Quantum data hiding in the presence of noise

Cosmo Lupo, Mark M. Wilde, and Seth Lloyd

Abstract—When classical or quantum information is broadcast to separate receivers, there exist codes that encrypt the encoded data such that the receivers cannot recover it when performing local operations and classical communication, but they can decode reliably if they bring their systems together and perform a collective measurement. This phenomenon is known as quantum data hiding and hitherto has been studied under the assumption that noise does not affect the encoded systems. With the aim of applying the quantum data hiding effect in practical scenarios, here we define the data-hiding capacity for hiding classical information using a quantum channel. Using this notion, we establish a regularized upper bound on the data hiding capacity of any quantum broadcast channel, and we prove that coherent-state encodings have a strong limitation on their data hiding rates. We then prove a lower bound on the data hiding capacity of channels that map the maximally mixed state to the maximally mixed state (we call these channels "mictodiactic"—they can be seen as a generalization of unital channels when the input and output spaces are not necessarily isomorphic) and argue how to extend this bound to generic channels and to more than two receivers.

I. INTRODUCTION

One of the primary goals of theoretical quantum information science is to identify significant separations between the classical and quantum theories of information. Many grand successes in this spirit have already been achieved: Bell inequalities [1], unconditionally secure communication [2], quantum teleportation [3], super-dense coding [4], communication complexity [5], quantum data locking [6], and so on.

Another such notable example is the quantum data hiding effect [7], [8], [9], [10]. Quantum data hiding is a communication protocol that allows for the encoding of classical or quantum information into a two-party system such that the information can be decoded if the systems are located in the same physical laboratory, whereas the information cannot be decoded if the systems are located in spatially separated laboratories, even if the laboratories are allowed to exchange classical messages. Such a task is impossible in the classical world, simply because the exchange of a classical message from one laboratory to the other allows for sending the entire classical system, such that it is subsequently located in the same physical laboratory and can thus be decoded. Arguing from the uncertainty principle of quantum mechanics, quantum data hiding allows for the encoding of information such that

C. Lupo is with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA (email: clupo@mit.edu).

M. M. Wilde is with the Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA (email: mwilde@lsu.edu).

S. Lloyd is with the Research Laboratory of Electronics and the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA (email: slloyd@mit.edu).

it is indistinguishable by local measurements and classical communication and one instead requires a global measurement to distinguish the states (i.e., information can be "hidden" in this way).

The original proposal for quantum data hiding was an impressive observation [7], the idea being to hide information in Bell states such that it can be recovered by a joint quantum operation but not by local operations and classical communication. The scheme from [7] has a low rate of hiding, and a quantum optical implementation was suggested. However, the proposal rested on there not being any loss or noise in the system and was thus impractical in several regards. Nevertheless, this striking phenomenon prompted further research, which eventually culminated in the high-rate schemes of [11]. These schemes relied upon the use of collective random unitary operations to hide both classical and quantum data, invoking mathematical techniques such as the concentration of measure. The newly proposed schemes still required no loss or noise in the encoded systems, which exclude their application in practice. These schemes were later extended to the case of hiding information from multiple parties [12].

Even though this fascinating protocol has been discussed for some years now in the literature and a variety of applications are now known, no one (as far as we are aware) has yet to consider the most pertinent practical question: Is *quantum data hiding possible in the presence of noise and loss, and if so, what is its performance?* Here, we address this fundamental question using the language of quantum Shannon theory and recently developed techniques that have established quantum data locking protocols in the presence of noise [13], [14], [15], [16], [17] (see also [18] for another advance in the same direction). The channels of primary interest to us are those which have the property of mapping the maximally mixed state in the input space into the maximally mixed state in the output space. We called these channels *mictodiactic* (literally, Greek for "mixtures passing through") as they preserve the maximally mixed state. Mictodiactic channels coincide with unital channels if the input and output space have the same dimension. We focus on these channels mainly because they are relatively simple to analyze.

The main development of this paper begins with a formal definition of the data hiding capacity of a quantum broadcast channel. From there, we establish a regularized upper bound on the data hiding capacity of any broadcast channel, and we then show that coherent-state protocols cannot offer a high rate of data hiding (this is perhaps to be expected, given that coherent states are "pseudo-classical" and in light of our prior work in which we showed that they do not offer high rates for quantum data locking [13]). We next establish a lower bound on the data hiding capacity of a mictodiactic quantum broadcast channel and argue that this scheme can be

extended to generic channels and to a multipartite scenario. The scheme involves random coding arguments, and we invoke concentration of measure bounds in order to establish security of the scheme.

II. Preliminaries

Here we recall some basic facts before beginning the main development. A quantum state is represented by a density operator, which is a positive semi-definite operator acting on a Hilbert space $\mathcal{H}$ and with trace equal to one. A multipartite quantum state acts on a tensor product of Hilbert spaces. A classical-quantum state ρ_{XB} has the following form:

$$\rho_{XB} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x, \quad (1)$$

where p_X is a probability distribution, $\{|x\rangle_X\}$ is an orthonormal basis, and $\{\rho_B^x\}$ is a set of quantum states. A quantum channel is modeled as a completely positive, trace-preserving linear map, taking operators acting on one Hilbert space to operators acting on another one. A quantum channel is unital if it preserves the identity operator (note that the input and output space of such a channel must have the same dimension). A quantum measurement is a special kind of quantum channel which accepts a quantum input system and outputs a classical system. That is, it can be described as follows:

$$\mathcal{M}(\rho) = \sum_y \mathrm{Tr}\{\Lambda^y \rho\} |y\rangle\langle y|, \quad (2)$$

where $\Lambda^y \geq 0$ for all y, $\sum_y \Lambda^y = I$, and $\{|y\rangle\}$ is some known orthonormal basis. The collection $\{\Lambda^y\}$ is called a positive operator-valued measure (POVM). A quantum instrument is a quantum channel that accepts a quantum input and outputs a quantum system and a classical system.

A quantum broadcast channel is defined as a quantum channel that accepts one input quantum system A and has two output quantum systems B and C [19]. Formally, it is a completely positive trace preserving linear map from the space of operators that act on the Hilbert space for system A to the space of operators acting on the tensor-product Hilbert space for systems B and C. An important example of such a channel is one that mixes an input mode with two vacuum states at two beamsplitters such that the two receivers obtain two of the beamsplitter outputs while the environment obtains the other mode [20], [21]. One could incorporate noise effects in addition to loss by mixing with thermal states instead of vacuum states.

The class of local operations and classical communication (LOCC) consists of compositions of the following operations:

1) Alice performs a quantum instrument, which has both a quantum and classical output. She forwards the classical output to Bob, who then performs a quantum channel conditioned on the classical data received.
2) The situation is reversed, with Bob performing the initial instrument, who forwards the classical data to Alice, who then performs a quantum channel conditioned on the classical data.

An LOCC measurement is a measurement channel that can be implemented by means of local operations and classical communication. Every LOCC channel can be written in the following form (as a separability preserving channel):

$$\mathcal{L}_{AB}(\rho_{AB}) = \sum_x (C_A^x \otimes D_B^x)\,\rho_{AB}\,(C_A^x \otimes D_B^x)^\dagger, \quad (3)$$

such that $\sum_x (C_A^x \otimes D_B^x)^\dagger (C_A^x \otimes D_B^x) = I_{AB}$. However, not every channel of the form above is an LOCC channel [22].

The trace norm of an operator A is defined as $\|A\|_1 \equiv \mathrm{Tr}[\sqrt{A^\dagger A}]$, and it induces the trace distance $\|\rho - \sigma\|_1$ as a measure of distinguishability between two quantum states ρ and σ. The von Neumann entropy of a state ρ on system A is equal to $S(A)_\rho \equiv S(\rho) \equiv -\mathrm{Tr}[\rho \log \rho]$. The quantum mutual information of a bipartite state ρ_{AB} is defined as

$$I(A;B)_\rho \equiv S(A)_\rho + S(B)_\rho - S(AB)_\rho. \quad (4)$$

The LOCC accessible information of a classical-quantum-quantum state ρ_{MBC} is as follows:

$$I_{\mathrm{acc,LOCC}}(M;BC)_\rho \equiv \max_{\mathcal{L}_{BC\to\hat{M}}\in\mathrm{LOCC}} I(M;\hat{M})_\omega, \quad (5)$$

where the mutual information on the RHS is with respect to the following classical-classical state:

$$\omega_{M\hat{M}} \equiv \mathcal{L}_{BC\to\hat{M}}(\rho_{MBC}), \quad (6)$$

and $\mathcal{L}_{BC\to\hat{M}}$ is an LOCC measurement channel.

III. Quantum data hiding capacity

In a quantum data-hiding protocol, the sender Alice communicates classical or quantum information to two spatially separated receivers Bob and Charlie via a quantum broadcast channel $\mathcal{N}_{A\to BC}$. The protocol satisfies the "correctness property" if Bob and Charlie can decode reliably when allowed to apply a joint quantum measurement (i.e., if they are located in the same laboratory). On the other hand, the "data-hiding/security property" is that the transmitted information cannot be accessed by Bob and Charlie when they are restricted to performing local operations and classical communication (i.e., if they are in different laboratories connected by a classical communication channel). When Alice sends classical information, we call this a *bit-hiding protocol* and define it as follows:

Definition 1 (Bit-hiding protocol): An $(n, M, \delta, \varepsilon)$ bit-hiding protocol for a quantum broadcast channel $\mathcal{N}_{A\to BC}$ consists of a collection of input states $\{\rho(x)\}_{x=1,\ldots,M}$, and a decoding measurement satisfying the following properties:

- *(Correctness)* It is possible to decode with high average success probability, that is,

$$\frac{1}{M}\sum_x \mathrm{Tr}\left[\Lambda_{B^nC^n}^x \mathcal{N}_{A\to BC}^{\otimes n}(\rho(x))\right] \geq 1-\delta, \quad (7)$$

where $\{\Lambda_{B^nC^n}^x\}$ is the POVM associated to the decoding measurement.

