Maximal Privacy without Coherence

Debbie Leung,[1] Ke Li,[2,3] Graeme Smith,[2] and John A. Smolin[2]
[1]Institute for Quantum Computing, University of Waterloo, Waterloo N2L 3G1, Ontario, Canada
[2]IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA
[3]Center for Theoretic Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
(Received 7 February 2014; published 16 July 2014)

Privacy is a fundamental feature of quantum mechanics. A coherently transmitted quantum state is inherently private. Remarkably, coherent quantum communication is not a prerequisite for privacy: there are quantum channels that are too noisy to transmit any quantum information reliably that can nevertheless send private classical information. Here, we ask how much private classical information a channel can transmit if it has little quantum capacity. We present a class of channels $\mathcal{N}_d$ with input dimension d^2, quantum capacity $Q(\mathcal{N}_d) \leq 1$, and private capacity $P(\mathcal{N}_d) = \log d$. These channels asymptotically saturate an interesting inequality $P(\mathcal{N}) \leq (1/2)[\log d_A + Q(\mathcal{N})]$ for any channel $\mathcal{N}$ with input dimension d_A and capture the essence of privacy stripped of the confounding influence of coherence.

DOI: 10.1103/PhysRevLett.113.030502 PACS numbers: 03.67.Hk, 03.67.Dd

Any communication link can be modeled as a (noisy) quantum channel. Given a sender, Alice, and a receiver, Bob, a quantum channel from Alice to Bob is a completely positive trace preserving map from an input space A to an output space B. The capability of a quantum channel for communication is measured by various capacities, one for each type of information to be transmitted. The classical capacity $C(\mathcal{N})$ quantifies the capability of a quantum channel $\mathcal{N}$ for transmitting classical information, in bits per channel use. The private capacity $P(\mathcal{N})$ gives the maximum rate of private classical communication and quantifies the optimal performance for key exchange. Finally, the quantum capacity $Q(\mathcal{N})$, measured in qubits per channel use, establishes the ultimate limit for transmitting quantum information and the performance of quantum error correction.

The three capacities mentioned above clearly satisfy $\mathcal{Q}(\mathcal{N}) \leq \mathcal{P}(\mathcal{N}) \leq \mathcal{C}(\mathcal{N})$. The analogies between coherent transmission and privacy, and between entanglement and a private key, strongly suggest that $\mathcal{Q}(\mathcal{N}) = \mathcal{P}(\mathcal{N})$. Surprisingly, it was shown in [1] that not only can P and Q differ, there are channels too noisy to transmit *any* quantum information [that is, $Q(\mathcal{N}) = 0$] but that can nevertheless be used to establish privacy [$P(\mathcal{N}) > 0$]. The central idea of [1] concerns *private* states that by their structure embody perfectly secure classical key, much as maximally entangled pure states embody perfectly coherent correlation.

While [1] draws a qualitative distinction between the private and the quantum capacities, it remains unclear how big the difference can be. If the capacities were always close, then privacy and coherence would still be closely related concepts and the distinction would have little practical relevance. Our contribution is to present a class of channels with $Q(\mathcal{N}_d) \leq 1$ and $P(\mathcal{N}_d) = \log d$, where d^2 is the input dimension and log is taken base 2 throughout the Letter. We further establish that such a separation is tight, by proving an inequality,

$$P(\mathcal{N}) \leq \frac{1}{2}[\log d_A + Q(\mathcal{N})], \tag{1}$$

for any channel $\mathcal{N}$ with input dimension d_A, quantifying the effect of incoherence in the channel transmission on privacy: inasmuch as a channel cannot simply transmit quantum information, it must devote half of its input space to acting as a "shield" system (as defined in [1]). While Eq. (1) can be inferred from properties of squashed entanglement of quantum states [2,3], this particular form appears to be new. Our relatively simple proof involves very different techniques.

As an aside, our channels combine features of private states from [1] and retrocorrectable or random-phase-coupling channels of [4–7] (these channels have large assisted capacities but small C, P, and Q). In addition to finding a very tight bound on $Q(\mathcal{N}_d)$, we can also compute both $P(\mathcal{N}_d)$ and $C(\mathcal{N}_d)$, a relative rarity in quantum information, especially for a highly nontrivial channel.

