Variational estimation of capacity bounds for quantum channels

Junaid ur Rehman [1], Seongjin Hong,[2] Yong-Su Kim,[2,3] and Hyundong Shin [1,*]

[1]*Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin 17104, Republic of Korea*
[2]*Center for Quantum Information, Korea Institute of Science and Technology (KIST), Seoul 02792, Republic of Korea*
[3]*Division of Nano & Information Technology, KIST School, Korea University of Science and Technology, Seoul 02792, Republic of Korea*

We propose a framework to variationally obtain detectable capacity bounds for quantum channels. The proposal of this framework is motivated by the difficulty of estimating the von Neumann entropy of an unknown quantum state. Instead of estimating the von Neumann entropy at the channel output, we propose to estimate the state purity—which can be measured by a single measurement setting—followed by bounding the von Neumann entropy from above and below. This procedure leads to new upper and lower bounds on various communication rates of quantum channels for some fixed input states. Then, by utilizing the variational method to find optimal input states we obtain lower bounds on (i) the quantum capacity of arbitrary channels, (ii) the entanglement-assisted classical capacity of arbitrary channels, and (iii) the classical capacity of covariant channels. Corresponding to these lower bounds, we also obtain upper bounds on (i) N-shot coherent information, (ii) the entanglement-assisted classical capacity, and (iii) the N-shot Holevo capacity of arbitrary quantum channels. All these bounds can be estimated by a single measurement setting without needing a full process tomography or any further *a priori* knowledge, e.g., preferred basis of the channel.

I. INTRODUCTION

Over the last two decades we have witnessed a rapid growth of quantum technologies both in theory and in practice [1]. This growth in the field of quantum metrology, quantum computation, and quantum communication necessitates that one accurately certify and benchmark the available quantum resources [2]. However, as our system of interest grows in size, it becomes exponentially difficult to characterize these resources on a classical computer. For example, the complexity of quantum state (channel/process) tomography, a task to characterize a quantum state (channel), scales as d^2 (d^4), where d is the dimension of the state (channel) under question. Obtaining a classical description of these resources is required when one wants to characterize some inherent property, e.g., the entanglement content of a quantum state or the capacity of a quantum channel.

This problem has been previously addressed in the literature and some approaches to directly characterize the quantities of interest without obtaining a classical description of the system of interest have been proposed [3–5]. These approaches began with the direct characterization of properties of quantum states [3,4]. Later on, this approach was extended for direct estimation of properties of quantum channels, e.g., capacities or certain structural properties of quantum channels. To the best of our knowledge, the first work to propose experimentally accessible bounds on the capacity of an unknown quantum channel was Ref. [6], which lower bounded the quantum capacity. This method was later experimentally

implemented for amplitude damping, dephasing, depolarizing, and Pauli channels [7]. In Ref. [8], a method for capacity estimation and verification of qubit channels with arbitrary correlated errors was proposed. This proposal also included the experimental demonstration on a transmon qubit, which exhibits a noisy quantum channel [9]. An experimentally accessible tight lower bound on the classical (Holevo) capacity of discrete Weyl/generalized Pauli channels of arbitrary dimensions was proposed in Ref. [10]. Recently, a lower bound on the classical capacity of arbitrary quantum channels was proposed in Ref. [11] with the experimental demonstration reported in Ref. [12].

Since these quantities of interest are entropic in nature, one major challenge in the aforementioned works is to accurately and reliably estimate the entropy of the quantum states. Measuring quantum entropy is known to be a challenging problem [13,14]. In fact, this problem is known to be as hard as estimating the actual quantum state [15]. In particular, the aforementioned bounds of quantum channel capacities [6,10,11] provide tight estimates of the actual capacities if the basis of the channel input and the measurement basis are exactly aligned with some specific basis. For example, the classical capacity of a qubit dephasing channel is 1 bit/channel use. The aforementioned accessible bounds [10,11] give 1 bit/channel use if at least one of the three measurement bases (Pauli X, Y, or Z) is exactly matched with the basis of dephasing. In a practical scenario with no *a priori* knowledge on the channel structure, it is difficult to justify the alignment of input and the measurement basis with the dephasing basis of the channel. Thus, the estimated bound may turn out to be considerably smaller than the actual value. Similar challenges arise in

---

*Corresponding author: hshin@khu.ac.kr

TABLE I. A comparison of this work with existing literature on detectable bounds. Quantum capacity, entanglement-assisted classical capacity, and the classical capacity are denoted by Q, C_{EA}, and C, respectively. Coherent information and Holevo capacity are denoted by I_c and χ, respectively. Superscripts ub and lb indicate an upper and a lower bound on the corresponding capacities, respectively.

Work	Result	$Q(\cdot)$		$C_{EA}(\cdot)$		$C(\cdot)$
	$Q^{lb}(\cdot)$	$I_c^{ub}(\cdot)$	$C_{EA}^{lb}(\cdot)$	$C_{EA}^{ub}(\cdot)$	$C^{lb}(\cdot)$	$\chi^{ub}(\cdot)$
[6]	✓	✗	✗	✗	✗	✗
[8]	✓	✗	✗	✗	✗	✗
[10]	✗	✗	✗	✗	✓	✗
[11]	✗	✗	✗	✗	✓	✗
This work	✓	✓	✓	✓	✓	✓

the estimation of quantum capacities of unknown quantum channels [6,8].

In this work, we circumvent this difficulty in the entropy estimation by estimating the purity of the output state of the channel and then utilizing the bounds on the entropy of a quantum state whose purity is known. We supplement this strategy with a variational approach, i.e., we parametrize the input state to the channel whose parameters are varied with the help of a classical optimizer. This variational approach has recently seen great success in several modern quantum computing and communications tasks.

In terms of the required number of measurement setting, purity estimation is an easier problem than the entropy estimation in both gate-based quantum computers [16–18] and photonic quantum systems [19–21]. In fact, purity can be estimated by using only a single fixed measurement setting on two copies of the arbitrary-dimensional quantum state. Furthermore, we demonstrate in the numerical results that our proposed approach not only provides simple to estimate, basis-independent bounds on the capacities of quantum channels but also these bounds coincide with the known analytical capacities of several important quantum channels. Additionally, this approach is general enough to provide detectable bounds on multiple capacities including quantum capacity, entanglement-assisted capacity, and classical capacity of quantum channels. A comparison of our work with existing works is summarized in Table I.

II. PURITY-BASED CAPACITY BOUNDS FOR QUANTUM CHANNELS

By nature, communication capacities of quantum channels are information-theoretic quantities. Thus, estimating the channel capacity requires estimating and *optimizing* some *entropic* quantity at the channel output. In this work, our proposed strategy for capacity estimation has two main components. First, we recast the problem of entropy estimation with purity estimation at the channel output. From the estimated output purity we are able to bound the output entropy, from above and below, by utilizing the known bounds on the von Neumann entropy of a quantum state with known purity [22–24]. By appropriately replacing these bounds in

the entropic expression for the communication rate, we obtain upper and lower bounds on the rate of interest. Similar approaches have been previously employed for entanglement spectroscopy (see, e.g., Refs. [18,25]). Second, we utilize the method of variation to optimize this expression. To this end, we parametrize the input state of the channel and then iteratively vary these parameters with the help of a classical computer to optimize the entropic expression at the channel output. At the end of this iterative process, we can lower bound the quantum capacity and the entanglement-assisted classical capacity of arbitrary channels and the classical capacity of covariant channels. Additionally, we can upper bound the coherent information, the entanglement-assisted classical capacity, and the Holevo capacity of arbitrary channels.

