Communication Complexity of One-Shot Remote State Preparation *

Shima Bab Hadiashar † U. Waterloo
Ashwin Nayak ‡ U. Waterloo
Renato Renner § ETH Zurich

February 2018

Abstract

Quantum teleportation uses prior shared entanglement and classical communication to send an unknown quantum state from one party to another. Remote state preparation (RSP) is a similar distributed task in which the sender knows the entire classical description of the state to be sent. (This may also be viewed as the task of *non-oblivious* compression of a single sample from an ensemble of quantum states.) We study the communication complexity of *approximate* remote state preparation, in which the goal is to prepare an approximation of the desired quantum state.

Jain [Quant. Inf. & Comp., 2006] showed that the worst-case communication complexity of approximate RSP can be bounded from above in terms of the *maximum possible information* in an encoding. He also showed that this quantity is a lower bound for communication complexity of (*exact*) remote state preparation. In this work, we tightly characterize the worst-case and average-case communication complexity of remote state preparation in terms of non-asymptotic information-theoretic quantities.

We also show that the average-case communication complexity of RSP can be much smaller than the worst-case one. In the process, we show that n bits cannot be communicated with less than n transmitted bits in LOCC protocols. This strengthens a result due to Nayak and Salzman [J. ACM, 2006] and may be of independent interest.

1 Introduction

Quantum teleportation [3] is an archetypical protocol in information processing that is impossible in the absence of quantum resources like shared entanglement. Through quantum teleportation, one party is able to communicate an *arbitrary* qubit state to another party using only two classical bits of communication and a previously shared maximally entangled pair of qubits. The two classical bits of communication and a maximally entangled pair of qubits are both necessary and sufficient for the task. This is a remarkable phenomenon, as the entire classical description of the state being communicated is potentially infinite in length.

In Ref. [33], Lo introduced a similar distributed task in which the sender (called Alice in the literature) knows a classical description of the quantum state. This task is called *remote state preparation* (RSP). In particular,

*Much of the work in this article was reported in S.B.'s Master's thesis [2].

†Institute for Quantum Computing, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 3G1, Canada. Email: sbabhadi@uwaterloo.ca . Research supported in part by NSERC Canada.

‡Department of Combinatorics and Optimization, and Institute for Quantum Computing, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 3G1, Canada. Email: ashwin.nayak@uwaterloo.ca . Research supported in part by NSERC Canada.

§Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Str. 27, 8093 Zurich, Switzerland. Email: renner@itp.phys.ethz.ch .

remote state preparation is a task involving two parties, Alice and Bob, who share qubits in an entangled state. Alice is given the description of a state, $Q(x)$, chosen from a subset of quantum states $\{Q(1), \ldots, Q(n)\}$, and their goal is to prepare that quantum state on Bob's side using only local operations and classical communication (LOCC). This may also be viewed as the task of compression (which is non-oblivious at the sender's end), of a single sample from an ensemble of quantum states with entanglement-assisted classical communication.

We say an RSP protocol is *oblivious to Bob* if he can get no more information about the prepared state than what is contained in a single copy of the state [32]. A relaxed version of RSP is *approximate remote state preparation* (ARSP) in which we wish to prepare an approximation σ_x of the specified quantum state $Q(x)$. We define the error of a protocol for approximate remote state preparation in terms of the fidelity between $Q(x)$ and σ_x. We say a protocol has *worst-case error* at most ϵ, if for every $x \in \{1, \ldots, n\}$, $\mathrm{F}(Q(x), \sigma_x) \geq \sqrt{1-\epsilon^2}$. Similarly, a protocol has *average-case error* at most ϵ with respect to a probability distribution p, if $\sum_{x=1}^{n} p_x \mathrm{F}(Q(x), \sigma_x) \geq \sqrt{1-\epsilon^2}$.

Lo [33] gave several examples of ensembles which can be remotely prepared using a one-way communication protocol with classical communication cost less than that in quantum teleportation. However, he conjectured that to prepare arbitrary pure n-qubit states remotely, Alice has to necessarily send the same number of classical bits as in quantum teleportation i.e., $2n$ classical bits. The task has been studied extensively since then, largely in the asymptotic setting.

Bennett *et al.* [4] showed that in the presence of a large amount of shared entanglement, Alice can prepare general quantum states on Bob's side with the asymptotic classical communication rate of one bit per qubit. This amount of classical communication from Alice to Bob is also necessary by causality [33]. They also showed that unlike for quantum teleportation, there is a trade-off between the communication cost and the amount of entanglement in remote state preparation. In particular, they proved that at the cost of using more entanglement, the communication cost of preparing a one-qubit state ranges from one bit in the high entanglement limit to an infinite number of bits in the case of no previously shared entanglement. In addition, they suggested that the Lo conjecture is true in a more restricted setting, such as when the protocol is *faithful* and oblivious to Bob [4]. (A protocol is said to be *faithful* if it is exact and deterministic.)

Devetak and Berger [17] found an analytic expression for the trade-off curve between the shared entanglement and classical communication of *teleportation based RSP protocols* in the *low-entanglement* region (less than 1 singlet state per qubit). They conjectured that teleportation based protocols are optimal among all low-entanglement protocols. Later, Leung and Shor [32] proved the Lo conjecture for a special case. They proved that if a one-way RSP protocol for a *generic ensemble* of pure states is faithful and oblivious to Bob, then it necessarily uses at least as much classical communication as in teleportation. (A *generic ensemble* is an ensemble of states whose density matrices span the operators in the input Hilbert space.) Hayashi, Hashimoto and Horibe [21] showed that in order to remotely prepare one qubit in an arbitrary state using a one-way faithful, but not necessarily oblivious protocol, Alice requires two classical bits of communication as in teleportation.

Berry and Sanders [7] studied ARSP, the approximation variant of RSP, of an ensemble $\mathscr{E}$ of mixed states (which might be entangled with some other system on Alice's part) such that their entanglement with other systems does not change significantly. They showed that approximate remote state preparation with arbitrary small average-case error ϵ can be done asymptotically using communication per prepared state arbitrarily close to the Holevo information $\chi(\mathscr{E})$ of the ensemble. (See Section 2.4 for a definition of Holevo information.) Later Bennett, Hayden, Leung, Shor, and Winter [6] proved that approximate remote state preparation with small worst-case error ϵ requires an asymptotic rate of one bit of classical communication per qubit from Alice to Bob. They also showed that this amount of classical communication is sufficient. Moreover, they derived the exact trade-off curve between shared entangled bits and classical communication bits for an arbitrary ensemble of candidate states.

Jain [25] studied remote state preparation in the one-shot scenario. He considered the total communication cost when given access to an arbitrary amount of entanglement. He showed that the communication cost

Protocol Type	Conditions	Entanglement	Classical Communication
Faithful RSP [4]	an arbitrary state, one-way communication, in asymptotics	high entanglement	= 1 classical bit per qubit
Faithful RSP [21]	one pure qubit in a general state, one-way communication	= 1 ebit(singlet) per qubit	= 2 classical bit
Faithful and oblivious RSP [32]	a generic ensemble of pure states, one-way communication	= 1 ebit(singlet) per qubit	= 2 classical bit per qubit
ARSP with small average-case error [7]	an ensemble $\mathscr{E}$ of mixed states preserving their entanglement, one-way communication, in asymptotics	no limit	$\approx \chi(\mathscr{E})$ classical bits per prepared state
ARSP with small worst-case error [6]	an arbitrary pure state, two-way communication, in asymptotics	= 1 ebit(singlet) per qubit	= 1 classical bit per qubit from Alice to Bob
Exact RSP [25]	an arbitrary state, two-way communication, in one-shot scenario	no limit	$\geq \mathsf{T}(Q)/2$
ARSP with worst-case error ϵ [25]	an arbitrary state, one-way communication, in one-shot scenario	no limit	$\leq \frac{8}{\left(1-\sqrt{1-\epsilon^2}\right)^2}(4\mathsf{T}(Q)+7)$

Table 1: A summary of previous works on communication cost of Remote State Preparation

required for exact remote state preparation is at least $\mathsf{T}(Q)/2$ and ARSP with worst-case error at most ϵ can be accomplished with communication at most $\frac{8}{\left(1-\sqrt{1-\epsilon^2}\right)^2}(4\mathsf{T}(Q)+7)$, where $\mathsf{T}(Q)$ denotes the *maximum possible information* in an encoding Q. (A precise definition can be found in Section 2.4.)