Upper bound on privacy.—Recall that any quantum channel can be expressed as an isometry followed by a partial trace, $\mathcal{N}(\rho) = \mathrm{tr}_E U\rho U^\dagger$, where $U: A \to BE$ with $U^\dagger U = I$. The complementary channel acts as $\hat{\mathcal{N}}(\rho) = \mathrm{tr}_B U\rho U^\dagger$, and allows us to define the coherent information of a channel as

$$Q^{(1)}(\mathcal{N}) = \max_{\phi_A} I_{\mathrm{coh}}(\mathcal{N}, \phi_A) \coloneqq \max_{\phi_A}[S(B) - S(E)],$$

where the maximization is taken over input quantum states ϕ_A, and $S(B)$, $S(E)$ are the von Neumann entropies of $\rho_B = \mathcal{N}(\phi_A)$ and $\rho_E = \hat{\mathcal{N}}(\phi_A)$, respectively. In turn, the

quantum capacity is proved [8–10] to be the regularized coherent information: $Q(\mathcal{N}) = \lim_{n\to\infty}(1/n)Q^{(1)}(\mathcal{N}^{\otimes n})$. We say that a quantum channel $\mathcal{N}$ is degradable if $\hat{\mathcal{N}} = \mathcal{D}\circ\mathcal{N}$ for some channel $\mathcal{D}$ [11] ($\mathcal{N}$ can be degraded to generate $\hat{\mathcal{N}}$). For degradable channels, $P(\mathcal{N}) = Q(\mathcal{N}) = Q^{(1)}(\mathcal{N})$ [12]. Degradable channels also provide a powerful tool for upper bounding the capacities of general channels [13]. If a channel $\mathcal{N} = \mathcal{L}\circ\mathcal{M}$ is a composition of two channels $\mathcal{L}$ and $\mathcal{M}$ with $\mathcal{M}$ degradable, we have

$$Q(\mathcal{N}) \le P(\mathcal{N}) = P(\mathcal{L}\circ\mathcal{M}) \le P(\mathcal{M}) = Q^{(1)}(\mathcal{M}). \quad (2)$$

We now have all the tools for proving Eq. (1). For any channel $\mathcal{N}$, define $\mathcal{M}$ as

$$\mathcal{M}(\rho) = \frac{1}{2}[\rho \otimes |0\rangle\langle 0| + \mathcal{N}(\rho) \otimes |1\rangle\langle 1|]. \quad (3)$$

Then, $\mathcal{N} = \mathcal{L}\circ\mathcal{M}$, where $\mathcal{L}(\sigma) = (\mathcal{N}\otimes\Pi_0 + \mathcal{I}\otimes\Pi_1)(\sigma)$ and $\Pi_i(\mu) = \langle i|\mu|i\rangle$. To see that $\mathcal{M}$ is degradable, note that the complementary channel of $\mathcal{M}$ is

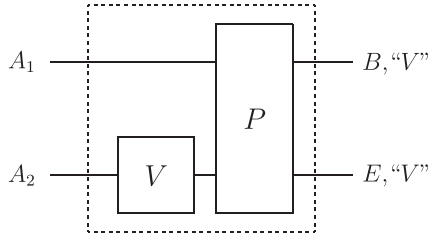
$$\hat{\mathcal{M}}(\rho) = \frac{1}{2}[|e\rangle\langle e| \otimes |0\rangle\langle 0| + \hat{\mathcal{N}}(\rho) \otimes |1\rangle\langle 1|], \quad (4)$$

where $|e\rangle\langle e|$ is an orthogonal erasure flag. Choose a degrading map $\mathcal{D}$ that first flips the flag qubit (the second register), and then conditioned on the flag being $|1\rangle$ or $|0\rangle$, applies $\hat{\mathcal{N}}$ to the first register or resets it to $|e\rangle\langle e|$. So, $\hat{\mathcal{M}} = \mathcal{D}\circ\mathcal{M}$. Now, applying Eq. (2),

$$\begin{aligned} P(\mathcal{N}) &\le Q^{(1)}(\mathcal{M}) \\ &= \max_{\phi_A}[S(B_1B_2) - S(E_1E_2)] \\ &= \max_{\phi_A}\frac{1}{2}[S(\phi_A) + S(\mathcal{N}(\phi_A)) - S(\hat{\mathcal{N}}(\phi_A))] \\ &\le \frac{1}{2}[\log d_A + Q^{(1)}(\mathcal{N})]. \end{aligned}$$

This bound is, in fact, stronger than Eq. (1), since $Q^{(1)}(\mathcal{N}) \le Q(\mathcal{N})$.