- *(Security)* For all LOCC measurements $\mathcal{M}_{\mathrm{LOCC}}$ on the bipartite system B^nC^n, there exists a state $\sigma_{B^nC^n}$ such that

$$\frac{1}{M}\sum_x \left\|\mathcal{M}_{\mathrm{LOCC}}(\mathcal{N}_{A\to BC}^{\otimes n}(\rho(x)) - \sigma_{B^nC^n})\right\|_1 \leq \varepsilon. \quad (8)$$

Remark 1: Due to the convexity of the trace norm, it is sufficient to prove the security property against rank-one LOCC measurements.

We can now define the bit-hiding capacity of a quantum broadcast channel $\mathcal{N}_{A \to BC}$:

Definition 2 (Bit-hiding capacity): A bit-hiding rate R is achievable for a quantum broadcast channel $\mathcal{N}_{A \to BC}$ if for all $\delta, \varepsilon, \zeta \in (0,1)$ and sufficiently large n, there exists an $(n, M, \delta, \varepsilon)$ bit-hiding protocol such that $\frac{1}{n} \log M \geq R - \zeta$. The bit-hiding capacity $\kappa(\mathcal{N})$ of $\mathcal{N}$ is equal to the supremum of all achievable bit-hiding rates.

In this paper we focus on bit-hiding protocols. One can also define qubit-hiding protocols, aimed at hiding quantum information (see [11]). Similarly, one could define a notion of qubit-hiding capacity of a quantum channel. Additionally, these definitions immediately generalize to the case of a channel with one sender and an arbitrary number of receivers.

IV. REGULARIZED UPPER BOUND ON QUANTUM DATA HIDING CAPACITY

We begin by establishing a regularized upper bound on the bit-hiding capacity of any channel:

Theorem 1: The bit-hiding capacity $\kappa(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is bounded from above as follows:

$$\kappa(\mathcal{N}) \leq \lim_{n \to \infty} \frac{1}{n} \kappa^{(u)}(\mathcal{N}^{\otimes n}), \qquad (9)$$

where

$$\kappa^{(u)}(\mathcal{N}) \equiv \max_{\{p_X(x), \rho_x\}} \left[I(X; BC)_\rho - I_{\text{acc,LOCC}}(X; BC)_\rho \right],$$

$$\rho_{XBC} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A \to BC}(\rho_x). \qquad (10)$$

Proof: This upper bound follows by using the definition of a bit-hiding protocol and a few well known facts. Let us consider an $(n, M, \delta, \varepsilon)$ bit hiding protocol. By a simple reduction (see Appendix A), the correctness criterion in (7) translates to the following criterion:

$$\frac{1}{2} \left\| \overline{\Phi}_{MM'} - \omega_{MM'} \right\|_1 \leq \delta, \qquad (11)$$

where $\overline{\Phi}_{MM'}$ is a maximally correlated state, defined as

$$\overline{\Phi}_{MM'} \equiv \frac{1}{M} \sum_x |x\rangle\langle x|_M \otimes |x\rangle\langle x|_{M'}, \qquad (12)$$

and $\omega_{MM'}$ is defined from the probability distribution on the LHS of (7) as

$$\sum_{x,x'} \frac{1}{M} \text{Tr} \left[\Lambda_{B^n C^n}^{x'} \mathcal{N}_{A \to BC}^{\otimes n}(\rho(x)) \right] |x\rangle\langle x|_M \otimes |x'\rangle\langle x'|_{M'}. \qquad (13)$$

Then

$$\log M = I(M; M')_{\overline{\Phi}} \qquad (14)$$
$$\leq I(M; M')_\omega + f(n, \delta) \qquad (15)$$
$$\leq I(M; B^n C^n)_\omega + f(n, \delta). \qquad (16)$$

The first inequality is an application of the Alicki-Fannes continuity of entropy inequality [23], with $f(n, \delta)$ a function such that $\lim_{\delta \to 0} \lim_{n \to \infty} \frac{1}{n} f(n, \delta) = 0$. The second

inequality is a consequence of the Holevo bound [24], where we have denoted the state after the channel (but before the decoding measurement) again by ω. As a consequence of the security criterion in (8) and the Alicki-Fannes inequality, we can conclude that

$$I_{\text{acc,LOCC}}(M; B^n C^n) \leq g(n, \varepsilon), \qquad (17)$$

with $g(n, \varepsilon)$ a function such that $\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} g(n, \varepsilon) = 0$. So this gives

$$\log M \leq I(M; B^n C^n) - I_{\text{acc,LOCC}}(M; B^n C^n) + f(n, \delta) + g(n, \varepsilon). \qquad (18)$$

We finally optimize over all possible channel inputs, divide by n, take the limit as $n \to \infty$ and then as $\varepsilon, \delta \to 0$ to establish the statement of the theorem. $\blacksquare$

V. UPPER BOUND FOR COHERENT-STATE DATA-HIDING PROTOCOLS

Here we show that data-hiding schemes making use of coherent-state encodings are highly limited in terms of the rate at which they can hide information. We assume that the mean input photon number is less than $N_S \in (0, \infty)$ and Alice is connected to Bob and Charlie by a bosonic broadcast channel [20], in which Alice has access to one input port of a beamsplitter, the vacuum is injected into the other input port, and Bob and Charlie have access to the outputs of the beamsplitter. The main idea behind the bound that we prove here is that the classical capacity of the pure-loss bosonic channel is limited from above by $g(N_S)$ [25], where $g(x) \equiv (x+1) \log_2(x+1) - x \log_2 x$. At the same time, if Bob and Charlie perform heterodyne detection on their outputs and coordinate their results, the rate at which they can decode information is equal to $\log_2(1 + N_S)$. Our proof of the following theorem makes this intuition rigorous. Notably, this bound is the same as that found in [13] for the strong locking capacity when using coherent-state encodings.

Theorem 2: The quantum data-hiding capacity of a bosonic broadcast channel when restricting to coherent-state encodings with mean photon number N_S is bounded from above by $g(N_S) - \log_2(1 + N_S) \leq \log_2(e) \approx 1.45$.

Proof: Consider a quantum data-hiding scheme consisting of coherent-state codewords: $\{|\alpha^n(x, k)\rangle\}_{x,k}$, where $|\alpha^n\rangle \equiv |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \cdots \otimes |\alpha_n\rangle$ and such that the mean photon number of the scheme is less than $N_S \in (0, \infty)$. Then the quantum codeword transmitted for message x is as follows:

$$\rho(x) \equiv \frac{1}{K} \sum_k |\alpha^n(x, k)\rangle\langle \alpha^n(x, k)|. \qquad (19)$$

So the classical-quantum state corresponding to the transmitted state is as follows:

$$\omega_{MKB^n C^n} \equiv$$
$$\frac{1}{MK} \sum_{x,k} |x\rangle\langle x|_M \otimes |k\rangle\langle k|_K \otimes |\sqrt{\eta}\alpha^n(x, k)\rangle \langle \sqrt{\eta}\alpha^n(x, k)|_{B^n}$$
$$\otimes \left|\sqrt{1 - \eta}\alpha^n(x, k)\right\rangle \left\langle \sqrt{1 - \eta}\alpha^n(x, k)\right|_{C^n}, \qquad (20)$$

where $\left|\sqrt{\eta}\alpha^n\right\rangle$ is a shorthand for $\left|\sqrt{\eta}\alpha_1\right\rangle \otimes \left|\sqrt{\eta}\alpha_2\right\rangle \otimes \cdots \otimes \left|\sqrt{\eta}\alpha_n\right\rangle$.

In order to obtain our upper bound, we suppose that Bob performs heterodyne detection, forwards the results to Charlie, who also performs heterodyne detection and coordinates the results to decode. Now begin from the upper bound in (18):

$$I(M;B^nC^n)_\omega - I_{\mathrm{acc,LOCC}}(M;B^nC^n)_\omega$$

$$\leq I(M;B^nC^n)_\omega - I_{\mathrm{het,LO}}(M;B^nC^n)_\omega \qquad (21)$$

$$= I(MK;B^nC^n)_\omega - I(K;B^nC^n|M)_\omega$$
$$\quad - [I_{\mathrm{het,LO}}(MK;B^nC^n)_\omega - I_{\mathrm{het,LO}}(K;B^nC^n|M)_\omega] \quad (22)$$

$$= I(MK;B^nC^n)_\omega - I_{\mathrm{het,LO}}(MK;B^nC^n)_\omega$$
$$\quad - [I(K;B^nC^n|M)_\omega - I_{\mathrm{het,LO}}(K;B^nC^n|M)_\omega] \qquad (23)$$

$$\leq I(MK;B^nC^n)_\omega - I_{\mathrm{het,LO}}(MK;B^nC^n)_\omega \qquad (24)$$

$$\leq \max_{p_Y(y)} [I(Y;B^nC^n)_\tau - I_{\mathrm{het,LO}}(Y;B^nC^n)_\tau] \qquad (25)$$

$$\leq n \max_{p_Y(y)} [I(Y;BC)_\sigma - I_{\mathrm{het,LO}}(Y;BC)_\sigma]. \qquad (26)$$

The first inequality follows by picking the LOCC measurement to be the classically coordinated heterodyne detection mentioned above. The first equality follows from the chain rule for conditional mutual information, and the second equality is a rewriting. The second inequality follows because $I(K;B^nC^n|M)_\omega - I_{\mathrm{het,LO}}(K;B^nC^n|M)_\omega \geq 0$, which is a consequence of data processing: the classical heterodyne detection measurement outcomes are from measuring systems B^nC^n. The second-to-last inequality follows by optimizing over all input distributions and the information quantities are evaluated with respect to a state of the following form:

$$\tau_{YB^nC^n} \equiv \sum_y p_Y(y)|y\rangle\langle y|_Y \otimes |\sqrt{\eta}\alpha^n(y)\rangle \langle\sqrt{\eta}\alpha^n(y)|_{B^n}$$
$$\otimes \left|\sqrt{1-\eta}\alpha^n(y)\right\rangle \left\langle\sqrt{1-\eta}\alpha^n(y)\right|_{C^n}. \quad (27)$$