These abovementioned results on remote state preparation are summarized in Table 1 .

1.1 Our results

Intuitively, relaxing the remote state preparation problem so that Bob produces some approximation to the ideal state should lower the communication complexity of the task. This suggests that the bounds provided by Jain [25] are not tight.

In this work, we characterize the communication complexity of remote state preparation in two different cases. First, we consider ARSP with average-case error at most ϵ, and bound its communication complexity by the *smooth max-information* Bob has about Alice's input. (See Section 2.4 for a precise definition of this quantity.) Then we consider ARSP with worst-case error at most ϵ, and give lower and upper bounds for its communication complexity in terms of *smooth max-relative entropy* and show that these bounds may be arbitrarily tighter than that in Ref. [25].

Our main results about the remote state preparation problem are summarized below, using notions introduced in Section 2. Recall that a protocol has worst-case error at most ϵ, if for every $x \in \{1,\ldots,n\}$, $\mathrm{F}(Q(x),\sigma_x) \geq \sqrt{1-\epsilon^2}$, and a protocol has average-case error at most ϵ with respect to a probability distribution p, if $\sum_{x=1}^{n} p_x \mathrm{F}(Q(x),\sigma_x) \geq \sqrt{1-\epsilon^2}$. We denote the average-case communication complexity of ARSP by $\mathsf{Q}_p^*(\mathrm{RSP}(S,Q),\epsilon)$, and the worst-case communication complexity of ARSP by $\mathsf{Q}^*(\mathrm{RSP}(S,Q),\epsilon)$.

Theorem 1.1. *For any finite set S, and set of quantum states $\{Q(x) : x \in S\}$, let p be a probability distribution over S and $\rho_{AB}(p) \in \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ be the bipartite quantum state $\rho_{AB}(p) = \sum_{x\in S} p_x |x\rangle\langle x|_A \otimes Q(x)_B$. Then*

1. *For any fixed $\epsilon \in (0,1]$, we have*
$$\mathrm{I}^{\epsilon}_{\max}(A:B)_{\rho(p)} \quad \leq \quad \mathsf{Q}^*_p(\mathrm{RSP}(S,Q),\epsilon) \quad \leq \quad \mathrm{I}^{\frac{\epsilon}{2\sqrt{2}}}_{\max}(A:B)_{\rho(p)} + f(\epsilon) \ ,$$
where $f(\epsilon) \in \Theta(\log\log\frac{1}{\epsilon})$ is a function of ϵ, and $\mathrm{I}^{\epsilon}_{\max}(A:B)$ denotes the smooth max-information part B has about part A.

2. *For any fixed $\epsilon \in (0,1]$ and for any $0 < \delta < 1-\epsilon^2$, we have*
$$\min_{\sigma\in\mathsf{D}(\mathcal{H})} \max_{x\in S} \mathrm{D}^{\sqrt{2(\epsilon^2+\delta)}}_{\max}(Q(x)\|\sigma) + g_1(\epsilon,\delta) \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(S,Q),\epsilon)$$
$$\leq \quad \min_{\sigma\in\mathsf{D}(\mathcal{H})} \max_{x\in S} \mathrm{D}^{\frac{\epsilon}{\sqrt{1+\epsilon^2}}}_{\max}(Q(x)\|\sigma) + g_2(\epsilon) \ ,$$
where g_1, g_2 are functions such that $g_1(\epsilon,\delta) \in \Theta\left(\log\frac{\delta^3}{\epsilon^2+\delta}\right)$, $g_2(\epsilon) \in \Theta(\log\log\frac{1}{\epsilon})$, and $\mathrm{D}^{\epsilon}_{\max}(Q(x)\|\sigma)$ denotes the smooth max-relative entropy of $Q(x)$ with respect to σ.

It is relatively straightforward to show that the one-shot information expressions appearing in the above theorem are continuous in ϵ. This indicates the tightness of the characterization. In fact, a bound on the difference between lower and upper bounds in the above theorem, in terms of the ensemble, may be inferred from the continuity property.

We remark that the quantity appearing in the second part of the theorem is similar to the notion of *information radius*. It may be possible to relate the quantity to smooth max-information with respect to a distribution over S using ideas from Ref. [18, Lemma 3] (which extends Ref. [50, Lemma 14]), and the connection between max-relative entropy and the *sandwiched Rényi relative entropy*. Finally, earlier works have considered remote state preparation of states drawn from infinite sets of states. We discuss how the bounds in Theorem 1.1 may be applied to that case in Appendix B.

The communication cost of ARSP may decrease dramatically when more error is allowed,and if we consider average-case error instead of worst-case error. In particular, we show that for every $\epsilon \in [0, \frac{1}{\sqrt{2}})$, there exists a set of n quantum states for which there is a $\log n$ gap between the worst-case error and average-case error remote preparation of that set. In addition, for a special set of quantum states, we derive a gap between the worst-case error and average-case communication complexity in terms of ϵ. This confirms our intuition that the more skewed the probability distribution is, the bigger the gap between worst-case and average-case error variants may be.

In the process of establishing the first gap described above, we strengthen a result due to Nayak and Salzman [35]; we prove a bound on the communication required by any LOCC protocol for transmitting a uniformly random n bit string with some probability p. This bound is optimal, and may be of independent interest.

Theorem 1.2. *Let Y be the output of Bob in any two-way LOCC protocol in which Alice receives a uniformly distributed n-bit input X (that is not known to Bob, and is independent of their joint quantum state). Let m_A be the total number of bits Alice sends to Bob and $p := \Pr[Y = X]$ be the probability that Bob obtains the output X. Then*
$$m_A \quad \geq \quad n + \log p \ .$$

Worst-case protocols for ARSP capture precisely the task of compression in one-way communication complexity. Average-case protocols for ARSP are relevant in the distributional setting in communication complexity, and in asymptotic information theory. The results in this paper thus supercede those due to Jain, Radhakrishnan, and Sen [29] (and due to Touchette [47] for the same setting). We also show how a characterization due to Berry and Sanders [7] may be reproduced from ours, via a quantum asymptotic equipartition property (cf. Theorem 2.6). Thus, we believe the results presented here have wider ramifications.

1.2 Organization

The organization of this paper is as follows. In Section 2, we review some concepts, fix notation, and the terminology used in the paper. Then we define remote state preparation, and explain an efficient protocol for this problem introduced in Ref. [29]. In Section 3 and Section 4, we give bounds on average-case error and worst-case communication complexity of ARSP, respectively. We make some observations, including a comparison with previously known results in Section 5. We analyze LOCC protocols for communicating a uniformly random n bit string in Section 6. The paper ends with a summary of our results and an outlook in Section 7. In the Appendix, we present the proofs of some properties of information-theoretic quantities, and discuss remote state preparation of states drawn from an infinite set.