Channel construction.—The family of channels $\mathcal{N}_d$ asymptotically saturating Eq. (1) is given by



(5)

The isometric extension of the channel $\mathcal{N}_d$ is given by the operations in the dashed box. $\mathcal{N}_d$ has two input registers A_1 and A_2, each of dimension d. A random unitary V is applied to A_2, followed by a controlled phase gate $P = \sum_{i,j}\omega^{ij}|i\rangle\langle i| \otimes |j\rangle\langle j|$ acting on A_1A_2, where ω is a primitive dth root of unity. Bob receives only A_1 (now relabeled B) and "V," which denotes a classical register with a description of V. The A_2 register is discarded. The complementary channel has outputs A_2 and "V." More formally, let $W_V = P(I\otimes V)$, $\mathcal{N}_V(\rho) = \mathrm{tr}_E W_V\rho W_V^\dagger$, and $\mathcal{N}_d = \mathbb{E}_V\mathcal{N}_V \otimes |V\rangle\langle V|_{V_B}$, where the register V_B holds "V." The isometric extension is given by

$$U_d|\psi\rangle_{A_1A_2} = \sum_V \sqrt{\mathrm{pr}(V)}(W_V|\psi\rangle_{A_1A_2}) \otimes |V\rangle_{V_B} \otimes |V\rangle_{V_E},$$

and the complementary channel acts as $\hat{\mathcal{N}}_d(\rho) = \mathrm{tr}_{BV_B}U_d\rho U_d^\dagger = \mathbb{E}_V\mathrm{tr}_B W_V\rho W_V^\dagger \otimes |V\rangle\langle V|_{V_B}$.

Here is the intuition behind the construction: The classical capacity of this channel is at least $\log d$, since the d computation basis states of A_1 are transmitted without error. Furthermore, we will see that inserting a maximally mixed state into A_2 keeps the environment ignorant of the transmitted message so $P(\mathcal{N}_d) \ge \log d$. However, as the classical capacity is no greater than the output entropy, and "V" is independent of the input, $C(\mathcal{N}_d) \le \log d$, so, $C(\mathcal{N}_d) = P(\mathcal{N}_d) = \log d$. However, the channel is constructed to suppress the quantum capacity, since without knowing V, Alice cannot avoid the P gate from entangling A_1 with A_2, thereby dephasing A_1. We will prove $Q(\mathcal{N}_d) \le 1$.

Our proofs of the above statements hold for any V drawn from a so-called exact unitary 2-design, and thus, V can be a random Clifford gate [14]. In our work to *lower* bound $Q(\mathcal{N})$, a Haar distributed V is analyzed as a first attempt. We expect a similar conclusion for random Clifford gate V.

Private capacity.—For an ensemble $\mathcal{E} = \{p_i, \phi_i\}$ and channel $\mathcal{N}$, the private information is defined as

$$P^{(1)}(\mathcal{N}, \mathcal{E}) = \chi(\mathcal{N}, \mathcal{E}) - \chi(\hat{\mathcal{N}}, \mathcal{E}), \quad (6)$$

with Holevo information $\chi(\mathcal{N}, \mathcal{E}) = S(\rho) - \sum_i p_i S(\rho_i)$ evaluated on the induced ensemble $\mathcal{N}(\mathcal{E}) = \{p_i, \rho_i = \mathcal{N}(\phi_i)\}$ and average state $\rho = \sum_i p_i\rho_i$ [similarly for $\chi(\hat{\mathcal{N}}, \mathcal{E})$]. For any ensemble $\mathcal{E}$, $P^{(1)}(\mathcal{N}, \mathcal{E})$ is an achievable rate for private communication and thus a lower bound on $P(\mathcal{N})$ [10].

For our channel $\mathcal{N}_d$, choosing probabilities $p_i = 1/d$ and states $\phi_i = |i\rangle\langle i|_{A_1} \otimes (I/d)_{A_2}$ for $i = 1, \ldots, d$ gives $\chi(\mathcal{N}_d, \mathcal{E}) = \log d$ and $\chi(\hat{\mathcal{N}}_d, \mathcal{E}) = 0$, so $P(\mathcal{N}_d) \ge \log d$, as claimed.