A. Purity-based entropy bounds for quantum states

The state ρ of a quantum system is represented by a positive semidefinite matrix of unit trace, i.e., $\rho \geqslant 0$ and $\mathrm{tr}\,\rho = 1$. A quantum channel $\mathcal{N}$ is a trace-preserving completely positive map corresponding to physical operations on the set of quantum states. The capacity of a quantum channel is the maximum achievable rate of reliable communication over the channel. This channel capacity can be defined depending on the available resources and the nature of information that we want to communicate. For example, the quantum capacity $Q(\mathcal{N})$ is defined for quantum information; the classical capacity $C(\mathcal{N})$ is defined for classical information; and the entanglement-assisted classical capacity $C_{EA}(\mathcal{N})$ is defined for classical information when communicating parties have access to an arbitrary amount of shared entanglement.

The von Neumann entropy $S(\rho)$ of the quantum state ρ can be represented as the Shannon entropy of its eigenvalue distribution, i.e.,

$$S(\rho) = -\mathrm{tr}\,\rho \log_2 \rho = -\sum_i \lambda_i \log_2 \lambda_i,$$

where λ_i's are the eigenvalues of ρ. Similarly, the purity $T(\rho)$ of the quantum state ρ can be defined as the sum of its squared eigenvalues, i.e.,

$$T(\rho) = \mathrm{tr}\,\rho^2 = \sum_i \lambda_i^2,$$

which is called the correlation figure for correlation matrices [26] or the index of coincidence for classical probability distributions [23]. For $T(\rho) = 1$ (pure states) and $T(\rho) = 1/d$ (maximally mixed states), we have the relation

$$S(\rho) = -\log_2 T(\rho). \tag{1}$$

In general, there is no one-to-one relation between the entropy and the purity of quantum states, except for these two extreme cases. However, we can bound the von Neumann entropy $S(\rho)$ when the purity $T(\rho)$ is known [22–24].

Let ρ be an arbitrary d-dimensional quantum state with $1/d \leqslant T(\rho) \leqslant 1$ and $H_{\pm}(\rho)$ be the quasiuniform Shannon entropy function of ρ, defined as

$$H_{\pm}(\rho) = -(\kappa_{\pm} - 1)p_{\pm} \log_2 p_{\pm} - q_{\pm} \log_2 q_{\pm}, \tag{2}$$

where $\kappa_+ = \lceil 1/T(\rho) \rceil$, $\kappa_- = d$, and

$$p_\pm = \frac{1 \pm \sqrt{1 - \frac{\kappa_\pm[1-T(\rho)]}{\kappa_\pm - 1}}}{\kappa_\pm},$$

$$q_\pm = 1 - (\kappa_\pm - 1)p_\pm. \tag{3}$$

Then,

$$L(\rho) \leqslant S(\rho) \leqslant U(\rho), \tag{4}$$

where $L(\rho) = H_+(\rho)$ and $U(\rho) = H_-(\rho)$. Note that for $T(\rho) = 1$ (pure states), lower and upper bounds coincide with $S(\rho) = 0$, while for $T(\rho) = 1/d$ (maximally mixed states), these two bounds coincide with $S(\rho) = \log_2 d$.

The main motivation for using the purity instead of the entropy at the channel output is the ease of its experimental estimation as compared to the entropy estimation. The purity $T(\rho)$ can be operationally seen as the expected value of the SWAP operator $W = \sum_{jk} |jk\rangle \langle kj|$ on the two copies of ρ, i.e.,

$$T(\rho) = \mathrm{tr}(\rho \otimes \rho)W. \tag{5}$$

The spectrum of W consists of symmetric and antisymmetric eigenspaces with eigenvalues 1 and -1, respectively. Hence, the expectation in Eq. (5) can be obtained by implementing a two-outcome projective measurement on the symmetric and antisymmetric subspaces of the joint Hilbert space. One example of such measurement for a qubit case is the Bell-state measurement where the singlet lies in the antisymmetric eigenspace of W and the remaining Bell states correspond to the symmetric eigenspace. We remark that obtaining the expectation of the SWAP operator requires a single but nonlocal measurement setting. On the other hand, if we employ some other technique of direct entropy estimation, for example, quantum state tomography, it requires only local but exponentially many (order of 2^{2n}) measurement settings for an n-qubit state. In particular, the purity $T(\rho)$ can be efficiently estimated by using the well-known swap test on a gate-based quantum computer [16–18] or by measuring the visibility of the Hong-Ou-Mandel dip in the photonic systems [19–21].

The second main component of this work is to use a method of variations for the capacity bounds. We input a physically available but unknown quantum channel with a parametrized quantum state with some randomly chosen parameter values. At the channel output, we measure the purity of the output state, calculate the bounds for the communication rate, and input the obtained parameters to the classical optimizer, which then generates a new set of parameters for the input state in the next round. The classical optimizer attempts to maximize the bounds on the communication rate of the channel output by varying the parameters of the state. Performing this optimization iteratively tightens the variational capacity bounds for the given unknown quantum channel. A pictorial representation and description of the proposed method is provided in Fig. 1 and its caption, respectively.

B. Bounds on the quantum capacity

Now we are ready to transform the capacity expressions of quantum channels to a variational bound framework. The quantum capacity of a quantum channel $\mathcal{N}$ is defined as
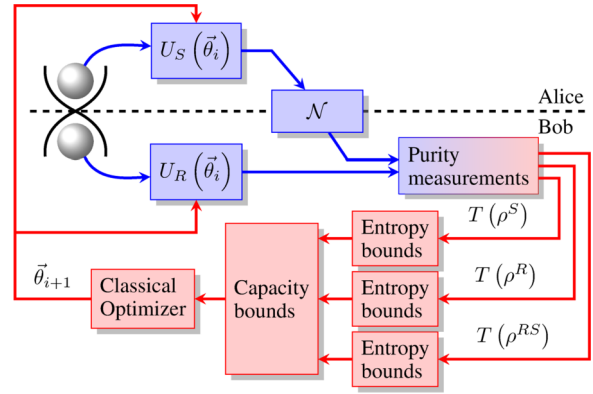

FIG. 1. Variational estimation of capacity bounds. The exact protocol for variational estimation of capacity bounds for $Q(\cdot)$ and $C_{\mathrm{EA}}(\cdot)$ is as follows. (i) Alice and Bob begin with the maximally entangled state on $\mathcal{H}_R \otimes \mathcal{H}_S$, where the system qudit (S) is held by Alice and the reference qudit (R) is held by Bob. (ii) They apply local operations and classical communication (LOCC) according to parameter values $\vec{\theta}_i$ to obtain $|\psi(\vec{\theta}_i)\rangle^{RS}$. (iii) Alice sends her qudit to Bob through $\mathcal{N}$, whereas Bob's qudit undergoes no evolution $\mathcal{I}$. (iv) Upon receiving Alice's qudit, Bob performs purity measurements and bounds the marginal entropies of S and R, as well as the joint entropy of RS by utilizing Eq. (4). (v) Then, he computes the bound of interest by computing Eq. (10), (14), or (19). (vi) Bob feeds the obtained value to the optimizer to obtain a new set of parameter $\vec{\theta}_{i+1}$, which he communicates to Alice. (vii) They calculate their parts of the LOCC and perform the next iteration. Once some convergence criterion is met, they output the estimated bounds on the capacity of interest of $\mathcal{N}$.

[27–31]

$$Q(\mathcal{N}) = \lim_{N \to \infty} \frac{Q_N(\mathcal{N})}{N}, \tag{6}$$

where

$$Q_N(\mathcal{N}) = \max_\rho I_c(\rho, \mathcal{N}^{\otimes N}) \tag{7}$$

is the maximum coherent information over N channel uses of $\mathcal{N}$. The coherent information of a channel $\mathcal{N}$ is [32]

$$I_c(\rho, \mathcal{N}) = S[\mathcal{N}(\rho)] - S_e(\rho, \mathcal{N}), \tag{8}$$

where $S_e(\rho, \mathcal{N}) = S[\mathcal{I} \otimes \mathcal{N}(|\psi_\rho\rangle \langle \psi_\rho|)]$ is the entropy exchange, where $|\psi_\rho\rangle^{RS}$ is any purification of the system qubit ρ^S with a reference system R, and $\mathcal{I}$ is the identity channel that acts on the reference part of the purification.