The last inequality follows by realizing that the difference between the mutual informations can be understood as being equal to the private information of a quantum wiretap channel in which the state is prepared for the receiver while the heterodyned version of this state (a classical variable) is prepared for the eavesdropper. Such a quantum wiretap channel has pure product input states (they are coherent states) and it is degraded. Thus, we can apply Theorem 35 from [13] to conclude that this private information is subadditive. The information quantities in the last line are with respect to a state of the following form:

$$\sigma_{YBC} \equiv \sum_y p_Y(y)|y\rangle\langle y|_Y \otimes |\sqrt{\eta}\alpha(y)\rangle \langle\sqrt{\eta}\alpha(y)|_B$$
$$\otimes \left|\sqrt{1-\eta}\alpha(y)\right\rangle \left\langle\sqrt{1-\eta}\alpha(y)\right|_C. \quad (28)$$

Now by a development nearly identical to that given in Eqs. (34)-(42) of [13], we can conclude that a circularly symmetric, Gaussian mixture of coherent states with variance N_S optimizes the quantity in (26). For such a distribution, the quantity $I(Y;BC)_\sigma$ evaluates to

$$I(Y;BC)_\sigma = g(N_S). \qquad (29)$$

We now evaluate the quantity $I_{\mathrm{het,LO}}(Y;BC)_\sigma$. Consider that the output for Bob is a random variable $\sqrt{\eta}\alpha + z_B$, where α is a zero-mean Gaussian random variable (RV) with variance N_S and z_B is a zero-mean Gaussian RV with variance 1. The output for Charlie is a RV $\sqrt{1-\eta}\alpha + z_C$, where α is the same Gaussian RV as above and z_C is a zero-mean Gaussian RV with variance 1 (note that α, z_B, and z_C are independent RVs). The covariance matrix for the real part of these RVs is as follows:

$$\begin{bmatrix} \eta N_S/2 + 1/2 & \sqrt{\eta(1-\eta)}N_S/2 \\ \sqrt{\eta(1-\eta)}N_S/2 & (1-\eta)N_S/2 + 1/2 \end{bmatrix}, \qquad (30)$$

with determinant equal to

$$\frac{1}{4}(N_S + 1). \qquad (31)$$

The determinant of the covariance matrix of the real parts of the RVs z_B and z_C is equal to $1/4$. By modeling the real and imaginary components as two independent parallel channels and plugging into the Shannon formula for the capacity of a Gaussian channel, we find that

$$I_{\mathrm{het,LO}}(Y;BC)_\sigma = \log(N_S + 1). \qquad (32)$$

So we can finally conclude the upper bound on the data hiding capacity of coherent-state schemes. ∎

VI. LOWER BOUND ON QUANTUM DATA HIDING CAPACITY

Here we prove a lower bound for the bit-hiding capacity of mictodiactic quantum broadcast channels. We consider a channel $\mathcal{N}_{A\to BC}$ from an input quantum system of dimensions $\dim A = d_A$ to the output systems with dimensions $\dim B = d_B$ and $\dim C = d_C$. A mictodiactic channel is defined from the following property:

$$\mathcal{N}_{A\to BC}(I_A/d_A) = I_{BC}/(d_B d_C), \qquad (33)$$

where I_A and I_{BC} denote the identity operators acting on the input and output spaces, respectively.

Theorem 3: The bit-hiding capacity $\kappa(\mathcal{N})$ of a mictodiactic broadcast channel $\mathcal{N}_{A\to BC}$, with $\dim A = d_A$, $\dim B = d_B$ and $\dim C = d_C$ is bounded from below as follows:

$$\kappa(\mathcal{N}) \geq \kappa^{(l)}(\mathcal{N}) \equiv \chi(\mathcal{N}) - \log d_+ - \log\gamma. \qquad (34)$$

where

$$\chi(\mathcal{N}) = S\left(\int d\psi\, \mathcal{N}(\psi)\right) - \int d\psi\, S(\mathcal{N}(\psi)) \qquad (35)$$

is the Holevo information of the channel computed for a uniform ensemble of input states ($d\psi$ denotes the uniform measure on the sphere of unit vectors), and $d_+ = \max\{d_B, d_C\}$. The parameter γ is given by

$$\gamma = \frac{2d_B^2 d_C^2}{d_A(d_A+1)}\, \|(\mathcal{N}\otimes\mathcal{N})(P_{\mathrm{sym}})\|_\infty, \qquad (36)$$

where $\|\cdot\|_\infty$ is the ∞-norm and P_{sym} is the projector onto the symmetric subspace.

A proof is given in the next section.

As an example consider the d-dimensional depolarizing channel, with $d = d_A = d_B d_C$:

$$\mathcal{N}(X) = pX + (1-p)\operatorname{Tr}(X)I/d. \qquad (37)$$

Let us first compute the Holevo information for the uniform distribution of input states. A straightforward calculation yields

$$\chi(\mathcal{N}) = \log d + \left(p + \frac{1-p}{d}\right)\log\left(p + \frac{1-p}{d}\right)$$
$$+ (d-1)\left(\frac{1-p}{d}\right)\log\left(\frac{1-p}{d}\right). \qquad (38)$$

Then consider the density matrix $\frac{P_{\text{sym}}}{\operatorname{Tr}(P_{\text{sym}})} = 2\frac{P_{\text{sym}}}{d(d+1)}$, which is mapped by $(\mathcal{N} \otimes \mathcal{N})$ into

$$(\mathcal{N} \otimes \mathcal{N})\left(\frac{P_{\text{sym}}}{\operatorname{Tr}(P_{\text{sym}})}\right) = p^2 \frac{P_{\text{sym}}}{\operatorname{Tr}(P_{\text{sym}})} + (1-p^2)\frac{I \otimes I}{d^2}, \qquad (39)$$

where we have used the fact that the partial trace of $\frac{P_{\text{sym}}}{\operatorname{Tr}(P_{\text{sym}})}$ is the maximally mixed state I/d. We hence have

$$\|(\mathcal{N} \otimes \mathcal{N})(P_{\text{sym}})\|_\infty$$
$$= \operatorname{Tr}(P_{\text{sym}})\left\|(\mathcal{N} \otimes \mathcal{N})\left(\frac{P_{\text{sym}}}{\operatorname{Tr}(P_{\text{sym}})}\right)\right\|_\infty \qquad (40)$$
$$= \frac{d(d+1)}{2}\left(\frac{2p^2}{d(d+1)} + \frac{1-p^2}{d^2}\right) \qquad (41)$$
$$= \frac{d+1}{2d} + p^2\left(1 - \frac{d+1}{2d}\right). \qquad (42)$$

Finally, from (36), putting $d_B d_C = d$ we obtain

$$\gamma = 1 + p^2\left(\frac{2d}{d+1} - 1\right). \qquad (43)$$

For the case of unital channels, which are contractive with respect to the Schatten norms [26], we have $\|(\mathcal{N} \otimes \mathcal{N})(P_{\text{sym}})\|_\infty \le \|P_{\text{sym}}\|_\infty = 1$, from which we obtain the following looser bound:

$$\kappa^{(l)}(\mathcal{N}) \ge \chi(\mathcal{N}) - \log d_+ - \log\frac{2d}{d+1}. \qquad (44)$$

A. Proof of Theorem 3

We now provide a proof for Theorem 3. We generate a bit-hiding code for n uses of the quantum broadcast channel $\mathcal{N}_{A \to BC}$, by employing the well known method of random coding. For $x = 1, \ldots, M$, pick a collection of M pure quantum states at random, each having the following form:

$$|\psi(x)\rangle = \bigotimes_{j=1}^n |\psi_j(x)\rangle. \qquad (45)$$

These random unit vectors are each generated by sampling $|\psi_j(x)\rangle$ independently from the uniform distribution over the unit sphere in $\mathbb{C}^d$. A simple way of doing so is to single out a fiducial unit vector $|0\rangle \in \mathbb{C}^d$ and pick a unitary V_{xj} at random according to the Haar measure on $SU(d)$, in order to generate $|\psi_j(x)\rangle = V_{xj}|0\rangle$. Let

$$V_x = \bigotimes_{j=1}^n V_{xj}. \qquad (46)$$

We then pick Kn qudit unitaries U_{kj} i.i.d. from the Haar measure on $SU(d)$. The inputs to the channel have the following form:

$$\rho(x) = \frac{1}{K}\sum_{k=1}^K U_k\,\psi(x)\,U_k^\dagger, \qquad (47)$$

where

$$U_k = \bigotimes_{j=1}^n U_{kj}. \qquad (48)$$

We introduce the notation $\mathbf{V} = (V_1, \ldots, V_M)$ to denote the M-tuple of n-local qudit unitaries and the notation $\mathbf{U} = (U_1, \ldots, U_K)$ to denote the K-tuple of n-local qudit unitaries, and we further define $\mathcal{R}_{\mathbf{U}}$ as follows:

$$\mathcal{R}_{\mathbf{U}}(\sigma) = \frac{1}{K}\sum_{k=1}^K U_k\,\sigma\,U_k^\dagger. \qquad (49)$$

We also define the random variables $\mathcal{V}$ and $\mathcal{U}$, taking values on the set of M-tuples $\mathbf{V} = (V_1, \ldots, V_M)$ and K-tuples $\mathbf{U} = (U_1, \ldots, U_K)$, respectively, where each qudit unitary is statistically independent and uniformly distributed according to the distribution induced by the Haar measure on $SU(d)$.

Given the input codewords in (47), the state at the output of n channel uses is as follows:

$$\mathcal{N}^{\otimes n}(\rho(x)) = \mathcal{N}^{\otimes n}(\mathcal{R}_{\mathbf{U}}(\psi(x))) \qquad (50)$$
$$= \mathcal{N}^{\otimes n}(\mathcal{R}_{\mathbf{U}}(V_x|0\rangle\langle 0|V_x)). \qquad (51)$$

We split the proof into two parts: first we prove the correctness property and then the security property.