Acknowledgments

We are grateful to Matthias Christandl for discussions which led to the research reported in this article. S.B. thanks Marco Tomamichel for his help with the proof of Theorem 4.2 during her internship at CQT, Singapore. We also thank the reviewers and the Associate Editor, Mark Wilde, for their comments and suggestions.

2 Preliminaries

In this section, we review some notions in quantum computing and quantum information theory, such as LOCC protocols, quantum communication complexity, asymptotic and non-asymptotic quantum information theory, as well as some mathematical tools like the minimax theorem. We also define remote state preparation formally and describe a non-trivial protocol for this problem. We refer the reader to the books by Nielsen and Chuang [37] and Watrous [49] for basic notions and results in quantum information, and largely only describe the potentially non-standard notation and terminology we use.

2.1 Some basic notions

We denote Hilbert spaces either by capital script letters like $\mathcal{H}$ and $\mathcal{K}$, or as $\mathbb{C}^m$ where m is the dimension of the Hilbert space. We concern ourselves only with finite dimensional Hilbert spaces in this article. We denote the set of all linear operators from $\mathcal{H}$ to $\mathcal{K}$ by $\mathsf{L}(\mathcal{H}, \mathcal{K})$. We abbreviate $\mathsf{L}(\mathcal{H}, \mathcal{H})$ as $\mathsf{L}(\mathcal{H})$. We denote the set of all positive semidefinite operators in $\mathcal{H}$ by $\mathsf{Pos}(\mathcal{H})$. An operator A is called *sub-normal* if it is positive semidefinite and has trace at most 1. (The term "subnormalized" is also often used for such operators.)

We denote the identity operator on a Hilbert space by $\mathbb{1}$ and the set of all unitary operators on space $\mathcal{H}$ by $\mathsf{U}(\mathcal{H})$.

We call a physical quantum system with a finite number of degrees of freedom a *register*. Every register is associated with a Hilbert space. We denote registers by capital letters, e.g., X, Y and Z. We use the notation $|X|$ to denote the dimension of the Hilbert space associated with register X. The state of a register X is modelled as a *density operator*, i.e., a positive semidefinite operator with trace one, and is called a *quantum state*. We denote density operators by lower case Greek letters (e.g., ρ, σ, $\dots$), and the set of all density operators over a Hilbert space $\mathcal{H}$ by $\mathsf{D}(\mathcal{H})$. We may also denote a state by ρ_X to indicate its register X. A bipartite register XY with Hilbert space $\mathcal{H} \otimes \mathcal{K}$ is called a *classical-quantum* register in the context of an information processing task, if it only assumes states of the form $\sum_i p_i |e_i\rangle\langle e_i| \otimes \rho_i$ where $\{|e_i\rangle\}$ is the standard basis of $\mathcal{H}$ and p is a probability distribution over the basis. In that case we say that the states are classical on X. For any $\omega \in \mathsf{Pos}(\mathcal{H})$ with spectral decomposition $\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, we let $\sqrt{\omega} = \sum_i \sqrt{\lambda_i} |\psi_i\rangle\langle\psi_i|$.

We denote the partial trace over Hilbert space $\mathcal{K}$ of a quantum state $\rho_{AB} \in \mathsf{D}(\mathcal{H} \otimes \mathcal{K})$ by either $\mathrm{Tr}_{\mathcal{K}}(\rho_{AB})$ or $\mathrm{Tr}_B(\rho_{AB})$. We say that $\rho_{AB} \in \mathsf{D}(\mathcal{H} \otimes \mathcal{K})$ is an *extension* of $\rho_A \in \mathsf{D}(\mathcal{H})$ if $\mathrm{Tr}_{\mathcal{K}}(\rho_{AB}) = \rho_A$.

We call completely positive and trace preserving linear maps $\mathsf{L}(\mathcal{H}) \to \mathsf{L}(\mathcal{K})$ *quantum channels*. Quantum measurements are quantum channels with Kraus operators $\{\sqrt{E_a} \otimes |a\rangle : a \in \Gamma\}$, where Γ is the set of outcomes of the measurement and E_a is a positive semidefinite operator associated with the outcome $a \in \Gamma$ such that $\sum_{a \in \Gamma} E_a = \mathbb{1}$. We refer to the operators E_a as *measurement operators*.

The *fidelity* $\mathrm{F}(\rho, \sigma)$ between two quantum states ρ and σ, is defined as

$$\mathrm{F}(\rho, \sigma) \quad := \quad \mathrm{Tr}\sqrt{\sqrt{\rho}\,\sigma\sqrt{\rho}} \ .$$

In the literature, fidelity is sometimes defined as the square of the above quantity. Fidelity may be extended to sub-normal states ρ, σ as follows:

$$\mathrm{F}(\rho, \sigma) \quad := \quad \mathrm{Tr}\sqrt{\sqrt{\rho}\,\sigma\sqrt{\rho}} + \sqrt{(1 - \mathrm{Tr}(\rho))\,(1 - \mathrm{Tr}(\sigma))} \ .$$

The fidelity function is monotone under the application of quantum channels, and is jointly concave over the set of quantum states. Other useful properties of fidelity are stated in the following propositions.

Proposition 2.1. *For any quantum state ρ and sub-normal state σ, it holds that*

$$\mathrm{F}(\rho, \sigma)^2 \quad \leq \quad \mathrm{Tr}(\sigma) \ .$$

Proposition 2.2. *Let $\rho, \sigma \in \mathsf{D}(\mathcal{H})$ be two quantum states. Then*

$$1 + \mathrm{F}(\rho, \sigma) \quad = \quad \max\left\{\mathrm{F}(\rho, \xi)^2 + \mathrm{F}(\sigma, \xi)^2 : \xi \in \mathsf{D}(\mathcal{H})\right\} \ .$$

For a proof of the above property, see Ref. [36, Lemma 3.3].

We use the *purified distance* (see Ref. [45]) as a metric for sub-normal states. This is an extension of the metrics developed in Refs. [40, 41, 20, 42]. Suppose that ρ and σ are two sub-normal states. Then the purified distance of ρ and σ is defined as

$$\mathrm{P}(\rho, \sigma) \quad := \quad \sqrt{1 - \mathrm{F}(\rho, \sigma)^2} \ .$$

There are other metrics over sub-normal states, such as the trace distance. However, we choose purified distance since it turns out to be more convenient to use in non-asymptotic quantum information theory.

Let $\rho \in \mathsf{D}(\mathcal{H})$ be a quantum state and $\epsilon \in [0, 1)$. Then, we define

$$\mathsf{B}^{\epsilon}(\rho) \quad := \quad \{\tilde{\rho} \in \mathsf{Pos}(\mathcal{H}) : \mathrm{P}(\rho, \tilde{\rho}) \leq \epsilon, \mathrm{Tr}\,\tilde{\rho} \leq 1\}$$

as the ball of sub-normal states that are within purified distance ϵ of ρ. We say that σ is *ϵ-close* to ρ, or equivalently, σ is an *ϵ-approximation* of ρ, if $\sigma \in \mathsf{B}^{\epsilon}(\rho)$. The following property of purified distance states that any state ρ'_A that is ϵ-close to ρ_A may be extended to a state ρ'_{AB} that is ϵ-close to any given extension ρ_{AB} of ρ_A.

