_r^{(w)}\right)p_D^{(w)}$); and,
- PPM frame is not selected, but Willie records a dark click (probability $(1-\zeta)p_D^{(w)}$).

The probability of the union of these events is

$$p_s^{(w)}=\zeta p_r^{(w)}\left(1-p_D^{(w)}\right)+p_D^{(w)}, \tag{S52}$$

and, thus, Willie observes an independent Bernoulli $\left(p_s^{(w)}\right)$ random variable in the l^{th} mode of the i^{th} PPM frame. Since Alice only uses the l^{th} mode for transmission, in modes other than the l^{th}, Willie observes a set of $Q-1$ i.i.d. Bernoulli $\left(p_D^{(w)}\right)$ random variables corresponding to dark clicks. The click record $\mathbf{x}_i^{(w)}$ of the i^{th} PPM frame is independent of other PPM frame and, thus, the likelihood function of $\mathbf{x}_w$ under H_1 is $\mathbb{P}_1(\mathbf{x}_w=\mathbf{x})=\prod_{i=1}^n p_1(\mathbf{x}_i^{(w)})$ where

$$f_1(\mathbf{x}_w)=\prod_{i=1}^{n/Q}\frac{1}{Q}\sum_{l=1}^{Q}\left(p_s^{(w)}\right)^{x_{i,l}^{(w)}}\left(1-p_s^{(w)}\right)^{1-x_{i,l}^{(w)}}\left(p_D^{(w)}\right)^{\sum_{\substack{j=1\\ j\neq l}}^{Q}x_{i,j}^{(w)}}\left(1-p_D^{(w)}\right)^{Q-1-\sum_{\substack{j=1\\ j\neq l}}^{Q}x_{i,j}^{(w)}}.$$

Evaluation of the likelihood ratio yields:

$$\begin{aligned}\frac{f_1(\mathbf{x}_w)}{f_0(\mathbf{x}_w)}&=\prod_{i=1}^{n/Q}\frac{1-\zeta p_r^{(w)}}{Q}\sum_{l=1}^{Q}\left(1+\frac{\zeta p_r^{(w)}}{\left(1-\zeta p_r^{(w)}\right)p_D^{(w)}}\right)^{x_{i,l}^{(w)}}\\&=\prod_{i=1}^{n/Q}\left[1-\zeta p_r^{(w)}+\frac{\zeta p_r^{(w)}y_i^{(w)}}{Qp_D^{(w)}}\right]\end{aligned} \tag{S53}$$

where $y_i^{(w)}=\sum_{l=1}^{Q}x_{i,l}^{(w)}$ and the simplification yielding (S53) is due to $x_{i,l}^{(w)}\in\{0,1\}$. Taking the logarithm of equation (S53) yields the log-likelihood ratio in (1).

Gaussian approximation of $\mathbb{P}_e^{(w)}$

For small ζ, the Taylor series expansion of the summand in (1) at $\zeta=0$, $\log\left[1+\zeta p_r^{(w)}\left(\frac{y_i^{(w)}}{Qp_D^{(w)}}-1\right)\right]\approx\zeta p_r^{(w)}\left(\frac{y_i^{(w)}}{Qp_D^{(w)}}-1\right)$, yields an approximation for the log-likelihood ratio: $L\approx\frac{\zeta p_r^{(w)}}{Qp_D^{(w)}}\left(\sum_{i=1}^{n/Q}y_i^{(w)}-np_D^{(w)}\right)$. Thus, effectively, Willie uses the total click count $Y=\sum_{i=1}^{n/Q}y_i^{(w)}$ as a test statistic, which explains the lack of sensitivity of our test to the variations in the observed channel characteristics given in Table I.

This also provides an analytical approximation of $\mathbb{P}_e^{(w)}$. First consider the case when Alice is not transmitting. Then the total click count is a binomial random variable $Y\sim\mathcal{B}\left(y;p_D^{(w)},n\right)$ whose distribution, for large n, can be approximated using the central limit theorem by a Gaussian distribution $\Phi\left(y;\mu_0,\sigma_0^2\right)$ with $\mu_0=np_D^{(w)}$ and $\sigma_0^2=np_D^{(w)}\left(1-p_D^{(w)}\right)$, where $\Phi\left(x;\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{|t-\mu|^2}{2\sigma^2}}dt$ is the distribution function of a Gaussian random variable $\mathcal{N}(x;\mu,\sigma^2)$. Now consider the case when Alice is transmitting. Since $\mathcal{S}$ and $\mathbf{k}$ are unknown to Willie, the total click count is the sum of two independent but not identical binomial random variables $Y=X+Z$, where $X\sim\mathcal{B}\left(x;p_D^{(w)},n-\frac{n}{Q}\right)$ is the number of dark clicks in the $n-\frac{n}{Q}$ modes that Alice never uses in a PPM scheme and $Z\sim\mathcal{B}\left(z;p_s^{(w)},\frac{n}{Q}\right)$ is the contribution from the $\frac{n}{Q}$ modes that Alice can use to transmit, with $p_s^{(w)}$ defined in (S52). By the central limit theorem, for large n, the distribution of X can be approximated using a Gaussian distribution $\Phi(x;\mu_X,\sigma_X^2)$ where $\mu_X = \left(n-\frac{n}{Q}\right)p_D^{(w)}$ and $\sigma_X^2 = \left(n-\frac{n}{Q}\right)p_D^{(w)}\left(1-p_D^{(w)}\right)$. Similarly, the distribution of Z can be approximated by a Gaussian distribution $\Phi\left(z;\mu_Z,\sigma_Z^2\right)$ where $\mu_Z = \frac{n}{Q}\left(\zeta p_r^{(w)} + \left(1-\zeta p_r^{(w)}\right)p_D^{(w)}\right)$ and $\sigma_Z^2 = \frac{n}{Q}\left(\zeta p_r^{(w)} + \left(1-\zeta p_r^{(w)}\right)p_D^{(w)}\right)\left(1-\zeta p_r^{(w)}\right)\left(1-p_D^{(w)}\right)$. Thus, the distribution of Y can be approximated by a Gaussian distribution $\Phi\left(y;\mu_1,\sigma_1^2\right)$ with $\mu_1 = \mu_X + \mu_Z$ and $\sigma_1^2 = \sigma_X^2 + \sigma_Z^2$ via the additivity property of independent Gaussian random variables. Willie's probability of error is thus approximated by:

$$\tilde{\mathbb{P}}_e^{(w)} = \frac{1}{2}\min_S(1-\Phi(S;\mu_0,\sigma_0^2)+\Phi(S;\mu_1,\sigma_1^2)). \tag{S54}$$

The value of the threshold S^* that minimizes the RHS of (S54) satisfies $\frac{|S^*-\mu_0|^2}{\sigma_0^2} - \log(\sigma_1^2/\sigma_0^2) = \frac{|S^*-\mu_1|^2}{\sigma_1^2}$.