1) Correctness: The proof of the correctness property has a rather standard form, but we provide it for completeness. Let $\Pi_{\mathcal{N}(\pi)}^{n,\delta_0}$ denote the δ_0-weakly typical projection for the tensor-power state $[\mathcal{N}(\pi)]^{\otimes n}$, where π denotes the maximally mixed state and $\delta_0 > 0$ (see, e.g., [27] for a definition and properties). This projector has the following properties which hold for all $\varepsilon_0 \in (0,1)$ and sufficiently large n:

$$\operatorname{Tr}(\Pi_{\mathcal{N}(\pi)}^{n,\delta_0}[\mathcal{N}(\pi)]^{\otimes n}) \ge 1 - \varepsilon_0, \qquad (52)$$

$$\Pi_{\mathcal{N}(\pi)}^{n,\delta_0}[\mathcal{N}(\pi)]^{\otimes n}\Pi_{\mathcal{N}(\pi)}^{n,\delta_0} \le 2^{-[S(\mathcal{N}(\pi))-\delta_0]}\Pi_{\mathcal{N}(\pi)}^{n,\delta_0}. \qquad (53)$$

Let a spectral decomposition of the state $\mathcal{N}^{\otimes n}(U_k\,\psi(x)\,U_k^\dagger)$ be as follows:

$$\mathcal{N}^{\otimes n}(U_k\,\psi(x)\,U_k^\dagger) = \sum_{y^n} p(y^n|k,x)\,|y_{k,x}^n\rangle\langle y_{k,x}^n|, \qquad (54)$$

where the eigenvalues $p(y^n|k,x)$ form a product distribution because the state is tensor product. Let $\Pi_{k,x}^{n,\delta_1}$ denote the δ_1-weak conditionally typical projection corresponding to the state $\mathcal{N}^{\otimes n}(U_k\,\psi(x)\,U_k^\dagger)$, defined as the projection onto the following subspace:

$$\operatorname{span}\left\{|y_{k,x}^n\rangle : \left|-\frac{1}{n}\log p(y^n|k,x) - S_*\right| \le \delta_1\right\}. \qquad (55)$$

where $\delta_1 > 0$ and $S_* = \int d\psi\, S(\mathcal{N}(\psi))$ (see, e.g., [27] for more details). These projectors have the following property, which holds for all $\varepsilon_1 \in (0,1)$ and sufficiently large n:

$$\mathbb{E}_{U_k,V_x}\left\{\operatorname{Tr}\left(\Pi_{k,x}^{n,\delta_1}\mathcal{N}^{\otimes n}(U_k\psi(x)U_k^\dagger)\right)\right\} \ge 1 - \varepsilon_1. \qquad (56)$$

(In fact, both ε_0 and ε_1 can be taken to be exponentially decreasing in n, following from an application of the Chernoff bound.) So this means that we can build a reliable decoding measurement from the set $\{\Pi_{k,x}\}$ of projectors. Another property of the conditionally typical projector is the following bound on the size of each projector:

$$\mathrm{Tr}(\Pi_{k,x}^{n,\delta_1}) \leq 2^{n[S_*+\delta_1]}. \tag{57}$$

We then consider a "square-root" decoding measurement defined by the following POVM elements:

$$\Lambda_{k,x} \equiv \left(\sum_{k',x'} \Gamma_{k',x'}\right)^{-1/2} \Gamma_{k,x} \left(\sum_{k',x'} \Gamma_{k',x'}\right)^{-1/2}, \tag{58}$$

$$\Gamma_{k,x} \equiv \Pi_{\mathcal{N}(\pi)}^{n,\delta_0} \Pi_{k,x}^{n,\delta_1} \Pi_{\mathcal{N}(\pi)}^{n,\delta_0}. \tag{59}$$

The average probability of an error when decoding both x and k is as follows:

$$\bar{p}_{\mathrm{err}} = 1 - \frac{1}{MK} \sum_{k,x} \mathrm{Tr}(\Lambda_{k,x} \mathcal{N}^{\otimes n}(\psi_k(x))), \tag{60}$$

where we have introduced the shorthand $\psi_k(x) \equiv U_k \psi(x) U_k^\dagger$. Recall the Hayashi-Nagooka operator inequality [28]

$$I - (P+Q)^{-1/2} P (P+Q)^{-1/2} \leq 2(I-P) + 4Q, \tag{61}$$

which holds for $0 \leq P \leq I$ and $Q \geq 0$. Picking

$$P = \Gamma_{k,x}, \qquad Q = \sum_{(k',x')\neq(k,x)} \Gamma_{k',x'}, \tag{62}$$

we can apply this to (60) to bound it from above as

$$\bar{p}_{\mathrm{err}} \leq \frac{2}{MK} \sum_{k,x} \mathrm{Tr}([I - \Gamma_{k,x}] \mathcal{N}^{\otimes n}(\psi_k(x)))$$
$$+ \frac{4}{MK} \sum_{k,x} \sum_{(k',x')\neq(k,x)} \mathrm{Tr}(\Gamma_{k',x'} \mathcal{N}^{\otimes n}(\psi_k(x))). \tag{63}$$

Now taking an expectation over a random choice of both $\mathcal{V}$ and $\mathcal{U}$, by (52) and (56) and the gentle measurement lemma [29], [30] (see also [27]), it follows that

$$\mathbb{E}_{\mathcal{V},\mathcal{U}} \left[\mathrm{Tr}([I - \Gamma_{k,x}] \mathcal{N}^{\otimes n}(\psi_k(x)))\right] \leq 2\sqrt{\varepsilon_0} + \varepsilon_1. \tag{64}$$

Recalling that each unitary is independent of another, for $(k',x') \neq (k,x)$, we find that

$$\mathbb{E}_{\mathcal{V},\mathcal{U}} \left[\mathrm{Tr}(\Gamma_{k',x'} \mathcal{N}^{\otimes n}(\psi_k(x)))\right]$$
$$= \mathrm{Tr}(\mathbb{E}_{\mathcal{V},\mathcal{U}} [\Gamma_{k',x'}] \mathbb{E}_{\mathcal{V},\mathcal{U}} [\mathcal{N}^{\otimes n}(\psi_k(x))]) \tag{65}$$
$$= \mathrm{Tr}(\mathbb{E}_{\mathcal{V},\mathcal{U}} [\Gamma_{k',x'}] [\mathcal{N}(\pi)]^{\otimes n}) \tag{66}$$
$$= \mathbb{E}_{\mathcal{V},\mathcal{U}} \mathrm{Tr}(\Pi_{k',x'}^{n,\delta_1} \Pi_{\mathcal{N}(\pi)}^{n,\delta_0} [\mathcal{N}(\pi)]^{\otimes n} \Pi_{\mathcal{N}(\pi)}^{n,\delta_0}) \tag{67}$$
$$\leq 2^{-[S(\mathcal{N}(\pi))-\delta_0]} \mathrm{Tr}(\Pi_{k',x'}^{n,\delta_1} \Pi_{\mathcal{N}(\pi)}^{n,\delta_0}) \tag{68}$$
$$\leq 2^{-[S(\mathcal{N}(\pi))-\delta_0]} \mathrm{Tr}(\Pi_{k',x'}^{n,\delta_1}) \tag{69}$$
$$\leq 2^{-[S(\mathcal{N}(\pi))-\delta_0]} 2^{n[S_*+\delta_1]}. \tag{70}$$

Putting everything together, we find that

$$\mathbb{E}_{\mathcal{V},\mathcal{U}} [\bar{p}_{\mathrm{err}}] \leq 2[2\sqrt{\varepsilon_0} + \varepsilon_1]$$
$$+ 4KM 2^{-[S(\mathcal{N}(\pi))-\delta_0]} 2^{n[S_*+\delta_1]}. \tag{71}$$

By picking

$$\frac{1}{n} \log M + \frac{1}{n} \log K = S(\mathcal{N}(\pi)) - S_* - 2\delta_0 - 2\delta_1, \tag{72}$$

we can ensure that

$$\mathbb{E}_{\mathcal{V},\mathcal{U}} [\bar{p}_{\mathrm{err}}] \leq 2[2\sqrt{\varepsilon_0} + \varepsilon_1] + 4 \cdot 2^{-n(\delta_0+\delta_1)}. \tag{73}$$

Eventually, we will derandomize the random bit-hiding code in order to conclude the existence of one with this error probability bound.

2) Security: To simplify the notation we set $d = d_B d_C$. To prove the security of the protocol against an LOCC measurement we show that

$$\frac{1}{M} \sum_x \left\|\mathcal{M}_{\mathrm{LOCC}} \left(\mathcal{N}_{A\to BC}^{\otimes n}(\rho(x)) - I^{\otimes n}/d^n\right)\right\|_1 \leq \varepsilon_2, \tag{74}$$

for an arbitrarily small $\varepsilon_2 \in (0,1)$ and for all LOCC measurements. In fact, what we show is an even stronger bound by proving that the protocol is secure against all separable measurements $\mathcal{M}_{\mathrm{sep}}$.

For a given codeword $\psi(x)$ and a separable measurement $\mathcal{M}_{\mathrm{sep}}$ (with POVM elements $\{B^y \otimes C^y\}$), we define the conditional random variable $Y|\mathcal{U}$ from the following distribution:

$$p_{Y|\mathbf{U}}(y) = \mathrm{Tr}[(B^y \otimes C^y) \mathcal{N}^{\otimes n}(\mathcal{R}_{\mathbf{U}}(\psi(x)))]. \tag{75}$$

We then consider the mutual information of the random variables Y and $\mathcal{U}$, given that codeword $\psi(x)$ was transmitted:

$$I(Y;\mathcal{U}) = H(Y) - H(Y|\mathcal{U}). \tag{76}$$

We first consider the case of separable projective measurements $\mathcal{M}_{\mathrm{pro\text{-}sep}}$, which have separable POVM elements $B^y \otimes C^y = \phi_B^y \otimes \phi_C^y$, where ϕ_B^y and ϕ_C^y are rank-one projectors acting on Bob and Charlie's systems, respectively. By applying concentration inequalities to the conditional probability distribution $p_{Y|\mathbf{U}}(y)$, we show that (see Appendix B for details) for $K \gg K(n, d_B, d_C, \delta_2)$

$$\sup_{x, \mathcal{M}_{\mathrm{pro\text{-}sep}}} I(Y;\mathcal{U}) = O(\delta_2 \log d^n), \tag{77}$$

where $\delta_2 \in (0,1)$ and

$$K(n, d_B, d_C, \delta_2) = 8 d_+^n \gamma^n \delta_2^{-2} \log \frac{10 d^n}{\delta_2}, \tag{78}$$

with $d_+ = \max\{d_B, d_C\}$.