Bipartite states carrying a key (called private states) are defined and characterized in [1]. We note that the Choi state of $\mathcal{N}_d$ with Alice holding R_1, R_2 is an exact private state of key systems R_1 and B and a single shield system R_2. We refer interested readers to Ref. [1] for further details.

Upper bound on quantum capacity.—To get an upper bound on $Q(\mathcal{N}_d)$, we consider the asymptotic behavior of the coherent information, $Q^{(1)}(\mathcal{N}_d^{\otimes n})$, for arbitrarily large n. We first define suitable notations. We group together the first input A_1 from all n channel uses, call it A_1^n, and we similarly define A_2^n, B^n, V_B^n, and V_E^n. We use $\mathbf{x}$ to denote an n-tuple of integers $(x_1, x_2, \ldots, x_n)$, where each x_i has range $\{0, 1, \ldots, d-1\}$, and similarly for $\mathbf{y}$. Finally, a random V is drawn from each channel use, and we denote the tensor product of n such independent and identically drawn unitaries by $\mathbf{V}$.

We consider the n-shot coherent information $Q^{(1)}(\mathcal{N}_d^{\otimes n}) = \max_{\phi_{A_1^n A_2^n}} [S(B^n V_B^n) - S(E^n V_E^n)]$. Since Bob and the environment receive the same classical description "V," $Q^{(1)}(\mathcal{N}_d^{\otimes n}) = \max_{\phi_{A_1^n A_2^n}} [S(B^n|V_B^n) - S(E^n|V_E^n)]$. First, we show that the optimal input state has a special form.

Lemma 1: For the channel $\mathcal{N}_d$ of Eq (5), the coherent information $I_{\mathrm{coh}}(\mathcal{N}_d^{\otimes n}, \phi_{A_1^n A_2^n})$ is maximized on states of the form

$$\phi_{A_1^n A_2^n} = \sum_{\mathbf{x}} p_{\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{x}|_{A_1^n} \otimes |\varphi^{\mathbf{x}}\rangle\langle\varphi^{\mathbf{x}}|_{A_2^n}, \tag{7}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $|\mathbf{x}\rangle = \bigotimes_{i=1}^n |x_i\rangle$ is a standard basis state on A_1^n.

Proof: First, we show that the optimal state has the form

$$\sigma_{A_1^n A_2^n} = \sum_{\mathbf{x}} p_{\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{x}|_{A_1^n} \otimes \varphi_{A_2^n}^{\mathbf{x}}, \tag{8}$$

where $\varphi_{A_2^n}^{\mathbf{x}}$ is potentially mixed. To see this, let $\psi_{A_1^n A_2^n}$ be an arbitrary input state, and $\sigma_{A_1^n A_2^n} = (\mathcal{P}^{\otimes n} \otimes I_{A_2^n})(\psi_{A_1^n A_2^n})$, where $\mathcal{P}(\rho) = (1/d)\sum_{i=0}^{d-1} Z_i \rho Z_i^\dagger$ is the completely dephasing map. So, $\sigma_{A_1^n A_2^n}$ indeed has the form given by Eq. (8). Now,

$$\begin{aligned} I_{\mathrm{coh}}(\mathcal{N}_d^{\otimes n}, \sigma_{A_1^n A_2^n}) &= I_{\mathrm{coh}}(\mathcal{N}_d^{\otimes n} \circ (\mathcal{P}^{\otimes n} \otimes I_{A_2^n}), \psi_{A_1^n A_2^n}) \\ &= I_{\mathrm{coh}}(\mathcal{P}^{\otimes n} \circ \mathcal{N}_d^{\otimes n}, \psi_{A_1^n A_2^n}) \geq I_{\mathrm{coh}}(\mathcal{N}_d^{\otimes n}, \psi_{A_1^n A_2^n}), \end{aligned}$$

since $\mathcal{P}$ commutes with $\mathcal{N}_d$, and $\mathcal{P}$ is unital so the entropy cannot decrease. Meanwhile, the reduced state on $E^n V_E^n$ remains the same, so the coherent information cannot decrease.