From the purification of ρ, we can rewrite the last expression as

$$I_c(\rho, \mathcal{N}) = S(\mathrm{tr}_R[\mathcal{I} \otimes \mathcal{N}(|\psi_\rho\rangle \langle \psi_\rho|)])$$
$$- S[\mathcal{I} \otimes \mathcal{N}(|\psi_\rho\rangle \langle \psi_\rho|)], \tag{9}$$

where $\mathrm{tr}_R[\cdot]$ denotes the partial trace with respect to the reference system R. Then, the maximum of coherent information can be found by optimizing the pure states in d^2-dimensional Hilbert space, which is a more suitable form for the variational circuits, instead of optimizing mixed states on d-dimensional Hilbert space.

Finally, observing that both terms on the right of Eq. (9) are entropies of quantum states, we can readily bound them by

inserting the purity-based bounds discussed above. Then, we can obtain variational lower and upper bounds on the coherent information of arbitrary unknown quantum channels.

Specifically, let $\rho^{RS} = \mathcal{I} \otimes \mathcal{N}(|\psi\rangle)$ and $\rho^S = \mathrm{tr}_R(\rho^{RS})$. Then,

$$I_c^{\mathrm{lb}}(\rho, \mathcal{N}) := L(\rho^S) - U(\rho^{RS}) \leqslant I_c(\rho, \mathcal{N})$$
$$\leqslant U(\rho^S) - L(\rho^{RS}) := I_c^{\mathrm{ub}}(\rho, \mathcal{N}). \quad (10)$$

Maximizing all expressions in the last chain of inequalities gives

$$Q_1^{\mathrm{lb}}(\mathcal{N}) := \max_{|\psi\rangle^{RS}} I_c^{\mathrm{lb}}(\rho, \mathcal{N}) \leqslant Q_1(\mathcal{N})$$
$$\leqslant Q_1^{\mathrm{ub}}(\mathcal{N}) := \max_{|\psi\rangle^{RS}} I_c^{\mathrm{ub}}(\rho, \mathcal{N}), \quad (11)$$

where the d^2-dimensional $|\psi\rangle$ is input to $\mathcal{I} \otimes \mathcal{N}$. Note that $Q_1^{\mathrm{lb}}(\mathcal{N})$ and $Q_1^{\mathrm{ub}}(\mathcal{N})$ are the lower and upper VCBs to the one-shot quantum capacity of $\mathcal{N}$ that can be estimated without invoking the full process tomography of $\mathcal{N}$. Since we have the chain of inequalities $Q_1^{\mathrm{lb}}(\mathcal{N}) \leqslant Q_1(\mathcal{N}) \leqslant Q(\mathcal{N}), Q_1^{\mathrm{lb}}(\mathcal{N})$ also bounds the quantum capacity of $\mathcal{N}$ from below. Furthermore, since $\mathcal{N}$ is completely arbitrary, we can also estimate the N-shot coherent information (7) by substituting $\mathcal{N}$ with $\mathcal{N}^{\otimes N}$. Therefore, this method is also suitable for bounding the N-shot coherent information of $\mathcal{N}$ for any finite N. Furthermore, since the private information $P(\mathcal{N})$ of a channel $\mathcal{N}$ follows $P(\mathcal{N}) \geqslant Q_1(\mathcal{N})$, the lower bound we obtained above also lower bounds $P(\mathcal{N})$.

C. Bounds on the entanglement-assisted classical capacity

The entanglement-assisted classical capacity (EACC) of a quantum channel is the maximum rate of classical communication when the transmitter and the receiver have preshared entanglement [33,34]. The EACC is known to be additive and thus can be characterized by a single-letter formula, channel's mutual information, without the need of regularization [28,33,35]:

$$C_{\mathrm{EA}}(\mathcal{N}) = \max_{|\psi\rangle^{AA'}} I(A;B)_\rho, \quad (12)$$

where

$$I(A;B)_\rho = S(\rho^A) + S(\rho^B) - S(\rho^{AB}), \quad (13)$$

where $\rho^{AB} = \mathcal{I} \otimes \mathcal{N}(|\psi\rangle^{AA'})$, and ρ^A and ρ^B are the parts of ρ^{AB} by tracing out the B and A, respectively. Similar to the previous case, we readily bound the last expression as

$$I^{\mathrm{lb}}(A;B)_\rho := L(\rho^A) + L(\rho^B) - U(\rho^{AB}) \leqslant I(A;B)_\rho$$
$$\leqslant U(\rho^A) + U(\rho^B) - (\rho^{AB})$$
$$:= I^{\mathrm{ub}}(A;B)_\rho. \quad (14)$$

Then, the maximization of all expressions in the chain of inequalities gives

$$C_{\mathrm{EA}}^{\mathrm{lb}}(\mathcal{N}) := \max_{|\psi\rangle^{AA'}} I^{\mathrm{lb}}(A;B)_\rho \leqslant C_{\mathrm{EA}}(\mathcal{N})$$
$$\leqslant C_{\mathrm{EA}}^{\mathrm{ub}}(\mathcal{N}) := \max_{|\psi\rangle^{AA'}} I^{\mathrm{ub}}(A;B)_\rho. \quad (15)$$

In a practical setting, the estimation method of EACC between Alice and Bob is similar to that of the estimation method for the quantum capacity as shown in Fig. 1. The only difference between the two estimation techniques is the entropy estimates and the evaluated expression at Bob's end. For the quantum capacity, Bob estimates the constituents and then attempts to maximize Eq. (10) whereas for the EACC he operates with Eq. (14).

D. Bounds on the classical capacity

The classical capacity of a quantum channel $\mathcal{N}$ is the maximum rate of reliable classical communication over arbitrarily many uses of $\mathcal{N}$. It is given by

$$C(\mathcal{N}) = \lim_{N \to \infty} \frac{\chi_N(\mathcal{N})}{N}, \quad (16)$$

where

$$\chi_N(\mathcal{N}) = \max_{p_i, \rho_i} \left[\sum_i S[p_i \mathcal{N}_N(\rho_i)] - \sum_i p_i S[\mathcal{N}_N(\rho_i)] \right] \quad (17)$$

is the Holevo capacity of $\mathcal{N}$ over N channel uses. In Eq. (17), ρ_i denotes the signal states sent over the channel with the corresponding probability p_i. Similar to the quantum capacity, the classical capacity of an arbitrary channel is feared to be uncomputable. However, $\chi_1(\mathcal{N})$ is known to be a reliable lower bound on the classical capacity. It is well known that the Holevo capacity is bounded from above as

$$\chi_N(\mathcal{N}) \leqslant N \log_2 d - \min_\rho S[\mathcal{N}_N(\rho)], \quad (18)$$

with equality for unital qubit and covariant channels of arbitrary dimensions [36]. Thus, following the same reasoning as before, we have the bounds

$$\chi_N^{\mathrm{lb}}(\mathcal{N}_N) := N \log_2 d - \min_\rho U[\mathcal{N}_N(\rho)] \overset{(a)}{\leqslant} \chi_N(\mathcal{N})$$
$$\leqslant \chi_N^{\mathrm{ub}}(\mathcal{N}_C) := N \log_2 d - \min_\rho L[\mathcal{N}_N(\rho)], \quad (19)$$

where the lower bound (a) is valid only for unital qubit and qudit covariant channels [36].