The Pinsker inequality (see, e.g., [31]) states that

$$\|p_{Y,\mathbf{U}} - p_Y p_{\mathbf{U}}\|_1 = \int d\mathbf{U} \sum_y |p_{Y|\mathbf{U}}(y) - p_Y(y)|$$
$$\leq \sqrt{2\ln 2 I(Y;\mathcal{U})}. \tag{79}$$

Using (77) and the Pinsker inequality we then obtain that for all M codewords $\psi(x)$ and $\mathcal{M}_{\mathrm{pro\text{-}sep}}$

$$\int d\mathbf{U} \left\|\mathcal{M}_{\mathrm{pro\text{-}sep}} \left(\mathcal{N}(\mathcal{R}_{\mathbf{U}}(\psi(x))) - I^{\otimes n}/d^n\right)\right\|_1$$
$$= \int d\mathbf{U} \sum_y |p_{Y|\mathbf{U}}(y) - p_Y(y)| \leq O(\sqrt{\delta_2 \log d^n}). \tag{80}$$

Taking an average over all messages and exchanging this average with the expectation over unitaries, we find that

$$\mathbb{E}_{\mathcal{V},\mathcal{U}}\left\{\frac{1}{M}\sum_x \left\|\mathcal{M}_{\text{pro-sep}}\left(\mathcal{N}(\mathcal{R}_{\mathbf{U}}(\psi(x)))-I^{\otimes n}/d^n\right)\right\|_1\right\}$$
$$\leq O(\sqrt{\delta_2 \log d^n}). \quad (81)$$

Then, in Appendix C we show how this result can be extended to the case of generic separable measurements by applying some techniques developed in [32].

Finally, by choosing

$$\frac{1}{n}\log K = \log d_+ + \log\gamma + \lambda\frac{\log n}{n}, \quad (82)$$

where $\lambda > 0$ is a positive constant, the condition $K > K(n, d_B, d_C, \delta_2)$ is satisfied and implies that the LHS of (81) decreases sub-exponentially in n.

To conclude the proof we combine the conditions (72) and (82) to obtain that as long as the rate satisfies the following condition

$$\frac{1}{n}\log M = S\left(\mathcal{N}(\pi)\right) - S_* - \log d_+$$
$$- \log\gamma - \lambda\frac{\log n}{n} - 2\delta_0 - 2\delta_1, \quad (83)$$

then (73) and (81) are satisfied. This in turn implies that there exist choices of the unitaries $(U_1, \ldots, U_K)$ and $(V_1, \ldots, V_M)$ such that these conditions are satisfied and thus there exists an $(n, M, \delta, \varepsilon)$ bit-hiding code such that δ and ε are vanishing with increasing n. The asymptotic rate of the generated sequence of codes is then as given in the statement of the theorem.

B. Multipartite Generalization

The result of Theorem 3 is readily generalized to a multipartite setting. Suppose for instance that Alice sends information from a system of dimension d_A through a mictodiactic channel $\mathcal{N}$ to ℓ receivers, $B_1, \ldots, B_\ell$, with Hilbert space dimensions $d_j = \dim B_j$, and we require security against LOCC measurements. In this case we would obtain the following achievable bit-hiding rate:

$$\chi(\mathcal{N}) - \log d_+ - \log\gamma, \quad (84)$$

with $d_+ = \max_j d_j$ and

$$\gamma = \frac{2\prod_j d_j^2}{d_A(d_A+1)}\left\|\mathcal{N}^{\otimes 2}(P_{\text{sym}})\right\|_\infty. \quad (85)$$

VII. DISCUSSION

The main contribution of this paper is to provide a formal definition of the data hiding capacity of a quantum channel and upper and lower bounds on this operational quantity. Our work thus initiates the study of "noisy data hiding," and we now suggest several directions for future study.

It could be possible to generalize Theorem 3 such that it applies to generic, non-mictodiactic, channels. Now we show how this generalization is obtained assuming two unproven

statements. We notice that the mictodiactic condition is used only to obtain (111), that is,

$$\mathbb{E}_{\mathcal{U}}(X_k) = \frac{1}{d^n}. \quad (86)$$

where

$$X_k = \text{Tr}\left(\phi^y \mathcal{N}^{\otimes n}(U_k\psi(x)U_k^\dagger)\right). \quad (87)$$

The vector ϕ^y corresponds to one of the POVM elements of a decoding measurement.

The first unproven statement is that, for n large, ϕ^y belongs to the typical subspace associated to $\mathcal{N}^{\otimes n}(\mathbb{E}_{\mathcal{U}}(U_k\psi(x)U_k^\dagger)) = \mathcal{N}^{\otimes n}(\pi_A^{\otimes n}) = \mathcal{N}(\pi_A)^{\otimes n}$, where $\pi_A = I_A/d_A$ denotes the maximally mixed state on system A. Let $\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}$ then denote the projector onto the typical subspace. Putting $\phi^y = \Pi_{\mathcal{N}(\pi_A)}^{n,\delta}\phi^y\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}$, we have

$$\mathbb{E}_{\mathcal{U}}(X_k) = \text{Tr}\left(\phi^y\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}\mathcal{N}^{\otimes n}(\pi_A)\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}\right). \quad (88)$$

The equipartition property of the typical subspace (see e.g. [27]) yields

$$2^{-n(S+\delta)}\Pi_{\mathcal{N}(\pi_A)}^{n,\delta} \leq \Pi_{\mathcal{N}(\pi_A)}^{n,\delta}\mathcal{N}^{\otimes n}(\pi_A)\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}$$
$$\leq 2^{-n(S-\delta)}\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}, \quad (89)$$

with $S = S(\mathcal{N}(\pi_A))$. This in turn implies

$$2^{-n(S+\delta)} = 2^{-n(S+\delta)}\text{Tr}\left(\phi^y\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}\right) \leq \mathbb{E}_{\mathcal{U}}(X_k)$$
$$\leq 2^{-n(S-\delta)}\text{Tr}\left(\phi^y\Pi_{\mathcal{N}(\pi_A)}^{n,\delta}\right) = 2^{-n(S-\delta)}. \quad (90)$$

These bounds generalize the condition (111), where 2^S replaces $d_B d_C$ and plays the role of the effective dimension of the output system.

Similarly, consider the second moment, see (115),

$$\mathbb{E}_{\mathcal{U}}(X_k^2) = \text{Tr}\left[\phi^{y\otimes 2}\left(\mathcal{N}^{\otimes 2}\left(\frac{P_{\text{sym}}}{\text{Tr}(P_{\text{sym}})}\right)\right)^{\otimes n}\right]. \quad (91)$$

The second unproven statement is that the measurement vectors are of the form $\phi^{y\otimes 2} = \Pi_{\mathcal{N}^{\otimes 2}(P_{\text{sym}})}^{n,\delta}\phi^{y\otimes 2}\Pi_{\mathcal{N}^{\otimes 2}(P_{\text{sym}})}^{n,\delta}$, where $\Pi_{\mathcal{N}^{\otimes 2}(P_{\text{sym}})}^{n,\delta}$ is the typical projector associated to the state $\mathcal{N}^{\otimes 2}\left(\frac{P_{\text{sym}}}{\text{Tr}P_{\text{sym}}}\right)$. The equipartition property implies

$$\text{Tr}\left[\phi^{y\otimes 2}\Pi_{\mathcal{N}^{\otimes 2}(P_{\text{sym}})}^{n,\delta}\left(\mathcal{N}^{\otimes 2}\left(\frac{P_{\text{sym}}}{\text{Tr}P_{\text{sym}}}\right)\right)^{\otimes n}\Pi_{\mathcal{N}^{\otimes 2}(P_{\text{sym}})}^{n,\delta}\right]$$
$$\leq 2^{-n(S_2-\delta)}, \quad (92)$$

where $S_2 = S(\mathcal{N}^{\otimes 2}(\frac{P_{\text{sym}}}{\text{Tr}P_{\text{sym}}}))$.

In conclusion, by modifying (116), it could be possible to extend Theorem 3 to non-mictodiactic (and non-unital) channels with γ redefined as

$$\left(\frac{\mathbb{E}_{\mathcal{U}}(X_k^2)}{\mathbb{E}_{\mathcal{U}}(X_k)^2}\right)^{\frac{1}{n}} \leq \left(\frac{2^{-n(S_2-\delta)}}{2^{-2n(S+\delta)}}\right)^{\frac{1}{n}} = 2^{2S-S_2+3\delta} =: \gamma. \quad (93)$$

Notice that this extension of Theorem 3 also improves the bound for mictodiactic channels. We leave a full development of the above observations for future work.

There are several ways in which the lower bound of Theorem 3 could be further improved. Let us recall that a map $\mathcal{R}$ is said to be an approximately randomizing map with respect to a given norm $\|\cdot\|_*$, if for all states ψ,

$$\|\mathcal{R}(\psi) - \pi\|_* \leq \varepsilon, \tag{94}$$

where π is the maximally mixed state. There exists a close relation between data-hiding protocols and approximately randomizing maps. For instance, the data-hiding protocol of [11] was obtained by modifying a related approximately randomizing map with respect to the operator norm $\|\cdot\|_\infty$. Several constructions of approximately randomizing maps with respect to the operator norm and the trace norm were discussed in [33], [34], [35], [36]. It could indeed be possible that these maps can be used to develop quantum data hiding protocols robust to noise and achieving higher rates. If this is true, it might be possible to achieve a bit-hiding rate of

$$\chi(\mathcal{N}) - \log d_+. \tag{95}$$

APPENDIX A

Let M be a random variable. Let $p_{MM'}$ denote the distribution in which M' is a copy of M, so that

$$p_{MM'}(m, m') = p_M(m)\delta_{m,m'}. \tag{96}$$

Let $q_{MM'}$ denote the distribution that results from sending random variable M' through a channel $q_{M'|M}(m'|m)$, so that

$$q_{MM'}(m, m') = p_M(m)q_{M'|M}(m'|m). \tag{97}$$

Then the probability that $M' \neq M$ under $q_{MM'}$ is equal to the normalized trace distance between $p_{MM'}$ and $q_{MM'}$:

$$\Pr_q\{M' \neq M\} = \frac{1}{2}\|p_{MM'} - q_{MM'}\|_1. \tag{98}$$

Consider that

$$\Pr_q\{M'' \neq M\} = \sum_{m' \neq m} q_{MM'}(m, m') \tag{99}$$

$$= \sum_{m' \neq m} p_M(m)q_{M'|M}(m'|m). \tag{100}$$

Consider also that

$$\|p_{MM'} - q_{MM'}\|_1$$
$$= \sum_{m,m'} \left|p_M(m)\delta_{m,m'} - p_M(m)q_{M'|M}(m'|m)\right| \tag{101}$$

$$= \sum_{m,m'} p_M(m)\left|\delta_{m,m'} - q_{M'|M}(m'|m)\right| \tag{102}$$

$$= \sum_{m \neq m'} p_M(m)q_{M'|M}(m'|m)$$
$$+ \sum_m p_M(m)\left[1 - q_{M'|M}(m|m)\right] \tag{103}$$

$$= 2\sum_{m \neq m'} p_M(m)q_{M'|M}(m'|m). \tag{104}$$

So the equality in (98) follows from (99)–(100) and (101)–(104).