Next, we show that $\varphi_{A_2^n}^{\mathbf{x}}$ in Eq. (8) can be taken to be pure. Fix an arbitrary $\mathbf{x}$. Let $\varphi_{A_2^n}^{\mathbf{x}} = \sum_w q(w|\mathbf{x}) |\mu_w^{\mathbf{x}}\rangle\langle\mu_w^{\mathbf{x}}|$, and for each w, let

$$\eta_{A_1^n A_2^n}^{\mathbf{x},w} = p_{\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{x}| \otimes |\mu_w^{\mathbf{x}}\rangle\langle\mu_w^{\mathbf{x}}| + \sum_{\mathbf{y}\neq\mathbf{x}} p_{\mathbf{y}} |\mathbf{y}\rangle\langle\mathbf{y}| \otimes \varphi_{A_2^n}^{\mathbf{y}}.$$

We now show that

$$\exists w' \text{ s.t. } I_{\mathrm{coh}}(\mathcal{N}_d^{\otimes n}, \eta_{A_1^n A_2^n}^{\mathbf{x},w'}) \geq I_{\mathrm{coh}}(\mathcal{N}_d^{\otimes n}, \sigma_{A_1^n A_2^n}). \tag{9}$$

To see this, note that for each $\mathbf{x}$ and w, $\mathcal{N}_d^{\otimes n}(\sigma_{A_1^n A_2^n}) = \mathcal{N}_d^{\otimes n}(\eta_{A_1^n A_2^n}^{\mathbf{x},w})$, so those states have the same entropy. For the complementary channel, observe that by construction,

$$\sigma_{A_1^n A_2^n} = \sum_w q(w|\mathbf{x}) \eta_{A_1^n A_2^n}^{\mathbf{x},w},$$

so $\hat{\mathcal{N}}_d^{\otimes n}(\sigma_{A_1^n A_2^n}) = \sum_w q(w|\mathbf{x}) \hat{\mathcal{N}}_d^{\otimes n}(\eta_{A_1^n A_2^n}^{\mathbf{x},w})$, and by concavity of entropy,

$$S(E^n V_E^n)_{\hat{\mathcal{N}}_d^{\otimes n}(\sigma_{A_1^n A_2^n})} \geq \sum_w q(w|\mathbf{x}) S(E^n V_E^n)_{\hat{\mathcal{N}}_d^{\otimes n}(\eta_{A_1^n A_2^n}^{\mathbf{x},w})},$$

so Eq. (9) holds. Iterating this process gives an optimal state of the form given by Eq. (7). ∎

The optimal input in Eq. (7) gives a conditional output entropy of $S(B^n|V_B^n) = S(p_{\mathbf{x}}) = -\sum_{\mathbf{x}} p_{\mathbf{x}} \log p_{\mathbf{x}}$. (Taking classical distributions as diagonal density matrices, we use S to denote both the von Neumann and the Shannon entropies.) We need to compare $S(B^n|V_B^n)$ to the conditional output entropy of the environment, $S(E^n|V_E^n)$. To do so, note that the controlled-phase gate for one channel can be expressed as $P = \sum_i |i\rangle\langle i| \otimes Z_i$, where $Z_i|j\rangle = \omega^{ij}|j\rangle$. For the n-tuple $\mathbf{x}$, we define $Z^{\mathbf{x}} = Z_{x_1} \otimes \cdots \otimes Z_{x_n}$. The output state on the environment is

$$\rho_{E^n, V_E^n} = \mathbb{E}_{\mathbf{V}} \sum_{\mathbf{x}} p_{\mathbf{x}} \rho_{E^n}^{\mathbf{x},\mathbf{V}} \otimes |\mathbf{V}\rangle\langle\mathbf{V}|_{V_E^n}, \tag{10}$$

where

$$\rho_{E^n}^{\mathbf{x},\mathbf{V}} = Z^{\mathbf{x}} \mathbf{V} |\varphi^{\mathbf{x}}\rangle\langle\varphi^{\mathbf{x}}| \mathbf{V}^\dagger Z^{-\mathbf{x}}. \tag{11}$$

If, for each $\mathbf{V}$ and $\mathbf{x} \neq \mathbf{y}$, $\rho_{E^n}^{\mathbf{x},\mathbf{V}}$ and $\rho_{E^n}^{\mathbf{y},\mathbf{V}}$ were orthogonal, we would have $S(E^n|V_E^n) = S(p_{\mathbf{x}})$ and $Q^{(1)}(\mathcal{N}_d^{\otimes n}) = 0$. Instead, we prove in the Supplemental Material (Lemma I.2) [16] that the Hilbert-Schmidt inner product $\mathbb{E}_{\mathbf{V}} \mathrm{tr} \rho_{E^n}^{\mathbf{x},\mathbf{V}} \rho_{E^n}^{\mathbf{y},\mathbf{V}}$ is low on average (over the choice of $\mathbf{V}$).