III. NUMERICAL EXAMPLES

We exemplify the performance of our proposed bounds with variational estimation for the amplitude damping channel,

$$\mathcal{N}(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger, \quad (20)$$

the depolarizing channel,

$$\mathcal{N}(\rho) = (1 - p)\rho + p\frac{I}{2}, \quad (21)$$

and the Pauli channel,

$$\mathcal{N}(\rho) = \sum_{i=0}^3 p_i \sigma_i \rho \sigma_i^\dagger, \quad (22)$$

where

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, A_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}, \quad (23)$$

and σ_0, σ_1, σ_2, and σ_3 are identity, Pauli X, Pauli Y, and Pauli Z matrices, respectively. We set the probabilities $\{p_i\}_i$ of Pauli channels from the eigenvalues of 4×4 exponential correlation and constant tridiagonal correlation matrices, which are parametrized by a single parameter [26]. The exponential correlation matrix that we use here is

$$\Phi^{(\mathrm{exp})}(\eta) = \frac{1}{4} \begin{bmatrix} 1 & \eta & \eta^2 & \eta^3 \\ \eta & 1 & \eta & \eta^2 \\ \eta^2 & \eta & 1 & \eta \\ \eta^3 & \eta^2 & \eta & 1 \end{bmatrix}, \qquad (24)$$

where $\eta \in [0, 1]$. The constant tridiagonal correlation matrix is

$$\Phi^{(\mathrm{tri})}(\xi) = \frac{1}{4} \begin{bmatrix} 1 & \xi & 0 & 0 \\ \xi & 1 & \xi & 0 \\ 0 & \xi & 1 & \xi \\ 0 & 0 & \xi & 1 \end{bmatrix}, \qquad (25)$$

where $\xi \in [0, 0.5/\cos\frac{\pi}{5}]$. The main motivation for using eigenvalues of these correlation matrices for channel parameters is because of the majorization relation of obtained eigenvalue vectors by varying the parameters η and ξ. By varying these parameters, we can move smoothly from the maximally noisy channel to the noiseless channel.

Figures 2(a)–2(c) show the proposed bounds along with the actual values of the coherent information for amplitude damping, depolarizing, and Pauli (probabilities from exponential correlation matrix) channels, respectively. These three examples show three distinct behaviors of the proposed bounds. The upper and the lower bounds are saturated by the amplitude damping channel and the depolarizing channel, respectively. On the other hand, the Pauli channel whose parameters are taken from the eigenvalues of the correlation matrix saturates neither of the proposed bounds.

Figures 3(a)–3(d) show the proposed bounds and the corresponding EACC values for the amplitude damping, depolarizing, Pauli (exponential correlation), and Pauli (constant tridiagonal correlation) channels, as functions of the channel parameters. Here, we behavior similar to that of coherent information bounds. An interesting case emerges for the Pauli channel with constant tridiagonal correlation matrix. In this case, the EACC is saturated by the corresponding lower bound for the small values of η and is almost saturated by the upper bound for the higher values of η, whereas the difference between the two bounds remains constant.

A. Experiments on IBM QISKIT

We implement the proposed capacity estimation protocol on the IBM's Qiskit library and IBM quantum (IBMQ) devices [37]. To this end, we simulate the channel of interest in Qiskit [38], measure the purity at the channel output [17], and iteratively update the channel input. The implemented circuits for these experiments are shown in Figs. 4 and 5 for the amplitude damping and the depolarizing channel, respectively.

In the circuit for the amplitude damping channel, we have $\theta = \arccos(\sqrt{1-\gamma})$, where γ is the amplitude damping parameter from the main text. On the other hand, for the depolarizing channel we have $\theta = \arccos(1-2p)/2$, where p is the amplitude damping parameter. The purity of the joint
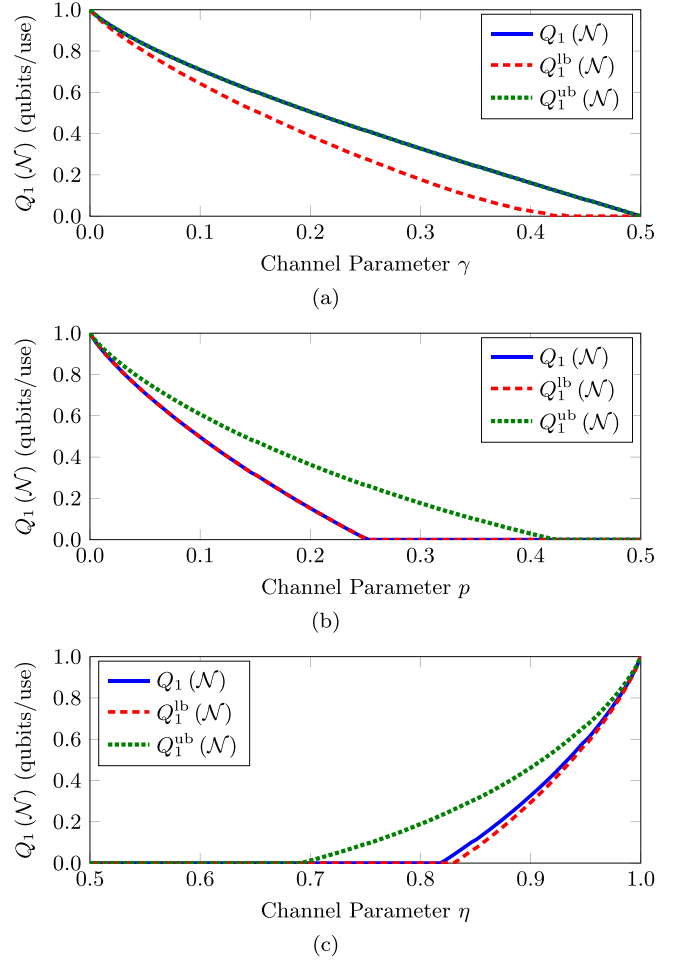

FIG. 2. Coherent information $Q_1(\mathcal{N})$ and the corresponding bounds for (a) amplitude damping, (b) depolarizing, and (c) Pauli channels whose parameters are determined from the 4×4 exponential correlation matrix.

state is computed by first obtaining $\vec{c}_{RS}$, the vector of normalized counts at the output, and taking the dot product with $\vec{b}_{RS} = [1, 1, 1, -1]^{\otimes 2}$ [17]. Similarly, the purity of system (resp. reference) qubit alone can be obtained by calculating $\vec{c}_S$ (respectively $\vec{c}_R$) from $\vec{c}_{RS}$ and taking the dot product with $\vec{b}_S$ (respectively $\vec{b}_R$) $= [1, 1, 1, -1]$.

We performed two types of experiments by implementing these circuits in Qiskit. First, we performed noiseless simulations on `QASM` simulator. Second, we performed experiments on `FakeMelbourne()` simulator provided by `qiskit.test.mock` module. This simulator mimics the behavior of `ibmq_16_melbourne`, which is a 16-qubit quantum device offered by IBMQ experience.