APPENDIX B
CONCENTRATION INEQUALITIES

In a crucial passage in the proof of Theorem 3 we have applied the bound

$$I(Y;\mathcal{U}) = O(\delta \log d^n), \tag{105}$$

holding for any separable measurement. Here and in Appendix C we prove that this bound holds true provided

$$K \gg K(n, d_B, d_C, \delta) = 8d_+^n \gamma^n \delta^{-2} \log \frac{10d^n}{\delta}, \tag{106}$$

with γ as in (36).

We do so by first showing that (105) holds for separable projective measurements $\mathcal{M}_{\text{pro-sep}}$, which have separable POVM elements $\phi_B^y \otimes \phi_C^y$, where ϕ_B^y and ϕ_C^y are rank-one projectors acting on Bob and Charlie's systems, respectively. This is done in Proposition 1, where we apply techniques analogous to those applied in [15], [16]. In Appendix C we show how this result can be extended to the case of generic separable measurements by applying some techniques developed in [32].

We will make use of two concentration inequalities. The first one is Maurer's tail bound [37]:

Theorem 4: Let $\{X_k\}_{k=1,\ldots,K}$ be K i.i.d. non-negative real-valued random variables, with $X_k \sim X$ and finite first and second moments: $\mathbb{E}[X], \mathbb{E}[X^2] < \infty$. Then, for any $\tau > 0$ we have that

$$\Pr\left\{\frac{1}{K}\sum_{k=1}^K X_k < \mathbb{E}[X] - \tau\right\} \leq \exp\left(-\frac{K\tau^2}{2\mathbb{E}[X^2]}\right). \tag{107}$$

The second one is the Chernoff bound [38]:

Theorem 5: Let $\{X_k\}_{k=1,\ldots,K}$ be K i.i.d. random variables, $X_k \sim X$, with $0 \leq X \leq 1$ and $\mathbb{E}[X] = \mu$. Then, for any $\tau > 0$ such that $(1+\tau)\mu \leq 1$ we have that

$$\Pr\left\{\frac{1}{K}\sum_{k=1}^K X_k > (1+\tau)\mu\right\} \leq \exp\left(-\frac{K\tau^2\mu}{4\ln 2}\right). \tag{108}$$

For given states $\psi(x)$ and $\phi^y = \phi_B^y \otimes \phi_C^y$, we apply these bounds to the quantity

$$X_k = \text{Tr}\left(\phi^y \mathcal{N}^{\otimes n}(U_k\psi(x)U_k^\dagger)\right). \tag{109}$$

This quantity is a random variable if the unitary U_k is chosen randomly. Taking the average over the random unitaries U_k we obtain

$$\mathbb{E}_{\mathcal{U}}(U_k \psi(x) U_k^{\dagger}) = \frac{I^{\otimes n}}{d_A^n}, \quad (110)$$

which for mictodiactic channels implies

$$\mathbb{E}_{\mathcal{U}}(X_k) = \frac{1}{d_B^n d_C^n}. \quad (111)$$

We also consider the second moment

$$\mathbb{E}_{\mathcal{U}}(X_k^2) = \int dU_k \operatorname{Tr}\left[\left(\phi^y \mathcal{N}^{\otimes n}(U_k \psi(x) U_k^{\dagger})\right)^{\otimes 2}\right]$$

$$= \operatorname{Tr}\left[\phi^{y \otimes 2} \bigotimes_{j=1}^{n} \mathcal{N}^{\otimes 2}\left(\int dU_{kj}\, U_{kj}^{\otimes 2} \psi_j(x)^{\otimes 2} U_{kj}^{\dagger \otimes 2}\right)\right]. \quad (112)$$

For any two-qudit pure state ρ we have

$$\int dU\, U^{\otimes 2} \rho^{\otimes 2} U^{\dagger \otimes 2} = \frac{2}{d_A(d_A+1)} P_{\text{sym}}, \quad (113)$$

where P_{sym} is the projector onto the symmetric subspace (see, e.g., [39]). We then obtain

$$\mathbb{E}_{\mathcal{U}}(X_k^2) = \left[\frac{2}{d_A(d_A+1)}\right]^n \operatorname{Tr}\left[\phi^{y \otimes 2}\left(\mathcal{N}^{\otimes 2}(P_{\text{sym}})\right)^{\otimes n}\right] \quad (114)$$

$$\leq \left[\frac{2}{d_A(d_A+1)}\right]^n \|\mathcal{N}^{\otimes 2}(P_{\text{sym}})\|_{\infty}^n \quad (115)$$

(where we have used the fact that the expectation value of an operator on a vector is upper bounded by the ∞-norm), from which it follows

$$\left(\frac{\mathbb{E}_{\mathcal{U}}(X_k^2)}{\mathbb{E}_{\mathcal{U}}(X_k)^2}\right)^{\frac{1}{n}} \leq \frac{2 d_B^2 d_C^2}{d_A(d_A+1)} \|\mathcal{N}^{\otimes 2}(P_{\text{sym}})\|_{\infty} = \gamma. \quad (116)$$

We will also make use of the notion of a net of unit vectors [11]. Let us recall that an ε-net is a finite set of unit vectors $N_\varepsilon = \{\tilde{\phi}\}$ in a D-dimensional Hilbert space such that for any unit vector ϕ there exists $\widetilde{\phi} \in N_\varepsilon$ for which

$$\|\phi - \widetilde{\phi}\|_1 \leq \varepsilon. \quad (117)$$

As discussed in [11] there exist ε-nets with less than $(5/\varepsilon)^{2D}$ elements. Let us consider a pair of vectors ϕ_B, ϕ_C and the corresponding vectors $\widetilde{\phi}_B$, $\widetilde{\phi}_C$ in the net such that

$$\|\phi_B - \widetilde{\phi}_B\|_1, \|\phi_C - \widetilde{\phi}_C\|_1 \leq \varepsilon. \quad (118)$$

We consider the quantity defined in (75):

$$p_{Y|\mathbf{U}}(y) = \operatorname{Tr}[(\phi_B^y \otimes \phi_C^y) \mathcal{N}^{\otimes n}(\mathcal{R}_{\mathbf{U}}(\psi(x)))], \quad (119)$$

and

$$\tilde{p}_{Y|\mathbf{U}}(y) = \operatorname{Tr}[(\tilde{\phi}_B^y \otimes \tilde{\phi}_C^y) \mathcal{N}^{\otimes n}(\mathcal{R}_{\mathbf{U}}(\psi(x)))]. \quad (120)$$

We have

$$\left|p_{Y|\mathbf{U}}(y) - \tilde{p}_{Y|\mathbf{U}}(y)\right| \leq 2\varepsilon. \quad (121)$$

Finally, we consider the function $\eta(\cdot) = -(\cdot)\log(\cdot)$. Using the inequality $|\eta(r) - \eta(s)| \leq \eta(|r-s|)$ (which holds for $|r-s| < 1/2$) we obtain

$$\left|\eta(p_{Y|\mathbf{U}}(y)) - \eta(\tilde{p}_{Y|\mathbf{U}}(y))\right| \leq \eta(2\varepsilon). \quad (122)$$

From now on, to make notation lighter, we put $d := d_B d_C$. We are now ready to prove the following:

Proposition 1: For any $\delta \in (0,1)$ and $K > K(n, d_B, d_C, \delta)$,

$$\sup_{x, \mathcal{M}_{\text{pro-sep}}} I(Y; \mathcal{U}) = O(\delta \log d^n), \quad (123)$$

where the sup is over separable projective measurements $\mathcal{M}_{\text{pro-sep}}$.

First of all, we notice that for a projective measurement with associated POVM elements $\{\phi^y = \phi_B^y \otimes \phi_C^y\}_{y=1,\ldots,d^n}$ the mutual information (76) reads

$$I(Y; \mathcal{U}) = \log d^n - H(Y|\mathcal{U}), \quad (124)$$

where we have also used the condition (111).

For a given vector $\psi(x)$ and separable projective measurement $\mathcal{M}_{\text{pro-sep}}$ we consider the quantity

$$H(Y)_{\mathbf{U}} = -\sum_y p_{Y|\mathbf{U}}(y) \log p_{Y|\mathbf{U}}(y) \quad (125)$$

$$= \sum_y \eta[p_{Y|\mathbf{U}}(y)]. \quad (126)$$

For a given y, we want to bound the probability that $H(Y)_{\mathbf{U}} \leq (1-\delta)\log d^n$. In order to do that we bound the probability that $\eta[p_{Y|\mathbf{U}}(y)] \leq \eta[\frac{1-\delta}{d^n}]$. Notice that $\eta[x]$ is a concave function and the equation $\eta[x] = \eta[\frac{1-\delta}{d^n}]$ has two roots: $x_- = \frac{1-\delta}{d^n}$, and $x_+ = 1 - \eta[\frac{1-\delta}{d^n}] + O(\eta[\frac{1-\delta}{d^n}])$, where for d^n large enough $x_+ \geq 1 - 2\eta[\frac{1-\delta}{d^n}]$. Therefore we have

$$\operatorname{Pr}_{\mathcal{U}}\left\{\eta[p_{Y|\mathbf{U}}(y)] \leq \frac{1-\delta}{d^n}\log d^n\right\}$$

$$\leq \operatorname{Pr}_{\mathcal{U}}\{p_{Y|\mathbf{U}}(y) \leq x_-\} + \operatorname{Pr}_{\mathcal{U}}\{p_{Y|\mathbf{U}}(y) \geq x_+\}$$

$$\leq \operatorname{Pr}_{\mathcal{U}}\left\{p_{Y|\mathbf{U}}(y) \leq \frac{1-\delta}{d^n}\right\}$$

$$+ \operatorname{Pr}_{\mathcal{U}}\{p_{Y|\mathbf{U}}(y) \geq 1 - \delta'\}, \quad (127)$$

where $\delta' \equiv 2\eta[\frac{1-\delta}{d^n}]$.