Proposition 2.3. *Let $\rho_A \in \mathsf{D}(\mathcal{H}_A)$ be a quantum state in the Hilbert space $\mathcal{H}_A$ and $\rho_{AB} \in \mathsf{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be an extension of ρ_A over the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, i.e. $\rho_A = \mathrm{Tr}_B(\rho_{AB})$. Let $\rho'_A \in \mathsf{B}^\epsilon(\rho_A)$ be an ϵ-approximation of ρ_A. Then there exists $\rho'_{AB} \in \mathsf{B}^\epsilon(\rho_{AB})$ such that $\rho'_A = \mathrm{Tr}_B(\rho'_{AB})$.*

Proof: Let $|v\rangle \in \mathsf{D}(\mathcal{H}_{A'} \otimes \mathcal{H}_{B'} \otimes \mathcal{H}_A \otimes \mathcal{H}_B)$ be a purification of ρ_{AB} and therefore also of ρ_A, and $|v'\rangle \in \mathsf{D}(\mathcal{H}_{A'} \otimes \mathcal{H}_{B'} \otimes \mathcal{H}_A \otimes \mathcal{H}_B)$ be a purification of ρ'_A, such that $\mathrm{F}(\rho_A, \rho'_A) = |\langle v|v'\rangle|$. Such $|v\rangle$ and $|v'\rangle$ exist by the Uhlmann theorem. Define $\rho'_{AB} = \mathrm{Tr}_{A'B'}(|v'\rangle\langle v'|)$. By definition, we have $\mathrm{F}(\rho_A, \rho'_A) = \mathrm{F}(\rho_{AB}, \rho'_{AB})$. Therefore $\rho'_{AB} \in \mathsf{B}^\epsilon(\rho_{AB})$. ■

The above property is in fact an extension of the Uhlmann theorem for purified distance.

2.2 LOCC protocols

The notion of *LOCC*, short for *local operations and classical communication*, plays an important role in quantum information, especially in the study of properties of entanglement (see, e.g., Ref. [5]). This notion has been described formally in terms of *quantum instruments* in Ref. [14]. In this article, we only study two-party LOCC protocols, in which one party receives a classical input, and the other party produces a quantum output. We describe these protocols informally below.

Suppose we have two parties, Alice and Bob, who communicate with each other using only classical bits, share parts of a possibly entangled quantum state, and are allowed to perform any local quantum channels on their registers. We call the registers (or qubits) accessible by only one of the parties *private* registers (or qubits). Alice is given a classical input; Bob does not receive any input. Let A be the register which holds Alice's input, $Y_0 \coloneqq P_0 V_0$ and $Z_0 \coloneqq Q_0 W_0$ be Alice's and Bob's initial classical-quantum private registers, respectively. Registers P_i and Q_i are classical registers with Alice and Bob, respectively, after the ith message. These registers hold the message transcript thus far. Initially, P_0, Q_0 are both empty. Registers V_0 and W_0 are initialized to a quantum state independent of the inputs. Note that the state in $V_0 W_0$ might be entangled across the registers. If there are k messages, P_{k+1} and V_{k+1} denote Alice's final classical and quantum registers, respectively, and $Q_{k+1} W_{k+1}$ denote Bob's, potentially after a local operation. Register A remains unchanged throughout the protocol. Bob produces the output, which is a sub-register B of $Q_{k+1} W_{k+1}$.

A *one-way* LOCC protocol is an LOCC protocol in which the communication consists of one message from Alice to Bob. The three steps of the protocol are:

1) Alice measures her register V_0, obtains the outcome in register P_1 (and a residual state in V_1). The measurement is controlled by her input in A.
2) Alice sends a copy of her measurement outcome to Bob, in classical register M. Bob sets $Q_1 = M$.
3) Bob measures his register W_1 (which is the same as W_0), controlled by the register Q_1. The outcome and residual state are stored in classical-quantum registers $Q_2 W_2$, where Q_2 includes Q_1. The output of the protocol is a designated sub-register B of his registers $Q_2 W_2$

A *two-way* LOCC protocol is a protocol with communication in both directions, from Alice to Bob and Bob to Alice. It has several rounds of communication in which the two parties alternately do a local measurement and send a message. Either party may start or end the protocol. Suppose in round i, it is Alice's turn. Then

- First, Alice measures her quantum register in that round, V_{i-1}, controlled by her input A and her classical register P_{i-1}. She copies the outcome M_i in a fresh register N_i. The register $P_i \coloneqq P_{i-1} N_i$.

- Alice then sends M_i to Bob using m_i classical bits, and Bob includes the received message M_i in his transcript register: $Q_i := Q_{i-1}M_i$.

Bob's actions are similar in a round in which it is his turn (except that he does not have any input), using registers Q_iW_i. At the end of a protocol with k rounds of communication, Bob makes a measurement on the quantum register W_k controlled by Q_k, and he includes the outcome M_{k+1} of the measurement in the register Q_{k+1}. A pre-designated sub-register B of $Q_{k+1}W_{k+1}$ is the output of the protocol.

2.3 Quantum communication complexity

Quantum communication complexity was introduced by Yao [51], and has been studied extensively since. Here we describe it in the context of LOCC protocols.

Let X, Y be two finite sets, Z be a set (not necessarily finite), and $f \subseteq X \times Y \times Z$ be a relation such that for every $(x, y) \in X \times Y$, there exists some $z \in Z$ such that $(x, y, z) \in f$. The sets X, Y, Z might be sets of quantum states. For example, in remote state preparation Z is the set of quantum states over some space. In an LOCC protocol, Alice and Bob get as their inputs $x \in X$ and $y \in Y$, respectively, and their goal is to output an element $z \in Z$ such that $(x, y, z) \in f$. In the protocols we consider, one party may not get any input, e.g., Y may be empty. Also, in general the output of the protocol is probabilistic. If $W_{x,y}$ is the random output that the protocol produces on inputs (x, y), we define the error of the protocol as

$$\delta \quad := \quad \max_{x \in X, y \in Y} \Pr((x, y, W_{x,y}) \notin f) \ .$$

We then say the protocol *computes f with error δ*.

Definition 2.1. *The entanglement-assisted communication complexity of f with error δ is defined as the minimum number of bits exchanged in an LOCC protocol computing f with error δ.*

Now consider a relation $f \subseteq X \times Y \times Z$, with $Z = \mathsf{D}(\mathcal{H})$, the set of quantum states over $\mathcal{H}$. In this context we may allow a protocol to produce an approximation to the desired quantum state. Suppose the output quantum state that an LOCC protocol for f produces on inputs (x, y) is denoted by w_{xy}. Let p be a probability distribution over $X \times Y$. We say a protocol computes an approximation of f with average-case error at most ϵ if there are quantum states $\{z_{xy} : x \in X, y \in Y, (x, y, z_{xy}) \in f\}$ such that

$$\sum_{x \in X, y \in Y} p_{xy} \, \mathrm{F}(w_{xy}, z_{xy}) \quad \geq \quad \sqrt{1 - \epsilon^2} \ .$$

The above condition may equivalently be written as $\mathrm{P}(\zeta, \omega) \leq \epsilon$, where $\zeta := \sum_{x,y} p_{xy} |xy\rangle\langle xy| \otimes z_{xy}$ is an ideal input-output state, and $\omega := \sum_{x,y} p_{xy} |xy\rangle\langle xy| \otimes w_{xy}$ is the actual input-output state of the protocol.