We parametrized the state as [39]

$$\begin{aligned} |\psi\rangle_{RS} = {} & \cos(\theta_1/2)\cos(\theta_2/2)\,|00\rangle \\ & + \cos(\theta_1/2)\sin(\theta_2/2)e^{\iota\phi_2}\,|01\rangle \\ & + \sin(\theta_1/2)\cos(\theta_3/2)e^{\iota\phi_1}\,|10\rangle \\ & + \sin(\theta_1/2)\sin(\theta_3/2)e^{\iota(\phi_1+\phi_3)}\,|11\rangle \end{aligned}$$
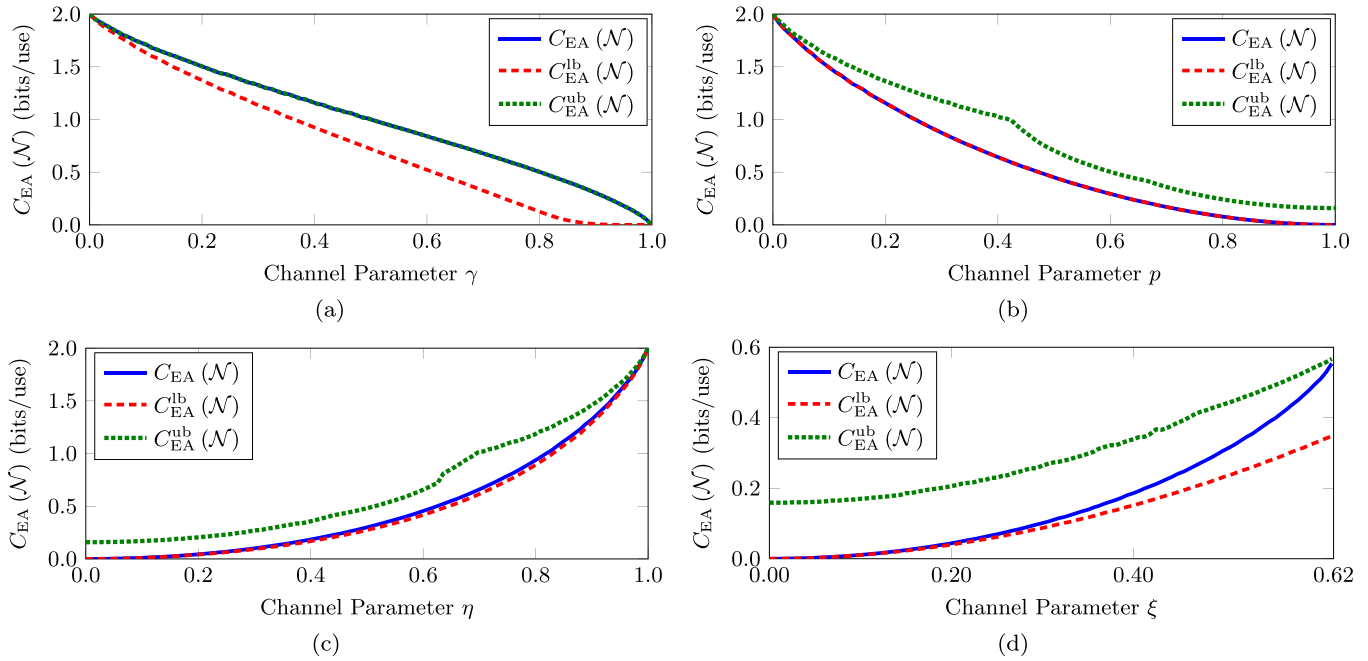
FIG. 3. Entanglement-assisted classical capacity C_{EA} and the corresponding bounds for (a) amplitude damping, (b) depolarizing, (c) Pauli (exponential correlation), and (d) Pauli (constant tridiagonal correlation) channels. Solid lines show the entanglement-assisted classical capacity for the corresponding channels.

and utilized the "Powell" optimizer for varying and optimizing these parameters [40]. In both types of experiments, we set the number of measurements `n_shots = 8192` per iterations and set `maxfev = 100` in `scipy.optimize` module per channel. The parameter `maxfev` specifies the maximum number of allowed evaluations of the objective function. In our case, this corresponds to $\sim$140–150 iterations of state

preparation, evolution, and then measurement. In these examples, we initialized our ansatz (26) to the parameters $[\theta_1, \theta_2, \theta_3, \phi_1, \phi_2, \phi_3] = [\pi/2, 0, \pi, 2, 1, 0.2]$, which served as a somewhat arbitrary but good starting point for these optimization runs. We left all other optimizer options to their default values. The optimization terminated when either the maximum number of iterations was achieved or the
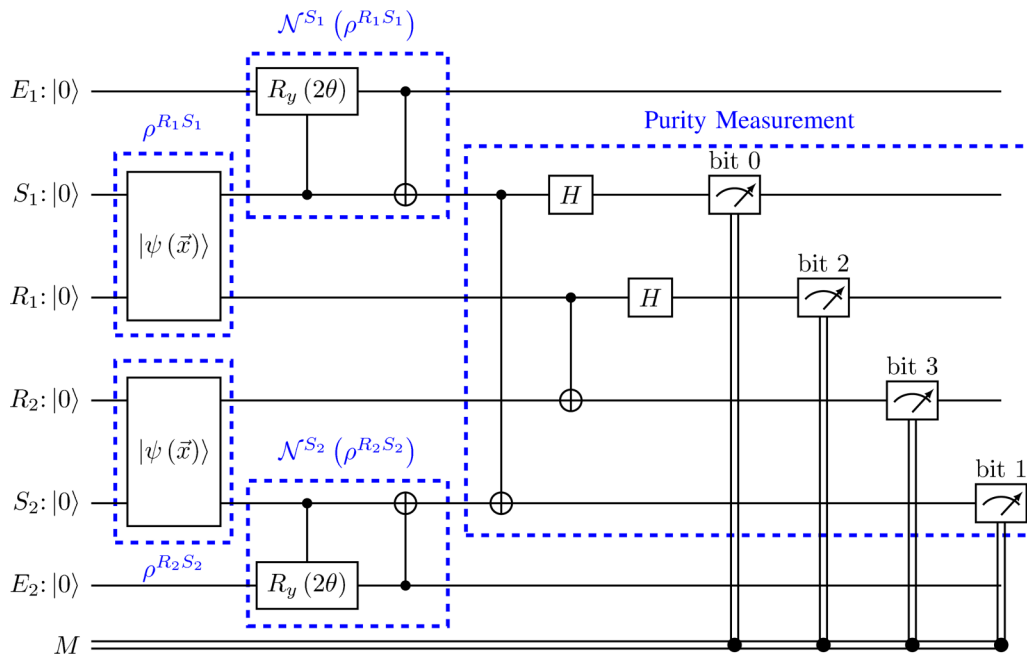