Lemma 2: The states $\rho_{E^n}^{\mathbf{x},\mathbf{V}}$ given by Eq. (11) satisfy

$$\mathbb{E}_{\mathbf{V}} \mathrm{tr} \rho_{E^n}^{\mathbf{x},\mathbf{V}} \rho_{E^n}^{\mathbf{y},\mathbf{V}} \leq \frac{1}{(d-1)^{d_H(\mathbf{x},\mathbf{y})}},$$

where $d_H(\mathbf{x}, \mathbf{y}) = |\{i | x_i \neq y_i\}|$ is the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$.

Next, we derive a lower bound on the output entropy of the environment, by considering the Rényi-2 entropy of many copies of $\rho_{E^n}^{\mathbf{V}}$ using Lemma 2.

Lemma 3: For an input given by Eq. (7), conditioned on $\mathbf{V}$, the output state on the environment $\rho_{E^n}^{\mathbf{V}} = \sum_{\mathbf{x}\in[d]^n} p_{\mathbf{x}} \rho_{E^n}^{\mathbf{x},\mathbf{V}}$ satisfies

$$\mathbb{E}_{\mathbf{V}}\mathrm{tr}(\rho_{E^n}^{\mathbf{V}})^2 \le 2^n \sum_{\mathbf{x}} p_{\mathbf{x}}^2, \tag{12}$$

and

$$S(E^n|V_E^n) = \mathbb{E}_{\mathbf{V}} S(\rho_{E^n}^{\mathbf{V}}) \ge S(\mathbf{X}) - n. \tag{13}$$

Proof: To prove the first statement Eq. (12),

$$\begin{aligned}\mathbb{E}_{\mathbf{V}}\mathrm{tr}(\rho_{E^n}^{\mathbf{V}})^2 &= \mathbb{E}_{\mathbf{V}} \sum_{\mathbf{x},\mathbf{y}} p_{\mathbf{x}} p_{\mathbf{y}} \mathrm{tr}(\rho_{E^n}^{\mathbf{x},\mathbf{V}} \rho_{E^n}^{\mathbf{y},\mathbf{V}}) \\ &\le \sum_{\mathbf{x},\mathbf{y}} p_{\mathbf{x}} p_{\mathbf{y}} \frac{1}{(d-1)^{d_H(\mathbf{x},\mathbf{y})}} \\ &= \sum_{w=0}^{n} \frac{1}{(d-1)^w} \sum_{\mathbf{x}} \sum_{\mathbf{y}|d_H(\mathbf{x},\mathbf{y})=w} p_{\mathbf{x}} p_{\mathbf{y}} \\ &= \sum_{w=0}^{n} \frac{1}{(d-1)^w} \sum_{|\mathbf{e}|=w} \sum_{\mathbf{x}} p_{\mathbf{x}} p_{\mathbf{x}+\mathbf{e}},\end{aligned}$$

where the first inequality follows from Lemma 2. By the Cauchy-Schwartz inequality,

$$\sum_{\mathbf{x}} p_{\mathbf{x}} p_{\mathbf{x}+\mathbf{e}} \le \sqrt{\sum_{\mathbf{x}} p_{\mathbf{x}}^2} \sqrt{\sum_{\mathbf{x}} p_{\mathbf{x}+\mathbf{e}}^2} = \sum_{\mathbf{x}} p_{\mathbf{x}}^2.$$

We thus have

$$\begin{aligned}\mathbb{E}_{\mathbf{V}}\mathrm{tr}(\rho_{E^n}^{\mathbf{V}})^2 &\le \sum_{w=0}^{n} \frac{1}{(d-1)^w} \sum_{|\mathbf{e}|=w} \sum_{\mathbf{x}} p_{\mathbf{x}}^2 \\ &= \sum_{w=0}^{n} \frac{1}{(d-1)^w} \binom{n}{w} (d-1)^w \sum_{\mathbf{x}} p_{\mathbf{x}}^2 \\ &= 2^n \sum_{\mathbf{x}} p_{\mathbf{x}}^2.\end{aligned}$$