Definition 2.2. *The* average-case communication complexity *of f is defined as the minimum number of bits exchanged in an LOCC protocol computing an approximation of f with average-case error at most ϵ, and is denoted by $\mathsf{Q}^*_p(f, \epsilon)$.*

Similarly, we say a protocol computes an approximation of f with worst-case error at most ϵ if there are quantum states $\{z_{xy} : x \in X, y \in Y, (x, y, z_{xy}) \in f\}$ such that

$$\max_{x \in X, y \in Y} \mathrm{P}(w_{xy}, z_{xy}) \quad \leq \quad \epsilon \ .$$

Definition 2.3. *The* worst-case communication complexity *of f is defined as the minimum number of bits exchanged in an LOCC protocol computing an approximation of f with worst-case error at most ϵ, and is denoted by $\mathsf{Q}^*(f, \epsilon)$.*

Note that "error" here refers to the quality of approximation in the output state. The result of any probabilistic error made by the protocol is included in the output state, and hence this kind of error is reflected in the quality of approximation.

2.4 Quantum information theory

Let X be a register in quantum state $\rho \in \mathsf{D}(\mathcal{H})$. Then the *von Neumann entropy* $\mathrm{S}(\rho)$ of X is defined as

$$\mathrm{S}(\rho) \quad := \quad -\mathrm{Tr}(\rho \log \rho) \ .$$

Let X and Y be two registers in quantum states $\rho_X \in \mathsf{D}(\mathcal{H})$ and $\sigma_Y \in \mathsf{D}(\mathcal{H})$, respectively. The *relative entropy* denoted by $\mathrm{S}(\rho_X \| \sigma_Y)$ is defined as

$$\mathrm{S}(\rho_X \| \sigma_Y) \quad := \quad \mathrm{Tr}\left(\rho_X \log \rho_X - \rho_X \log \sigma_Y\right)$$

if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, and as ∞ otherwise. Suppose that $\rho_{XY} \in \mathsf{D}(\mathcal{H} \otimes \mathcal{K})$ is the joint state of registers X and Y, then the *mutual information* of X and Y is defined as

$$\mathrm{I}(X:Y)_\rho \quad := \quad \mathrm{S}(\rho_X) + \mathrm{S}(\rho_Y) - \mathrm{S}(\rho_{XY}) \ ,$$

where $\rho_X = \mathrm{Tr}_Y(\rho_{XY})$ and $\rho_Y = \mathrm{Tr}_X(\rho_{XY})$. When the register whose state is ρ is clear from the context, we may omit it from the subscript of ρ. Similarly, when the state ρ of the registers XY is clear from the context, we may omit it from the subscript of $\mathrm{I}(X:Y)$.

For $\rho, \sigma \in \mathsf{D}(\mathcal{H})$, the *observational divergence* [28] between ρ and σ is defined as

$$\mathrm{D}_{\mathrm{obs}}(\rho \| \sigma) \quad := \quad \sup \left\{ \mathrm{Tr}(M\rho) \log \frac{\mathrm{Tr}(M\rho)}{\mathrm{Tr}(M\sigma)} \ : \ 0 \leq M \leq \mathbb{1}, \mathrm{Tr}(M\sigma) \neq 0 \right\} \ .$$

Let $\mathscr{E} = \big((p_j, \rho_j) : 1 \leq j \leq n\big)$ be an ensemble of quantum states, i.e., $0 \leq p_j \leq 1$ for $1 \leq j \leq n$, $\sum_{j=1}^n p_j = 1$, and $\rho_j \in \mathsf{D}(\mathcal{H})$ are quantum states over the same space. The *Holevo information* of $\mathscr{E}$, denoted as $\chi(\mathscr{E})$, is defined as

$$\chi(\mathscr{E}) \quad := \quad \sum_{j=1}^n p_j \, \mathrm{S}(\rho_j \| \rho) \ ,$$

where ρ is the ensemble average, i.e., $\rho = \sum_{j=1}^n p_j \rho_j$. Similarly, we define the *divergence information* of $\mathscr{E}$, denoted as $\mathrm{D}_{\mathrm{obs}}(\mathscr{E})$, as

$$\mathrm{D}_{\mathrm{obs}}(\mathscr{E}) \quad := \quad \sum_{j=1}^n p_j \, \mathrm{D}_{\mathrm{obs}}(\rho_j \| \rho) \ .$$

Let S be a set, and $Q : S \to \mathsf{D}(\mathcal{H})$ be a function which "encodes" each $x \in S$ as a quantum state. Let p be a probability distribution over S, and $\rho_{AB}(p)$ be the bipartite state $\rho_{AB}(p) := \sum_x p_x |x\rangle\langle x|_A \otimes Q(x)_B$. We define the *maximum possible information* in Q [25], denoted by $\mathsf{T}(Q)$, as

$$\mathsf{T}(Q) \quad := \quad \max_p \ \mathrm{I}(A:B)_{\rho(p)} \ ,$$

where the maximum is taken over all probability distributions p over S.

Note that for a classical-quantum state $\rho_{AB} = \sum_{j=1}^n p_j |j\rangle\langle j| \otimes \rho_j$, the mutual information of A and B is equal to the Holevo information of the quantum ensemble $\mathscr{E} = \big((p_j, \rho_j) : 1 \leq j \leq n\big)$, i.e., $\chi(\mathscr{E}) = \mathrm{I}(A:B)$, and therefore $\mathsf{T}(Q) \geq \chi(\mathscr{E})$.

Most of the entropic quantities defined above arise naturally in the analysis of information processing tasks in the *asymptotic* setting, i.e., when the available resources may be used to jointly complete arbitrarily long sequences of tasks on independent, identically distributed (iid) inputs. The asymptotic setting is an idealization that may not be realistic in certain scenarios. More often, we are faced with single instances of a task which we wish to accomplish with the fewest resources. Recently, researchers have begun to formally study tasks in the non-iid or *one-shot* setting, and the entropic notions that arise therein. Several one-shot entropic concepts have been implicit in traditional (iid) information theory and in communication complexity. For example, Jain, Radhakrishnan, and Sen implicitly studied the concept of *smooth max-relative entropy* in Ref. [28]. However, non-asymptotic concepts were formalized only later (see, e.g., Refs. [43, 44, 16]). In this work, we use one-shot entropic quantities to tightly characterize the communication complexity of remote state preparation.

Let $\rho, \sigma \in \mathsf{D}(\mathcal{H})$ be two quantum states. The *max-relative entropy* of ρ with respect to σ is defined as

$$\mathrm{D}_{\max}(\rho \| \sigma) \quad \coloneqq \quad \min\{\lambda : \rho \le 2^{\lambda}\sigma\} \ ,$$

when $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, and is ∞ otherwise [16]. This notion captures how two states ρ, σ behave relative to each other under the application of a measurement. For a bipartite quantum state $\rho_{AB} \in \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$, the *max-information* part B has about part A [9] is defined as

$$\mathrm{I}_{\max}(A : B)_{\rho} \quad \coloneqq \quad \min_{\sigma \in \mathsf{D}(\mathcal{H})} \mathrm{D}_{\max}(\rho_{AB} \,\|\, \rho_A \otimes \sigma_B) \ .$$

Note that this quantity is asymmetric with respect to the parts A and B. As for mutual information, we include the state as a subscript only when it is not clear from the context. The *smoothed* versions of these quantities come into play when approximations are allowed in the tasks at hand. *Smooth max-relative entropy* is defined as

$$\mathrm{D}^{\epsilon}_{\max}(\rho \| \sigma) \quad \coloneqq \quad \min_{\tilde{\rho} \in \mathsf{B}^{\epsilon}(\rho)} \mathrm{D}_{\max}(\tilde{\rho} \| \sigma) \ ,$$

and *smooth max-information* is defined as

$$\mathrm{I}^{\epsilon}_{\max}(A : B)_{\rho} \quad \coloneqq \quad \min_{\tilde{\rho} \in \mathsf{B}^{\epsilon}(\rho)} \mathrm{I}_{\max}(A : B)_{\tilde{\rho}} \ .$$

There are several ways to define max-information using max-relative entropy [15]. We choose the above definition in this work since it can be used to characterize average-case communication complexity of the remote state preparation problem.