FIG. 4. Quantum circuit for amplitude damping channel. Purity measurement requires two copies of the state $I^R \otimes \mathcal{N}^S(\rho^{RS})$. We prepare these two copies on wires 1–4, labeled as S_1, R_1, R_2, and S_2, respectively.
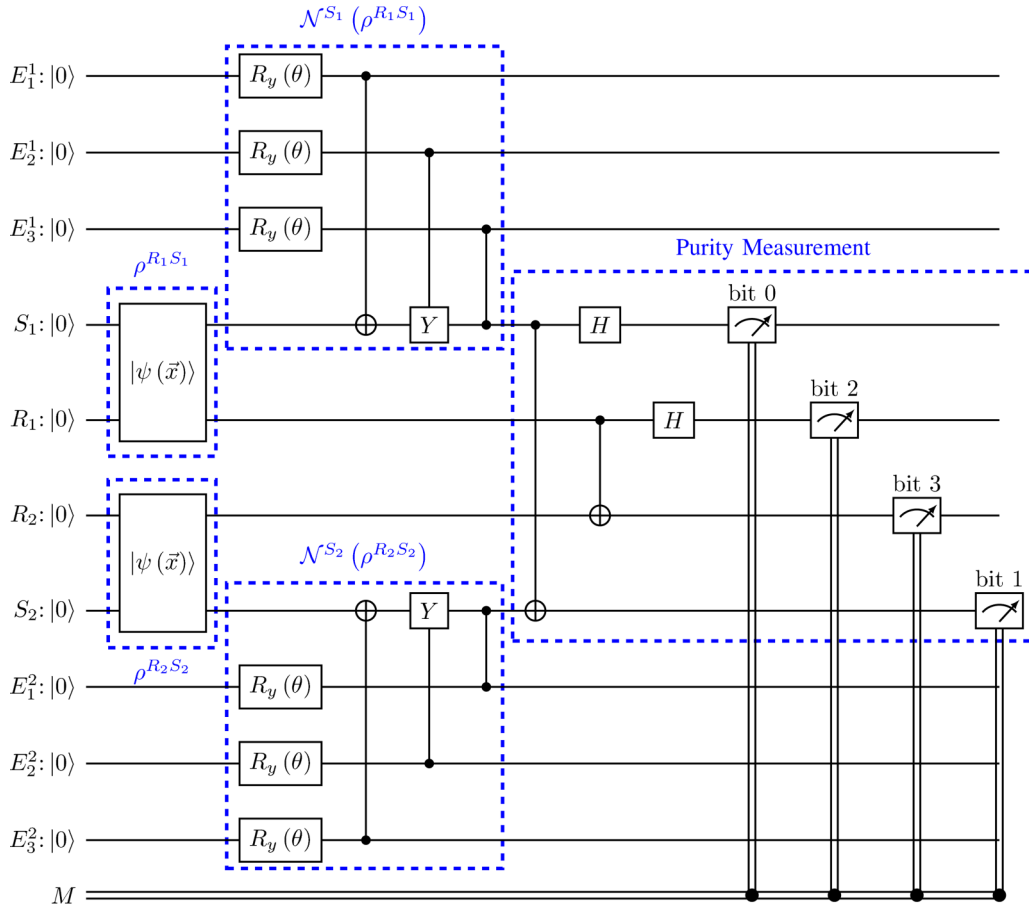
FIG. 5. Quantum circuit for depolarizing channel. Purity measurement requires two copies of the state $I^R \otimes \mathcal{N}^S(\rho^{RS})$. We prepare these two copies on wires 4–7, labeled as S_1, R_1, R_2, and S_2, respectively.
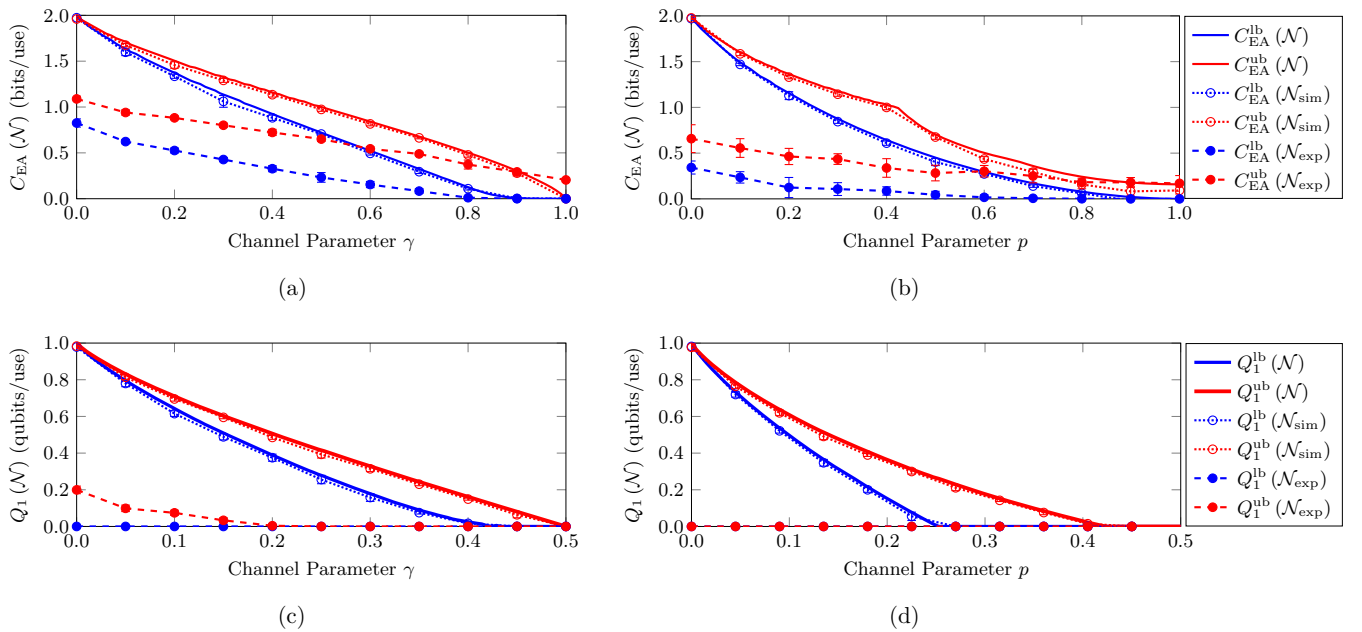

FIG. 6. Estimated lower and upper bounds on IBM QISKIT: (a) EACC bounds for the amplitude damping channel, (b) EACC bounds for the depolarizing channel, (c) coherent information bounds for the amplitude damping channel, and (d) coherent information bounds for the depolarizing channel. Solid lines show the proposed bounds for these cases. Dotted lines with empty markers ($\mathcal{N}_{\text{sim}}$) and dashed lines with solid markers ($\mathcal{N}_{\text{exp}}$) show the estimated bounds on QASM and FakeMelbourne() simulators, respectively.
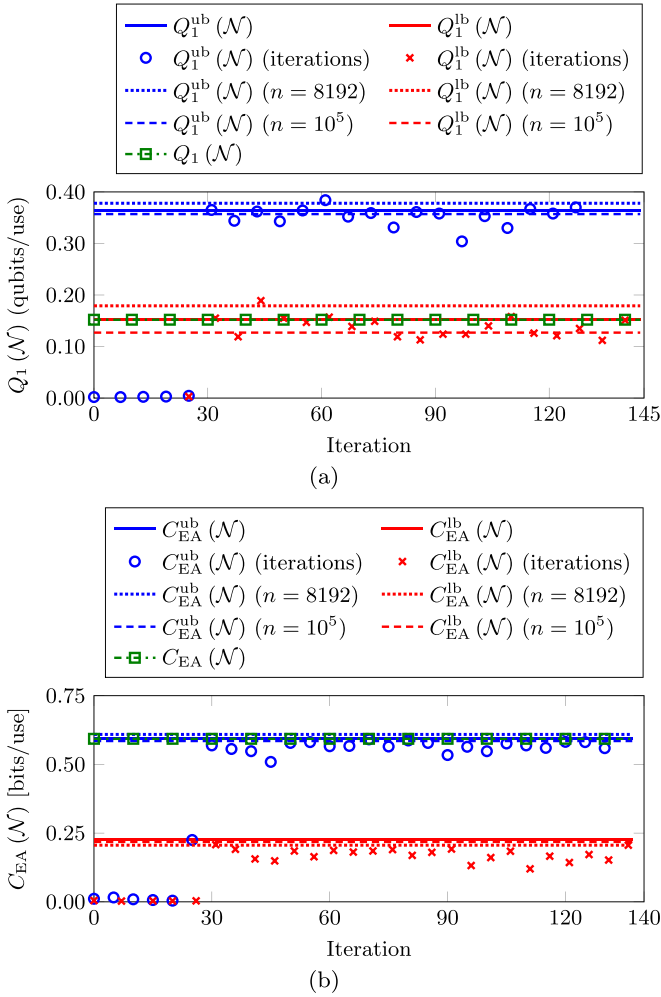
FIG. 7. Optimization of capacity bounds. An instance of optimization for capacity bounds is shown for (a) the quantum capacity of the depolarizing channel with $p = 0.2$ and (b) the entanglement-assisted classical capacity of the amplitude damping channel with $\gamma = 0.75$. Solid lines show the proposed bounds. Blue circles and red crosses respectively show the upper and lower bounds at a particular iteration of the variational method with $n = 8192$ shots per measurement. Dotted lines show the final value returned by the optimizer at the end of all iterations. Dashed lines show the bounds estimated by $n = 10^5$ shot measurement on the optimal solution returned by the optimizer. Green lines with square markers show the exact capacity values of these channels.