For the second statement Eq. (13), we first provide some intuition and a proof sketch. By the convexity of $-\log$, Eq. (12) translates to

$$\mathbb{E}_{\mathbf{V}} S_2(\rho_{E^n}^{\mathbf{V}}) \ge S_2(\mathbf{X}) - n, \tag{14}$$

where the Rényi-2 entropy S_2 is defined as $S_2(\rho) \coloneqq -\log \mathrm{tr}\rho^2$. If we can turn S_2 into S, we would have Eq. (13). We cannot do so on one copy of $\rho_{E^n}^{\mathbf{V}}$, but we can almost do so on m copies of $\rho_{E^n}^{\mathbf{V}}$ if we restrict to the "typical" approximation of $\rho_{E^n}^{\mathbf{V}}$, explained below.

Consider the case that we input m copies of the state given by Eq. (7) into mn copies of the channel $\mathcal{N}_d$. Conditioning on the random unitaries of the channels, this results in an output state $\rho_{E^{nm}}^{\mathbf{V}^m} = \otimes_{i=1}^m \rho_{E^n}^{\mathbf{V}_i}$ on the environment. The state $\rho_{E^{nm}}^{\mathbf{V}^m}$ involves a mixture over $(\mathbf{X})^m$, which is m i.i.d. copies of $\mathbf{X}$. Let $\tilde{\mathbf{X}}_m$ be the restriction of $(\mathbf{X})^m$ to its typical set $\{\mathbf{x}^m : |-(1/m)\log p_{\mathbf{x}^m} - S(\mathbf{X})| \le \epsilon\}$. The typical set contains only $\mathbf{x}^m$ with almost equal probabilities. The asymptotic equipartition theorem (cf. Ref. [15], for example) says that, for any $\epsilon > 0$, $\exists m$ such that with probability at least $1-\epsilon$, $\mathbf{x}^m$ is in the typical set. Consequently, the state $\rho_{E^{nm}}^{\mathbf{V}^m}$ can be well approximated by $\tilde{\rho}_{E^{nm}}^{\mathbf{V}^m}$, where the mixture is only taken over $\tilde{\mathbf{X}}_m$.

Note that Eq. (12) applies to $\tilde{\rho}_{E^{nm}}^{\mathbf{V}^m}$, so

$$\mathbb{E}_{\mathbf{V}^m}\mathrm{tr}(\tilde{\rho}_{E^{nm}}^{\mathbf{V}^m})^2 \le 2^{mn} \sum_{\mathbf{x}^m \in \tilde{\mathbf{X}}_m} \left(\frac{p_{\mathbf{x}^m}}{\sum_{\mathbf{x}^m \in \tilde{\mathbf{X}}_m} p_{\mathbf{x}^m}} \right)^2. \tag{15}$$

Taking $-\log$ of the above and using convexity,

$$\mathbb{E}_{\mathbf{V}^m} S_2(\tilde{\rho}_{E^{nm}}^{\mathbf{V}^m}) \ge S_2(\tilde{\mathbf{X}}_m) - mn. \tag{16}$$

Using the general property that $S \ge S_2$, and the continuity of S, the left-hand side of the above equation is approximately upper bounded by $m\mathbb{E}_{\mathbf{V}} S(\rho_{E^n}^{\mathbf{V}})$. On the right-hand side, $S_2(\tilde{\mathbf{X}}_m) \approx mS(\mathbf{X})$ because the typical set contains roughly equiprobably elements, so the two entropies are similar. Finally, dividing by m gives Eq. (13).

We provide an airtight proof in the Supplemental Material (Lemma I.3) [16]. There, we have a version of Eq. (13) that includes all the correction terms incurred by the two approximations taken in the last paragraph. Then, we show that when m is large and ϵ is small, the correction terms vanish, thereby proving Eq. (13). ∎

Together, $Q^{(1)}(\mathcal{N}_d^{\otimes n}) \le n$, so $Q(\mathcal{N}_d) \le 1$.

When proving the upper bound on Q, we cannot assume *a priori* that the entropy of $B^n V_B^n$ is maximal for the optimal input, ruling out the simpler path to show that the entropy of $E^n V_E^n$ is maximal. Instead, we have to show that $S(B^n V_B^n) - S(E^n V_E^n)$ is small for all distributions. Perhaps our technique has other applications. Also Lemma 3 effectively converts a statement concerning the Rényi-2 entropy into an analogue for the entropy for a large family of states, which may be of interest elsewhere.