The following are some properties of max-information we use. Both the exact and smooth versions of this quantity are monotonic under the application of a quantum channel [9].

Proposition 2.4 (Monotonicity under quantum channels)**.** *Let $\Phi : \mathrm{L}(\mathcal{H}') \to \mathrm{L}(\mathcal{K})$ be a quantum channel, ρ_{AB} a bipartite sub-normal state over $\mathcal{H}' \otimes \mathcal{H}$, $\sigma_{AB} \in \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ a bipartite quantum state, and $\epsilon \in [0, 1]$. Then*

$$\begin{aligned} \mathrm{I}_{\max}(A' : B)_{\rho'} \quad &\le \quad \mathrm{I}_{\max}(A : B)_{\rho} \ , \qquad \textit{and} \\ \mathrm{I}^{\epsilon}_{\max}(A' : B)_{\sigma'} \quad &\le \quad \mathrm{I}^{\epsilon}_{\max}(A : B)_{\sigma} \ , \end{aligned}$$

where A', B denote two parts of the states $\rho'_{A'B} \coloneqq (\Phi \otimes \mathbb{1})(\rho)$ and $\sigma'_{A'B} \coloneqq (\Phi \otimes \mathbb{1})(\sigma)$.

For a classical-quantum state ρ_{AB}, the value of smooth max-information is achieved by a classical-quantum state ρ'_{AB} that is ϵ-close to ρ_{AB}. A proof is included in Appendix A.

Proposition 2.5. *Let $\rho_{AB} \in \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ be a bipartite quantum state that is classical on A. For any $\epsilon \ge 0$, there exists $\rho'_{AB} \in \mathsf{B}^{\epsilon}(\rho_{AB}) \cap \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ classical on A such that*

$$\mathrm{I}^{\epsilon}_{\max}(A : B)_{\rho} \quad = \quad \mathrm{I}_{\max}(A : B)_{\rho'} \ .$$

Smooth max-information satisfies the Asymptotic Equipartition Property, as proven by Berta, Christandl, and Renner [9]. Let H denote the binary entropy function $\mathrm{H}(\alpha) \coloneqq -\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha)$.

Theorem 2.6 (Quantum Asymptotic Equipartition property)**.** *Let $\epsilon > 0$, n an integer such that $n \geq 2(1-\epsilon^2)$, and $\rho_{AB} \in \mathsf{D}(\mathcal{H}_{AB})$. Then*

$$\mathrm{I}(A:B)_\rho - \frac{3}{n}\,\mathrm{H}(\epsilon) - 2\epsilon \log(|A|\,|B|) \quad \leq \quad \frac{1}{n}\,\mathrm{I}^\epsilon_{\max}(A:B)_{\rho^{\otimes n}} \ , \tag{2.1}$$

and

$$\frac{1}{n}\,\mathrm{I}^\epsilon_{\max}(A:B)_{\rho^{\otimes n}} \quad \leq \quad \mathrm{I}(A:B)_\rho + \frac{\xi(\epsilon)}{\sqrt{n}} - \frac{2}{n}\log\frac{\epsilon^2}{24} \ , \tag{2.2}$$

where $\xi(\epsilon) = 8\sqrt{13 - 4\log\epsilon}\,(2 + \frac{1}{2}\log|A|)$. Therefore,

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} \frac{1}{n}\,\mathrm{I}^\epsilon_{\max}(A:B)_{\rho^{\otimes n}} \quad = \quad \mathrm{I}(A:B)_\rho \ .$$

For $\epsilon \in [0,1)$, the *ϵ-hypothesis testing relative entropy* [48] of two quantum states $\rho, \sigma \in \mathsf{D}(\mathcal{H})$ is defined as

$$\mathrm{D}^\epsilon_{\mathrm{h}}(\rho\|\sigma) \quad := \quad -\log\frac{\beta^\epsilon(\rho\|\sigma)}{1-\epsilon} \ ,$$

where

$$\beta^\epsilon(\rho\|\sigma) \quad := \quad \inf\ \{\langle Q, \sigma\rangle \mid 0 \leq Q \leq \mathbb{1} \text{ and } \langle Q, \rho\rangle \geq 1-\epsilon\} \ . \tag{2.3}$$

The infimum in the above definition is always achieved and $\beta^\epsilon(\rho\|\sigma)$ is between 0 and 1. In this definition, we interpret $(Q, \mathbb{1}-Q)$ as a measurement for distinguishing ρ from σ, i.e., as a strategy in *hypothesis testing*. So $\beta^\epsilon(\rho\|\sigma)$ corresponds to the minimum probability of *incorrectly* identifying σ when ρ is identified *correctly* with probability at least $1-\epsilon$. This one-shot entropic quantity has been studied for a long time either implicitly (see, e.g., Refs. [24, 38]) or explicitly, albeit without giving it a name (see, e.g., Refs. [13, 12]). It also arises in the context of *channel coding* [23, 48] and other tasks [22].

The error in hypothesis testing may only increase under the action of a quantum channel. This has been known for some time; see, e.g., Ref. [10, Eq. (44)] for a proof.

Proposition 2.7 (Data Processing Inequality)**.** *Let $\rho, \sigma \in \mathsf{D}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$, and $\Phi : \mathsf{L}(\mathcal{H}) \to \mathsf{L}(\mathcal{K})$ be a quantum channel. Then*

$$\beta^\epsilon(\rho\|\sigma) \quad \leq \quad \beta^\epsilon(\Phi(\rho)\,\|\,\Phi(\sigma)) \ .$$

The following two properties have been proved implicitly by Matthews and Wehner [34]. For completeness, we include their proofs in Appendix A.

Hypothesis testing error satisfies a restricted form of joint convexity in its two arguments.

Proposition 2.8. *Let $\rho_{AB}(p) \in \mathsf{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a state classical on A such that the distribution on A is given by the probability vector p. Let $\rho_A(p) = \mathrm{Tr}_B(\rho_{AB}(p))$, and $\sigma \in \mathsf{D}(\mathcal{H}_B)$ be a quantum state on Hilbert space $\mathcal{H}_B$. Then the function $\beta^\epsilon(\rho_{AB}(p)\,\|\,\rho_A(p)\otimes\sigma)$ is convex with respect to p.*

Hypothesis testing error is concave in its second argument.

Proposition 2.9. *For any fixed quantum state $\rho \in \mathsf{D}(\mathcal{H})$, the function $\beta^\epsilon(\rho\|\sigma)$ is a concave function with respect to σ.*

It turns out that hypothesis testing relative entropy is closely related to smooth max-relative entropy, as captured by the following theorem.

Theorem 2.10 ([19, 46]). *Let $\rho, \sigma \in \mathsf{D}(\mathcal{H})$ be two quantum states in Hilbert space $\mathcal{H}$. For any $\epsilon \in (0,1)$ and $\delta \in (0,\epsilon)$, the following inequalities hold:*

$$\mathrm{D}_{\max}^{\sqrt{2(1-\epsilon)}}(\rho\|\sigma) \quad \leq \quad \mathrm{D}_{\mathrm{h}}^{\epsilon}(\rho\|\sigma) \ , \qquad \textit{and} \tag{2.4}$$

$$\mathrm{D}_{\max}^{\sqrt{1-\epsilon}}(\rho\|\sigma) \quad \geq \quad \mathrm{D}_{\mathrm{h}}^{\epsilon-\delta}(\rho\|\sigma) - \log\frac{\epsilon(1-\epsilon+\delta)}{\delta^3} - 3\log 3 \ . \tag{2.5}$$

2.5 The minimax theorem

The minimax theorem is a powerful result that provides conditions under which switching the order of minimization and maximization in certain optimization problems does not change the optimum.