FIG. 8. Achievable rate of quantum information by states optimal for the lower and upper bounds. For (a) the amplitude damping channel, ρ_{ub} achieves the coherent information. On the other hand, both ρ_{ub} and ρ_{lb} achieve the coherent information for (b) the depolarizing and (c) the Pauli (exponential correlation) channel.

improvement in the objective function value was smaller than the default tolerance of `ftol = 0.0001`. For each fixed channel, i.e., a fixed value of γ, η, ξ, or p, we obtained ten estimates, whose means and standard deviations are reported in Fig. 6. In these experiments, we performed iterations with `n_shots = 8192`. However, once we obtained the optimal state as a final result, we performed one more iteration with `n_shots = 100 000`. This was done to reduce the statistical error in the final estimate and obtain a better assessment of the quality of the obtained result. The plots provided in Fig. 6 are results of this final iteration.

From these results, we note that the obtained estimates on the `FakeMelbourne()` simulator are considerably lower than
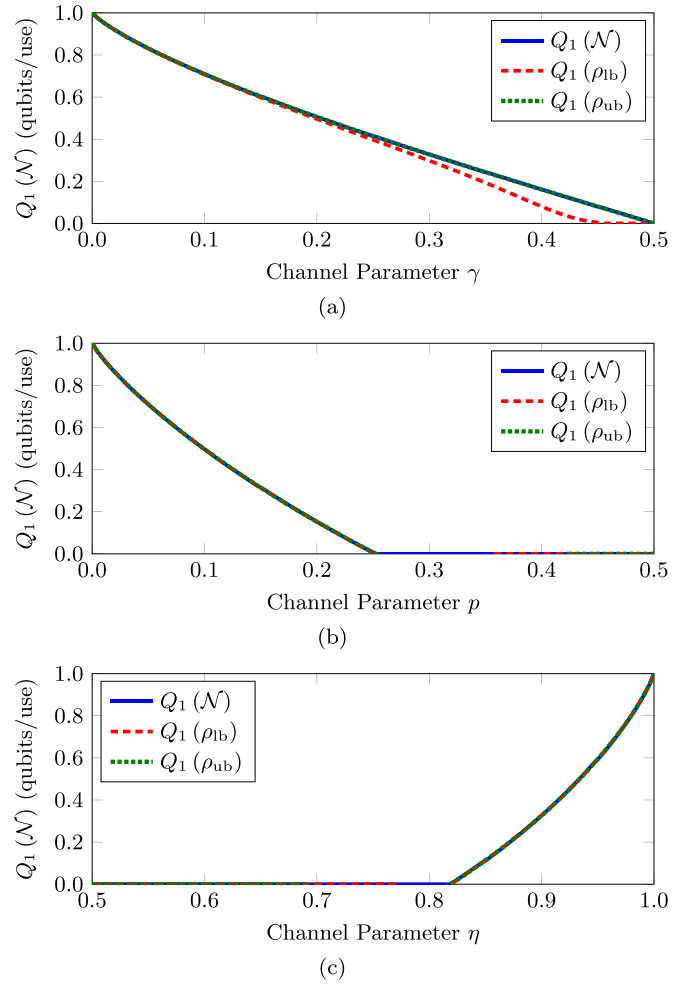
the actual bounds for the channels we are trying to simulate. This is not very surprising since the noisy device introduces additional noise in the system, as well as in the reference qubit. Thus, the channel we actually simulate is far noisier than the channel we are attempting to simulate. Furthermore, since the circuit for the depolarizing channel has larger width (number of involved qubits) and depth (roughly, the largest number of gates on a qubit from start to end), the amount of noise that can affect the final results is larger for this case. For this reason, the capacity estimates for the depolarizing channel are worse than the corresponding estimates for the amplitude damping channels.

Figures 7(a) and 7(b) show a single run of the variational capacity estimation method for the quantum capacity of the depolarizing channel ($p = 0.2$) and the entanglement-assisted classical capacity of the amplitude damping channel ($\gamma = 0.75$), respectively. We can clearly see the effect of statistical noise due to a modest number of shots (8192) per measurement. Furthermore, we notice that these effects of statistical noise can be well compensated by performing one final
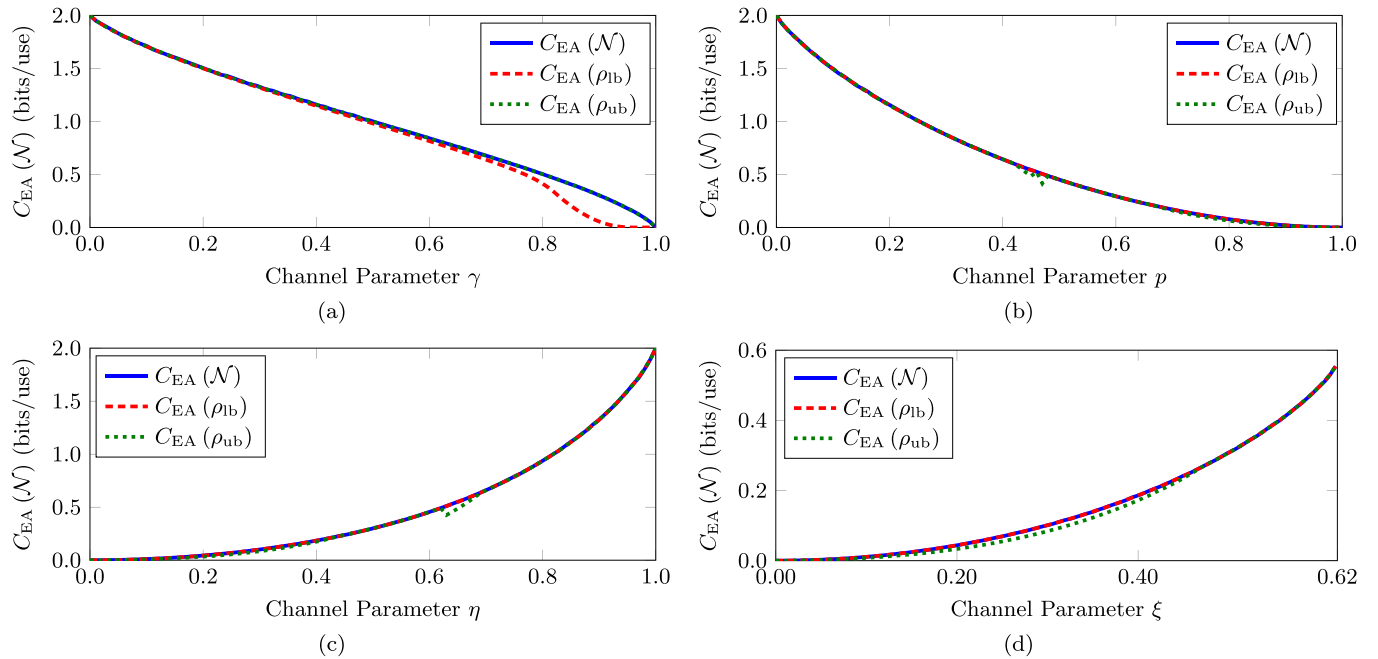
FIG. 9. Achievable rate of the entanglement-assisted classical communication by states optimal for the lower and upper bounds. For (a) the amplitude damping channel, ρ_{ub} achieves the EACC. On the other hand, both ρ_{ub} and ρ_{lb} achieve the capacity for (b) the depolarizing channel and (c) the Pauli (exponential correlation) channel. For (d) the Pauli (constant tridiagonal correlation) channel, ρ_{lb} saturates the EACC, whereas ρ_{lb} saturates it only for a small range of channel parameters.

measurement with $n = 10^5$ shots on the optimal state returned by the optimizer. We remark that for this particular example, we initialized our ansatz (26) with uniform random parameters to demonstrate the ability of the optimizer to converge to a good solution despite a rough starting point. Finally, note that some of the data points for the initial few iterations are not shown in plots, e.g., first four data points for the lower bound in Fig. 7(a). That is because we allow negative values for the bounds during the optimization. This allowance improves the feedback to the classical optimizer to produce better parameter values in the next iterations. If we clip the negative intermediate values to zero, the optimizer will not be able to assess the effect of modifying each of the optimization parameters.