Achievable quantum rate.—We have shown that $Q(\mathcal{N}_d) \le 1$, but could it actually be 0? It turns out that it cannot. In the Supplemental Material [16], we consider a specific input state $\phi_{A_1A_2} = (I/d)_{A_1} \otimes |0\rangle\langle 0|_{A_2}$ for one use of the channel, and prove an explicit lower bound $Q(\mathcal{N}_d) \ge Q^{(1)}(\mathcal{N}_d) \ge (1-\gamma)\log e \approx 0.61$ as $d \to \infty$, where $\gamma = \lim_{t\to\infty}(\sum_{k=1}^{t} \frac{1}{k} - \ln t)$ is the Euler-Mascheroni constant.

Here, we provide some intuition behind the choice of the input and the lower bound. First, note that the input (given for one use of the channel) still has the form given by Eq. (7) (with $p_{\mathbf{x}} = 1/d$ for all $\mathbf{x}$), so the subsequent analysis holds. We revisit the discussion after Eq. (11), but now with the converse in mind, and with $n = 1$ (omitted): If for some $\mathbf{V}$, $\rho_E^{\mathbf{x},\mathbf{V}}$ and $\rho_E^{\mathbf{y},\mathbf{V}}$ are *not* orthogonal, then $S(E|V_E) < S(p_{\mathbf{x}})$ and $Q^{(1)}(\mathcal{N}_d) > 0$. For the specific input state, $\phi_{A_1A_2} = (I/d)_{A_1} \otimes |0\rangle\langle 0|_{A_2}$, $|\varphi^{\mathbf{x}}\rangle = |0\rangle$ for

all $\mathbf{x}$. So, for any $\mathbf{x} \neq \mathbf{y}$, $\rho_E^{\mathbf{x},\mathbf{V}}$ and $\rho_E^{\mathbf{y},\mathbf{V}}$ are orthogonal only for a vanishing fraction of all possible $\mathbf{V}$, contributing to a significantly large $Q^{(1)}(\mathcal{N}_d)$.

Discussion.—In [1] it was shown that privacy and distillable entanglement can be different, indeed privacy can be nonzero even for bound-entangled states. What we have shown is similar, but somewhat incomparable. Our result is stronger in that the separation is maximal, saturating Eq. (1), but it only applies to the channel case, implicitly not allowing classical communication. The two-way assisted quantum capacity $Q_2(\mathcal{N}_d)$ is maximal (not zero!) and equal to the private capacity $\log d$. An open question is how big the separation can be in the two-way setting?

D. L. was supported by NSERC, DAS, CRC, and CIFAR, K. L. by NSF Grants No. CCF-1110961 and No. CCF-1111382, and G. S. and J. A. S. by DARPA QUEST.

[1] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 160502 (2005).
[2] M. Christandl, arXiv:quant-ph/0604183.
[3] M. Christandl, N. Schuch, and A. Winter, Phys. Rev. Lett. **104**, 240405 (2010).
[4] C. H. Bennett, I. Devetak, P. W. Shor, and J. A. Smolin, Phys. Rev. Lett. **96**, 150502 (2006).
[5] G. Smith and J. A. Smolin, Phys. Rev. Lett. **102**, 010501 (2009).
[6] K. Li, A. Winter, X. B. Zou, and G. Guo, Phys. Rev. Lett. **103**, 120501 (2009).
[7] G. Smith and J. A. Smolin, Phys. Rev. Lett. **103**, 120503 (2009).
[8] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
[9] P. W. Shor, in Proceedings of the MSRI Workshop on Quantum Computation, 2002, http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/.
[10] I. Devetak, IEEE Trans. Inf. Theory **51**, 44 (2005).
[11] I. Devetak and P. W. Shor, Commun. Math. Phys. **256**, 287 (2005).
[12] G. Smith, Phys. Rev. A **78**, 022306 (2008).
[13] G. Smith and J. Smolin, in *Proceedings of the IEEE Information Theory Workshop, 2008 (ITW'08)* (IEEE, New York, 2008), pp. 368–372.
[14] D. P. Divincenzo, D. W. Leung, and B. M. Terhal, IEEE Trans. Inf. Theory **48**, 580 (2002).
[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[16] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.113.030502, which includes Ref. [17], for proofs of Eq. (13) and the achievable quantum rate.
[17] K. M. R. Audenaert, J. Phys. A **40**, 8127 (2007).