For given x and y, we now apply the Maurer tail bound (Theorem 4) to the random variables

$$X_k := \operatorname{Tr}\left(\phi_B^y \otimes \phi_C^y \mathcal{N}^{\otimes n}(U_k \psi(x) U_k^{\dagger})\right), \quad (128)$$

whose first and second moments are given by (111) and (112) and obey the inequality in (116). We remark that

$$\frac{1}{K}\sum_k X_k = p_{Y|\mathbf{U}}(y). \quad (129)$$

Hence, applying (107) with $\tau = \delta \mathbb{E}[X] = \delta/d^n$, we obtain

$$\operatorname{Pr}_{\mathcal{U}}\left\{p_{Y|\mathbf{U}}(y) < \frac{1-\delta}{d^n}\right\} \leq \exp\left(-\frac{K\delta^2 \mathbb{E}[X]^2}{2\mathbb{E}[X^2]}\right) \quad (130)$$

$$\leq \exp\left(-\frac{K\delta^2}{2\gamma^n}\right). \quad (131)$$

Similarly we apply the Chernoff inequality, Theorem 5, and obtain

$$\Pr_{\mathcal{U}}\left\{p_{Y|\mathbf{U}}(y) \geq 1 - \delta'\right\} \leq \exp\left(-\frac{K d^n (1 - \delta' - 1/d^n)^2}{4\ln 2}\right)$$
$$\leq \exp\left(-\frac{K d^n (1/2)^2}{4\ln 2}\right), \quad (132)$$

where the last inequality holds for sufficiently small δ' and large d^n. We then have

$$\Pr_{\mathcal{U}}\left\{\eta(p_{Y|\mathbf{U}}(y)) \leq \frac{1-\delta}{d^n}\log d^n\right\}$$
$$\leq \exp\left(-\frac{K\delta^2}{2\gamma^n}\right) + \exp\left(-\frac{K d^n(1/2)^2}{4\ln 2}\right) \quad (133)$$
$$\leq 2\exp\left(-\frac{K\delta^2}{2\gamma^n}\right), \quad (134)$$

where the last inequality holds for any $d^n > 8\ln 2\gamma^{-n}\delta^2$.

This is true for given x and vectors ϕ_B^y, ϕ_C^y. To account for all possible codewords $\psi(x)$ and measurement vectors ϕ_B^y, ϕ_C^y, we introduce a $(\delta 2^{-1} d^{-n})$-net for Bob's system and a $(\delta 2^{-1} d^{-n})$-net for Charlie's one, containing in total no more than $(10 d^n/\delta)^{2 d_B^n + 2 d_C^n}$ elements. Applying the union bound on the net's vectors and on the M codewords we obtain

$$\Pr\left\{\inf_{x,\tilde{\phi}_B^y,\tilde{\phi}_C^y}\eta(p_{Y|\mathbf{U}}(y)) \leq \frac{1-\delta}{d^n}\log d^n\right\}$$
$$\leq 2M\left(\frac{10 d^n}{\delta}\right)^{2 d_B^n + 2 d_C^n}\exp\left(-\frac{K\delta^2}{2\gamma^n}\right) \quad (135)$$
$$\leq 2M\left(\frac{10 d^n}{\delta}\right)^{4 d_+^n}\exp\left(-\frac{K\delta^2}{2\gamma^n}\right) \quad (136)$$
$$\leq \exp\left(-\frac{K\delta^2}{2\gamma^n} + 4 d_+^n\log\frac{10 d^n}{\delta} + \log(2M)\right) \quad (137)$$
$$=: p, \quad (138)$$

where $d_+ = \max\{d_B, d_C\}$. From this, we can extend the infimum over all unit vectors by paying a small penalty given by (122). In this way we obtain

$$\Pr\left\{\inf_{\phi_B^y,\phi_C^y}\eta(p_{Y|\mathbf{U}}(y)) \leq \frac{1-\delta}{d^n}\log d^n - \eta(\delta d^{-n})\right\} \leq p. \quad (139)$$

The probability of such a bad event can be made arbitrarily small provided (we are assuming $\log(2M) \ll 4 d_+^n\log\frac{10 d^n}{\delta}$)

$$K \gg 8\gamma^n d_+^n \delta^{-2}\log\frac{10 d^n}{\delta}. \quad (140)$$

Under this condition we then have that, up to a small probability p, for all separable projective measurements $\mathcal{M}_{\text{pro-sep}}$,

$$H(Y)_{\mathbf{U}} \geq (1-\delta)\log d^n - d^n\eta(\delta d^{-n}). \quad (141)$$

This implies

$$H(Y|\mathcal{U}) = \int d\mathbf{U}\, H(Y)_{\mathbf{U}} \quad (142)$$
$$\geq \int d\mathbf{U}\, H(Y)_{\mathbf{U}}\,\Theta \quad (143)$$
$$\geq \left[(1-\delta)\log d^n - d^n\eta(\delta d^{-n})\right](1-p), \quad (144)$$

where $\Theta \equiv \Theta\left[H(Y)_{\mathbf{U}} - (1-\delta)\log d^n + d^n\eta(\delta d^{-n})\right]$ denotes the Heaviside step function, which finally yields

$$\sup_{x,\mathcal{M}_{\text{pro-sep}}} I(Y;\mathcal{U}) \leq \delta\log d^n + d^n\eta(\delta d^{-n})$$
$$+ \left[(1-\delta)\log d^n - d^n\eta(\delta d^{-n})\right]p$$
$$= O\left(\delta\log d^n\right). \quad (145)$$

APPENDIX C
FROM PROJECTIVE MEASUREMENTS TO GENERIC SEPARABLE MEASUREMENTS

In this appendix we discuss how Proposition 1 can be extended to include generic separable measurements via the notion of *separable quasi-measurements*. In order to do that we apply a simple modification of Lemma 6.2 in [32].

The notion of quasi-measurement was defined in [32] :

Definition 3 (Quasi-Measurement [32]): We call (s,f) quasi-measurement an incomplete measurement on a D-dimensional system such that the associated POVM elements are of the form $\frac{D^2}{s}\phi^y$, where ϕ^y, for $y = 1,\ldots,s$, are rank-one projectors and $\sum_{y=1}^s\frac{D^2}{s}\phi^y \leq fI$. We denote as $\mathcal{L}(s,f)$ the set of (s,f)-separable quasi-measurements.

Lemma 6.2 in [32] proves that quasi-measurements are almost as informative as generic measurements, that is, given a bipartite quantum state ρ_{UB} we have

$$\sup_{\mathcal{M}}\|\mathcal{M}(\rho_{UB} - \rho_U \otimes \rho_B)\|_1$$
$$\leq \sup_{\mathcal{M}' \in \mathcal{L}(s,f)}\|\mathcal{M}'(\rho_{UB} - \rho_U \otimes \rho_B)\|_1$$
$$+ 4D^2 e^{-s(\eta-1)^2/(D(2\ln 2))}, \quad (146)$$

where the sup on left hand side is over generic measurements $\mathcal{M}_{B\to Y}$ on the B system, and that on the right hand side is over quasi-measurements $\mathcal{M}'_{B\to Y}$.

To apply this result to our setting, we first define the notion of separable quasi-measurements by adapting Definition 3 to the LOCC setting:

Definition 4: [Separable quasi-measurement] We say that an incomplete measurement on a $D_B \times D_C$-dimensional bipartite system is an (s,f)-separable quasi-measurement if the associated POVM elements have the form $\frac{(D_B D_C)^2}{s}\phi_B^y \otimes \phi_C^y$, where ϕ_B^y, ϕ_C^y, for $y = 1,\ldots,s$, are rank-one projectors, such that $\sum_{y=1}^s\frac{(D_B D_C)^2}{s}\phi_B^y \otimes \phi_C^y \leq fI_{BC}$. We denote as $\mathcal{L}_{\text{sep}}(s,f)$ the set of (s,f)-separable quasi-measurements.

We notice that the proof in Appendix B, which considers separable projective measurements, also applies to the case of separable quasi-measurements. This implies that, by considering the set of (s,f)-separable quasi-measurements, a modified version of Proposition 1 will be obtained, i.e.,

$$\sup_{x,\mathcal{M}\in\mathcal{L}_{\text{sep}}(s,f)} I(Y;\mathcal{U}) = O(\delta\log s). \quad (147)$$

To move from separable quasi-measurements to generic separable measurements we apply a straightforward modification

of Lemma 6.2 in [32] in the LOCC setting, that is,

$$\sup_{\mathcal{M}_{\text{sep}}} \|\mathcal{M}\left(\rho_{UBC} - \rho_U \otimes \rho_{BC}\right)\|_1$$
$$\leq \sup_{\mathcal{M}'sep \in \mathcal{L}_{\text{sep}}(s,f)} \|\mathcal{M}'\left(\rho_{UBC} - \rho_U \otimes \rho_{BC}\right)\|_1$$
$$+ 4(D_B D_C)^2 e^{-s(\eta-1)^2/(D_B D_C (2\ln 2))}, \quad (148)$$

where the $\sup$ on the left hand side is over separable measurements $\mathcal{M}_{B,C \to Y}^{\text{sep}}$, and the one on the right is over separable quasi-measurements $\mathcal{M}_{B,C \to Y}'^{\text{sep}}$.