B. Achievable rates by states optimal for bounds

It is natural to question the achievable communication rates offered by the states optimal for the lower and the upper bounds. Here, we provide numerical examples of achievable communication rates by utilizing the states that optimize the proposed bounds. We denote by ρ_{lb} and ρ_{ub} the states optimal for the lower and the upper bounds for some capacity C. With a slight abuse of notation, we denote by $C(\rho_{lb})$ and $C(\rho_{ub})$ the achievable communication rates using these states.

In Figs. 8 and 9, we plot the achievable communication rates by using the states optimizing the lower and upper bounds. Surprisingly, the states optimized for the lower and upper bounds mostly saturate the corresponding capacities. Therefore, one can utilize these states to estimate the exact capacities by estimating the entropy at the channel output without needing the optimization. Also, this reassures that one does not sacrifice much of the communication rate if

one decides to utilize these optimal states in the design of communication protocols and error-correcting codes.

IV. CONCLUSIONS

We have developed a general variational framework for estimating capacity bounds of quantum channels. The main advantage of the proposed framework is the possibility of obtaining sharp bounds on the aforementioned quantities without any *a priori* knowledge, e.g., the channel structure or the preferred basis of dephasing of the quantum channel. The proposed framework provides a measurement-basis-independent technique aided by the variational approach to estimate lower bounds on the quantum capacity, the entanglement-assisted classical capacity, and the classical capacity of quantum channels. Correspondingly, we obtained upper bounds on the N-shot coherent information, the entanglement-assisted classical capacity, and the N-shot Holevo capacity of the quantum channels. The proposed bounds are completely generic and are applicable to arbitrary quantum channels, except for the classical capacity lower bound, which is applicable only for the unital qubit and qudit covariant quantum channels. Numerical examples, including IBM QISKIT experiments, of these bounds for different important classes of quantum channels demonstrate the effectiveness of this framework for different capacity notions of quantum channels.

The utility of the proposed framework is twofold. First, this framework is useful because of the scenario that has been discussed in detail throughout this paper, i.e., capacity estimation for unknown but physically available quantum channels. Second, this framework is also useful for capacity calculation of large quantum channels where numerical calculations are

not possible on a classical computer due to the large size of involved matrices and to the optimization problem. This second scenario is exactly the objective of the original proposal of hybrid quantum-classical algorithms.

Since the proposed framework, like other variational algorithms, utilizes heuristics-based methods for optimization, it is susceptible to the challenges commonly encountered in this class of algorithms. These challenges include difficulties in ansatz selection; optimization problems, e.g., barren plateaus, local optima, and insufficient iterations; statistical errors; and hardware noise in addition to what channel is being simulated. General progress in quantum computing hardware and in the domain of hybrid quantum-classical algorithms will help overcome these challenges for the proposed framework as well.

ACKNOWLEDGMENTS

[1] A. Acín, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, and F. K. Wilhelm, New J. Phys. **20**, 080201 (2018).

[2] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, Nat. Rev. Phys. **2**, 382 (2020).

[3] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2002).

[4] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).

[5] C. Macchiavello and M. Rossi, Phys. Rev. A **88**, 042335 (2013).

[6] C. Macchiavello and M. F. Sacchi, Phys. Rev. Lett. **116**, 140501 (2016).

[7] A. Cuevas, M. Proietti, M. A. Ciampini, S. Duranti, P. Mataloni, M. F. Sacchi, and C. Macchiavello, Phys. Rev. Lett. **119**, 100502 (2017).

[8] C. Pfister, M. A. Rol, A. Mantri, M. Tomamichel, and S. Wehner, Nat. Commun. **9**, 27 (2018).

[9] D. Ristè, S. Poletto, M.-Z. Huang, A. Bruno, V. Vesterinen, O.-P. Saira, and L. DiCarlo, Nat. Commun. **6**, 6983 (2015).

[10] J. ur Rehman, Y. Jeong, and H. Shin, Phys. Rev. A **99**, 042312 (2019).

[11] C. Macchiavello and M. F. Sacchi, Phys. Rev. Lett. **123**, 090503 (2019).

[12] M. A. Ciampini, Á. Cuevas, P. Mataloni, C. Macchiavello, and M. F. Sacchi, Phys. Rev. A **103**, 062414 (2021).

[13] T. Li and X. Wu, IEEE Trans. Inf. Theory **65**, 2899 (2019).

[14] E. M. Kontopoulou, G. P. Dexter, W. Szpankowski, A. Grama, and P. Drineas, IEEE Trans. Inf. Theory **66**, 5003 (2020).

[15] J. Acharya, I. Issa, N. V. Shende, and A. B. Wagner, IEEE J. Sel. Areas Inf. Theory **1**, 454 (2020).

[16] S. Johri, D. S. Steiger, and M. Troyer, Phys. Rev. B **96**, 195136 (2017).

[17] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, New J. Phys. **20**, 113022 (2018).

[18] Y. Subaşı, L. Cincio, and P. J. Coles, J. Phys. A: Math. Theor. **52**, 044001 (2019).

[19] R. B. A. Adamson, L. K. Shalm, and A. M. Steinberg, Phys. Rev. A **75**, 012104 (2007).

[20] J. C. Garcia-Escartin and P. Chamorro-Posada, Phys. Rev. A **87**, 052330 (2013).

[21] R. B. Patel, J. Ho, F. Ferreyrol, T. C. Ralph, and G. J. Pryde, Sci. Adv. **2**, e1501531 (2016).

[22] D. W. Berry and B. C. Sanders, J. Phys. A: Math. Gen. **36**, 12255 (2003).

[23] P. Harremoes and F. Topsoe, IEEE Trans. Inf. Theory **47**, 2944 (2001).

[24] T.-C. Wei, K. Nemoto, P. M. Goldbart, P. G. Kwiat, W. J. Munro, and F. Verstraete, Phys. Rev. A **67**, 022110 (2003).

[25] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, Nature (London) **528**, 77 (2015).

[26] H. Shin and M. Z. Win, IEEE Trans. Inf. Theory **54**, 2976 (2008).

[27] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).

[28] H. Barnum, M. A. Nielsen, and B. Schumacher, Phys. Rev. A **57**, 4153 (1998).

[29] I. Devetak and P. W. Shor, Commun. Math. Phys. **256**, 287 (2005).

[30] P. Hayden, M. Horodecki, A. Winter, and J. Yard, Open Syst. Inf. Dyn. **15**, 7 (2008).

[31] N. Ramakrishnan, R. Iten, V. Scholz, and M. Berta, IEEE Trans. Inf. Theory **67**, 946 (2021).

[32] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).

[33] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).

[34] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, IEEE Trans. Inf. Theory **48**, 2637 (2002).

[35] C. Adami and N. J. Cerf, Phys. Rev. A **56**, 3470 (1997).

[36] J. Cortese, Phys. Rev. A **69**, 022302 (2004).

[37] H. Abraham et al., Qiskit: An open-source framework for quantum computing (2019), doi: 10.5281/zenodo.2573505.

[38] G. García-Pérez, M. A. C. Rossi, and S. Maniscalco, npj Quantum Inf. **6**, 1 (2020).

[39] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5**, 4213 (2014).

[40] Interested readers might refer to Refs. [41–43] for introduction and comparison of different optimizers for variational circuits.

[41] D. Lee, J. Lee, S. Hong, H.-T. Lim, Y.-W. Cho, S.-W. Han, H. Shin, J. ur Rehman, and Y.-S. Kim, Optica **9**, 88 (2022).

[42] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Quantum Sci. Technol. **4**, 014008 (2018).

[43] A. Pellow-Jarman, I. Sinayskiy, A. Pillay, and F. Petruccione, Quantum Inf. Proc. **20**, 202 (2021).