Theorem 2.11 ([39]). *Let n be a positive integer, and A_1, A_2 be non-empty, convex and compact subsets of $\mathbb{R}^n$. Let $f : A_1 \times A_2 \to \mathbb{R}$ be a continuous function such that*

1. *$\forall a_2 \in A_2$, the set $\{a_1 \in A_1 : (\forall a_1' \in A_1)\ f(a_1,a_2) \geq f(a_1',a_2)\}$ is convex.*
2. *$\forall a_1 \in A_1$, the set $\{a_2 \in A_2 : (\forall a_2' \in A_2)\ f(a_1,a_2) \leq f(a_1,a_2')\}$ is convex.*

Then

$$\max_{a_1\in A_1}\min_{a_2\in A_2} f(a_1,a_2) \quad = \quad \min_{a_2\in A_2}\max_{a_1\in A_1} f(a_1,a_2) \ .$$

2.6 Remote state preparation

Let S be a finite, non-empty set, and let $Q : S \to \mathsf{D}(\mathcal{H})$ be a function that maps each element $x \in S$ to a quantum state $Q(x)$ over the Hilbert space $\mathcal{H}$. Recall that remote state preparation, denoted as $\mathrm{RSP}(S,Q)$, is a communication task in which one party, Alice, is given an input $x \in S$, and engages in an LOCC protocol with another party, Bob, so that Bob is able to prepare $Q(x)$. The function Q is known to both parties. In the approximate remote state preparation, we allow Bob to prepare an approximation $\sigma_x \in \mathsf{D}(\mathcal{H})$ to $Q(x)$. We consider two notions of error in approximation: worst case and average case. Let $\epsilon \in [0,1]$, and let p be a probability distribution on S. We say a protocol for $\mathrm{RSP}(S,Q)$ makes worst-case error ϵ if $\mathrm{P}(\sigma_x, Q(x)) \leq \epsilon$ for each $x \in S$. We say a protocol for $\mathrm{RSP}(S,Q)$ makes average-case error ϵ w.r.t. the distribution p over S if the purified distance between the ideal and actual joint input-output states is at most ϵ. By the definitions of purified distance and fidelity, this condition is equivalent to

$$\sum_{x\in S} p_x\,\mathrm{F}(\sigma_x, Q(x)) \quad \geq \quad \sqrt{1-\epsilon^2} \ .$$

In Sections 3 and 4, we characterize the communication complexity of this problem for the two different kinds of approximation. We emphasize that Alice and Bob communicate with a noiseless classical channel, they have access to an arbitrarily large amount of entanglement of their choice, and they have unlimited computational power.

A straightforward protocol for approximate remote state preparation is as follows. Alice sends her input x directly to Bob and Bob creates the desired state $Q(x)$. Thus Bob prepares the target state with zero error ($\epsilon = 0$) using $\lceil \log(n+1) \rceil$ bits of classical communication, where $n = |S|$.

Jain, Radhakrishnan, and Sen [28, 29] proposed the following, potentially more efficient protocol, which we call the JRS protocol in the sequel. Let $\mathcal{K}$ be a Hilbert space with $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$ and $\{\sigma_x\}_{x\in S} \subseteq \mathsf{D}(\mathcal{H})$

be a set of quantum states such that for all $x \in S$, $\mathrm{P}(\sigma_x, Q(x)) \leq \delta$ for some $\delta \in [0,1]$. Suppose that for some $\lambda \in [0,\infty)$ and some $\sigma \in \mathsf{D}(\mathcal{H})$, we have

$$\sigma_x \quad \leq \quad 2^{\lambda}\sigma \qquad \text{for all} \quad x \in S \ . \tag{2.6}$$

This can be rewritten for a fixed $x \in S$ as

$$\sigma \quad = \quad 2^{-\lambda}\sigma_x + (1-2^{-\lambda})\xi_x \ ,$$

where $\xi_x \in \mathsf{D}(\mathcal{H})$ is a quantum state. Let $|v_x\rangle \in \mathcal{K} \otimes \mathcal{H}$ be a purification of σ_x in the Hilbert space $\mathcal{K} \otimes \mathcal{H}$, and $|u_x\rangle \in \mathcal{K} \otimes \mathcal{H}$ be a purification of ξ_x. Then

$$|w_x\rangle \quad = \quad \sqrt{2^{-\lambda}}\,|0\rangle|v_x\rangle \quad + \quad \sqrt{1-2^{-\lambda}}\,|1\rangle|u_x\rangle \ ,$$

is a purification of σ. Let $|w\rangle$ be an arbitrary but fixed purification of σ in $\mathbb{C}^2 \otimes \mathcal{K} \otimes \mathcal{H}$. By the unitary equivalence of purifications, there is a unitary operation U_x on the space $\mathbb{C}^2 \otimes \mathcal{K}$ which transforms $|w\rangle$ to $|w_x\rangle$. We are ready to describe the JRS protocol.

JRS Protocol: Alice and Bob agree on a parameter t, that depends on the quality of approximation they desire. Initially, Alice and Bob share t copies of the quantum state $|w\rangle$. The registers corresponding to Hilbert spaces $\mathbb{C}^2$ and $\mathcal{K}$ in the ith copy of $|w\rangle$ are called C_i and K_i, respectively, and are held by Alice. The register corresponding to the Hilbert space $\mathcal{H}$ is called H_i and is held by Bob.

1. On getting input x, Alice performs the unitary operation U_x on registers C_iK_i for each $i \in [t]$. This transforms all copies of $|w\rangle$ to copies of $|w_x\rangle$. Then she measures the register C_i for all $i \in [t]$. If at least one of the measurement outcomes, say the jth, is equal to zero, she sends the index j to Bob, using $\lceil \log(t+1) \rceil$ bits. (She may choose to send any such index.) Otherwise, if the outcomes of all t measurements are equal to one, she sends 0 to Bob.
2. On receiving an integer k, where $0 \leq k \leq t$, Bob outputs the state in register H_k if $k \in [t]$, and outputs the maximally mixed state over $\mathcal{H}$ if $k = 0$.

The output of this protocol is $\frac{\mathbb{1}}{|H|}$ with probability $\left(1-2^{-\lambda}\right)^t$ and σ_x with the remaining probability. Hence, the output state is

$$\tilde{\sigma}_x \quad = \quad \left(1-\left(1-2^{-\lambda}\right)^t\right)\,\sigma_x + \left(1-2^{-\lambda}\right)^t\,\frac{\mathbb{1}}{|H|} \ .$$

By choosing the approximation parameter δ small enough and t large enough, Bob produces a state $\tilde{\sigma}_x$ with the desired accuracy. We use this protocol to give upper bounds on the worst-case error and average-case communication complexity of $\mathrm{RSP}(S,Q)$.

3 Average-case communication complexity

Let p be a probability distribution over S and $\mathsf{Q}^*_p(\mathrm{RSP}(S,Q),\epsilon)$ denote the average-case entanglement-assisted communication complexity of approximate remote state preparation (ARSP), with respect to p, and with (average) error at most ϵ. We characterize this quantity in terms of smooth max-information, a one-shot analogue of mutual information.