In conclusion, to extend Proposition 1 to generic separable measurements we proceed as follows. First recall that (see Eqs. (79) and (80)) given a separable measurement $\mathcal{M}_{\text{sep}}$, in our setting we have

$$\|\mathcal{M}_{\text{sep}}\left(\rho_{UBC} - \rho_U \otimes \rho_{BC}\right)\|_1 =$$
$$\int d\mathbf{U} \left\|\mathcal{M}_{\text{sep}}\left(\mathcal{N}(\mathcal{R}_{\mathbf{U}}(\psi(x))) - I^{\otimes n}/d^n\right)\right\|_1. \quad (149)$$

Then consider that Proposition 1, together with (79), states that

$$\sup_{\mathcal{M}_{\text{pro}-\text{sep}}} \|\mathcal{M}_{\text{pro}-\text{sep}}\left(\rho_{UBC} - \rho_U \otimes \rho_{BC}\right)\|_1$$
$$\leq O(\sqrt{\delta \log d^n}), \quad (150)$$

where the $\sup$ is over projective separable measurements. As noticed above, Proposition 1 is straightforwardly extended from projective separable measurements to separable quasi-measurements, yielding

$$\sup_{\mathcal{M} \in \mathcal{L}_{\text{sep}}(s,f)} \|\mathcal{M}\left(\rho_{UBC} - \rho_U \otimes \rho_{BC}\right)\|_1 \leq O(\sqrt{\delta \log s}). \quad (151)$$

Finally, using (148) with $D_B = d_B^n$, $D_C = d_C^n$, we obtain

$$\sup_{\mathcal{M}_{\text{sep}}} \|\mathcal{M}\left(\rho_{UBC} - \rho_U \otimes \rho_{BC}\right)\|_1$$
$$\leq O(\sqrt{\delta \log s}) + 4(d_B d_C)^{2n} e^{-s(\eta-1)^2/(d_B^n d_C^n (2\ln 2))}. \quad (152)$$

Putting, for example, $s = d^{2n}$ and $f = 1/2$, yields

$$\sup_{\mathcal{M}_{\text{sep}}} \|\mathcal{M}\left(\rho_{UBC} - \rho_U \otimes \rho_{BC}\right)\|_1$$
$$\leq O(\sqrt{\delta \log d^{2n}}) + 4d^{2n} \exp\left(-\frac{d^n}{8\ln 2}\right). \quad (153)$$

This result can then be used to extend the security to the case of generic separable measurements by paying a small penalty not larger than $4d^{2n} \exp\left(-\frac{d^n}{8\ln 2}\right)$.

REFERENCES

[1] J. S. Bell, "On the Einstein-Podolsky-Rosen paradox," *Physics*, vol. 1, pp. 195–200, 1964.

[2] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers Systems and Signal Processing*, Bangalore, India, December 1984, pp. 175–179.

[3] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, vol. 70, no. 13, pp. 1895–1899, March 1993.

[4] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Physical Review Letters*, vol. 69, no. 20, pp. 2881–2884, November 1992.

[5] R. Raz, "Exponential separation of quantum and classical communication complexity," *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pp. 358–367, 1999.

[6] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal, "Locking classical correlations in quantum states," *Physical Review Letters*, vol. 92, no. 6, p. 067902, February 2004, arXiv:quant-ph/0303088.

[7] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, "Hiding bits in Bell states," *Physical Review Letters*, vol. 86, no. 25, pp. 5807–5810, June 2001, arXiv:quant-ph/0011042.

[8] T. Eggeling and R. F. Werner, "Hiding classical data in multi-partite quantum states," *Physical Review Letters*, vol. 89, no. 9, p. 097905, 2002, arXiv:quant-ph/0203004.

[9] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, "Quantum data hiding," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 580–598, 2002, arXiv:quant-ph/0103098.

[10] D. P. DiVincenzo, P. Hayden, and B. M. Terhal, "Hiding quantum data," *Foundations of Physics*, vol. 33, no. 11, pp. 1629–1647, 2003, arXiv:quant-ph/0207147.

[11] P. Hayden, D. Leung, P. W. Shor, and A. Winter, "Randomizing quantum states: Constructions and applications," *Communications in Mathematical Physics*, vol. 250, no. 2, pp. 371–391, July 2004, arXiv:quant-ph/0307104.

[12] P. Hayden, D. Leung, and G. Smith, "Multiparty data hiding of quantum information," *Physical Review A*, vol. 71, no. 6, p. 062339, June 2005, arXiv:quant-ph/0407152.

[13] S. Guha, P. Hayden, H. Krovi, S. Lloyd, C. Lupo, J. H. Shapiro, M. Takeoka, and M. M. Wilde, "Quantum enigma machines and the locking capacity of a quantum channel," *Physical Review X*, vol. 4, no. 1, p. 011016, January 2014, arXiv:1307.5368.

[14] C. Lupo, M. M. Wilde, and S. Lloyd, "Robust quantum data locking from phase modulation," *Physical Review A*, vol. 90, no. 2, p. 022326, August 2014, arXiv:1311.5212.

[15] C. Lupo and S. Lloyd, "Quantum-locked key distribution at nearly the classical capacity rate," *Physical Review Letters*, vol. 113, no. 16, p. 160502, October 2014, arXiv:1406.4418.

[16] ——, "Quantum data locking for high-rate private communication," *New Journal of Physics*, vol. 17, no. 3, p. 033022, March 2015, arXiv:1410.4748.

[17] ——, "Continuous-variable quantum enigma machines for long-distance key distribution," *Physical Review A*, vol. 92, no. 6, p. 062312, December 2015, arXiv:1501.07212.

[18] A. Winter, "Weak locking capacity of quantum channels can be much larger than private capacity," *Accepted in Journal of Cryptology*, 2015, arXiv:1403.6361.

[19] J. Yard, P. Hayden, and I. Devetak, "Quantum broadcast channels," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 7147–7162, October 2011, arXiv:quant-ph/0603098.

[20] S. Guha, J. H. Shapiro, and B. I. Erkmen, "Classical capacity of bosonic broadcast communication and a minimum output entropy conjecture," *Physical Review A*, vol. 76, no. 3, p. 032303, September 2007, arXiv:0706.3416.

[21] K. P. Seshadreesan, M. Takeoka, and M. M. Wilde, "Bounds on entanglement distillation and secret key agreement for quantum broadcast channels," *Accepted in IEEE Transactions on Information Theory*, March 2015, arXiv:1503.08139.

[22] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, "Quantum nonlocality without entanglement," *Physical Review A*, vol. 59, no. 2, pp. 1070–1091, February 1999, arXiv:quant-ph/9804053.

[23] R. Alicki and M. Fannes, "Continuity of quantum conditional information," *Journal of Physics A: Mathematical and General*, vol. 37, no. 5, pp. L55–L57, February 2004, arXiv:quant-ph/0312081.

[24] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Probl. Peredachi Inf.*, vol. 9, no. 3, pp. 3–11, 1973.

[25] 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," *Physical Review Letters*, vol. 92, p. 027902, January 2004, arXiv:quant-ph/0308012.

[26] D. Perez-Garcia, M. M. Wolf, D. Petz, and M. B. Ruskai, "Contractivity of positive and trace-preserving maps under l_p norms," *Journal of Mathematical Physics*, vol. 47, no. 8, p. 083506, August 2006, arXiv:math-ph/0601063.

[27] M. M. Wilde, *Quantum Information Theory*. Cambridge University Press, 2013, arXiv:1106.1445.

[28] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1753–1768, July 2003, arXiv:quant-ph/0206186.

[29] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2481–2485, Nov. 1999, arXiv:1409.2536.

[30] T. Ogawa and H. Nagaoka, "Making good codes for classical-quantum channel coding via quantum hypothesis testing," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 2261–2266, June 2007, arXiv:quant-ph/9808063.

[31] A. A. Fedotov, P. Harremoes, and F. Topsoe, "Refinements of Pinsker's inequality," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1491–1498, June 2003.

[32] F. Dupuis, J. Florjanczyk, P. Hayden, and D. Leung, "Locking classical information," *Proceedings of the Royal Society A*, vol. 469, no. 2159, November 2013, arXiv:1011.1612.

[33] A. Ambainis and A. Smith, "Small pseudo-random families of matrices: Derandomizing approximate quantum encryption," *Proceedings of RANDOM 2004*, vol. LNCS 3122, pp. 249–260, August 2004, Cambridge, MA. arXiv:quant-ph/0404075.

[34] I. Kerenidis and D. Nagaj, "On the optimality of quantum encryption schemes," *Journal of Mathematical Physics*, vol. 47, no. 9, p. 092102, September 2006, arXiv:quant-ph/0509169.

[35] P. A. Dickinson and A. Nayak, "Approximate randomization of quantum states with fewer bits of key," *AIP Conference Proceedings*, vol. 864, pp. 18–36, March 2006, arXiv:quant-ph/0611033.

[36] G. Aubrun, "On almost randomizing channels with a short Kraus decomposition," *Communications in Mathematical Physics*, vol. 288, no. 3, pp. 1103–1116, June 2009, arXiv:0805.2900.

[37] A. Maurer, "A bound on the deviation probability for sums of non-negative random variables," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 4, no. 1, p. 15, January 2003.

[38] R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 569–579, March 2002, arXiv:quant-ph/0012127.

[39] A. W. Harrow, "The church of the symmetric subspace," August 2013, arXiv:1308.6595.

Cosmo Lupo received the Ph.D. degree in Fundamental and Applied Physics from the University of Napoli "Federico II". He has been a postdoctoral researcher at the University of Camerino and at the Massachusetts Institute of Technology. His research interests include quantum communication theory, quantum metrology, and quantum computing.

Mark M. Wilde (M'99–SM'13) was born in Metairie, Louisiana, USA. He is an Assistant Professor in the Department of Physics and Astronomy and the Center for Computation and Technology at Louisiana State University. His current research interests are in quantum Shannon theory, quantum optical communication, quantum computational complexity theory, and quantum error correction.

Seth Lloyd is Nam P. Suh Professor of Mechanical Engineering and Professor of Physics at the Massachusetts Institute of Technology, and on the external faculty of the Santa Fe Institute. His work focuses on quantum information theory, including quantum communications, quantum algorithms, quantum metrology, and methods for building quantum computers.