3.1 An upper bound

First, we show that the average-case communication complexity with error ϵ of ARSP is bounded above essentially by $\mathrm{I}^{\delta}_{\max}(A:B)_{\rho(p)}$, where $\rho(p)$ is the ideal joint state of Alice's input and Bob's output, and $\delta \in \Theta(\epsilon)$. To do so, we use the JRS protocol described in Section 2.6.

Theorem 3.1. *For any finite set S, function $Q : S \to \mathsf{D}(\mathcal{H})$, and $\epsilon \in (0,1]$, let p be a probability distribution over S and $\rho_{AB}(p) \in \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ be the bipartite classical-quantum state $\rho_{AB}(p) = \sum_{x\in S} p_x |x\rangle\langle x|_A \otimes Q(x)_B$. Then*

$$\mathsf{Q}^*_p(\mathrm{RSP}(S,Q),\epsilon) \quad \leq \quad \mathrm{I}^{\delta}_{\max}(A:B)_{\rho(p)} + \log_2 \ln \frac{8}{\epsilon^2} + 2 \ ,$$

where $\delta = \epsilon/2\sqrt{2}$.

Proof: Fix some $\epsilon \in (0,1]$, and let λ be equal to $\mathrm{I}^{\delta}_{\max}(A : B)_{\rho(p)}$ with δ as in the statement of the theorem. By Proposition 2.5, there exist quantum states $\rho'_{AB} \in \mathsf{B}^{\delta}(\rho_{AB})$ and $\sigma_B \in \mathsf{D}(\mathcal{H})$ such that $\rho'_{AB} \leq 2^{\lambda}\rho'_A \otimes \sigma_B$,where $\rho'_{AB} = \sum_x q_x |x\rangle\langle x| \otimes \sigma^x_B$ with $\sum_x q_x = 1$ and $\sigma^x_B \in \mathsf{D}(\mathcal{H})$, and $\rho'_A = \sum_x q_x |x\rangle\langle x|$. Then

$$\sigma^x_B \quad \leq \quad 2^{\lambda}\sigma_B \ , \tag{3.1}$$

for all $x \in S$ with $q_x \neq 0$. For each $x \in S$ with $q_x = 0$, we assume, w.l.o.g., that $\sigma^x_B = \sigma_B$. Inequality (3.1) is in the form of inequality (2.6) and therefore we may execute the JRS protocol with a suitable choice of parameter t. Initially, Alice and Bob share t copies of entangled state $|w\rangle$, where $|w\rangle$ is a purification of σ_B. Alice gets input x with probability p_x. They perform the protocol for approximating state σ^x_B from σ_B. The final joint state of Alice's input and Bob's output is

$$\tilde{\rho}_{AB} \quad = \quad \sum_{x\in S} p_x |x\rangle\langle x| \otimes \tilde{\sigma}^x_B \ ,$$

where

$$\tilde{\sigma}^x_B \quad = \quad \left(1 - \left(1 - 2^{-\lambda}\right)^t\right) \sigma^x_B + \left(1 - 2^{-\lambda}\right)^t \frac{\mathbb{1}}{\dim(\mathcal{H})} \ .$$

Therefore,

$$\begin{aligned}
\mathrm{F}(\tilde{\rho}_{AB}, \rho'_{AB}) \quad &= \quad \mathrm{F}\left(\sum_{x\in S} p_x |x\rangle\langle x| \otimes \tilde{\sigma}^x_B,\ \sum_{x\in S} q_x |x\rangle\langle x| \otimes \sigma^x_B\right) \\
&\geq \quad \left(1 - (1 - 2^{-\lambda})^t\right) \mathrm{F}\left(\sum_{x\in S} p_x |x\rangle\langle x| \otimes \sigma^x_B,\ \sum_{x\in S} q_x |x\rangle\langle x| \otimes \sigma^x_B\right) \\
&= \quad \left(1 - (1 - 2^{-\lambda})^t\right) \sum_{x\in S} \sqrt{p_x q_x} \\
&\geq \quad \left(1 - (1 - 2^{-\lambda})^t\right) \sqrt{1 - \delta^2} \ ,
\end{aligned}$$

where the first inequality follows from the joint concavity of fidelity. The last inequality follows from monotonicity under quantum channels:

$$\sum_{x\in S} \sqrt{p_x q_x} \quad = \quad \mathrm{F}(\rho'_A, \rho_A) \quad \geq \quad \mathrm{F}(\rho'_{AB}, \rho_{AB}) \ .$$

In addition, by Proposition 2.2,

$$\begin{aligned}
\mathrm{F}(\tilde{\rho}_{AB}, \rho_{AB}) \quad &\geq \quad \mathrm{F}(\tilde{\rho}_{AB}, \rho'_{AB})^2 + \mathrm{F}(\rho_{AB}, \rho'_{AB})^2 - 1 \\
&\geq \quad \left(1 - (1 - 2^{-\lambda})^t\right)^2 (1 - \delta^2) + (1 - \delta^2) - 1 \\
&\geq \quad \sqrt{1 - \epsilon^2} \ ,
\end{aligned}$$

where the last inequality is derived using inequalities $\ln(1-x) \leq -x$ and $\sqrt{1-x} \leq 1 - \frac{x}{2}$, which hold for $x \in [0,1)$, and the parameter values $\delta = \epsilon/2\sqrt{2}$ and $t = \lceil 2^{\lambda} \ln \frac{8}{\epsilon^2} \rceil$. Since $\mathrm{F}(\tilde{\rho}_{AB}, \rho_{AB}) = \sum_{x\in S} p_x \, \mathrm{F}(\tilde{\sigma}_x, Q(x))$, the protocol has average-case error at most ϵ.

The communication cost of this protocol is $\lceil \log(t+1) \rceil$. So the communication complexity of approximate remote state preparation with average-case error ϵ is

$$\mathsf{Q}^*_p(\mathrm{RSP}(S,Q),\epsilon) \quad \leq \quad \lceil \log(t+1) \rceil \quad \leq \quad \lambda + \log_2 \ln \frac{8}{\epsilon^2} + 2 \ ,$$

as required. ■

We have not attempted to optimize the upper bound derived above. It is possible that the parameter δ and the ϵ-dependent additive term be improved further.

3.2 A lower bound

Next, we show that the average-case communication complexity of any protocol for approximate remote state preparation is bounded from below by $\mathrm{I}^\epsilon_{\max}(A:B)_{\rho(p)}$. In order to do this, we strengthen a property of smooth max-information due to Berta, Christandl, and Renner [9, Lemma B.12], in the case of a tripartite state ρ_{MAB} that is classical on M.

Lemma 3.2. *Let $\epsilon \geq 0$ and $\rho_{MAB} \in \mathsf{D}(\mathcal{M} \otimes \mathcal{H}' \otimes \mathcal{H})$ be any tripartite quantum state over registers M, A and B such that ρ is classical on M. Then*

$$\mathrm{I}^\epsilon_{\max}(A:MB) \quad \leq \quad \mathrm{I}^\epsilon_{\max}(A:B) + \log|M| \ .$$

Proof: Fix $\sigma_B \in \mathsf{D}(\mathcal{H})$ and $\tilde{\rho}_{AB} \in \mathsf{B}^\epsilon(\rho_{AB})$ such that $\mathrm{I}^\epsilon_{\max}(A:B) = \mathrm{D}_{\max}(\tilde{\rho}_{AB} \,\|\, \tilde{\rho}_A \otimes \sigma_B)$. Let λ denote this max-relative entropy, i.e., λ is the minimum non-negative real number for which $\tilde{\rho}_{AB} \leq 2^\lambda \tilde{\rho}_A \otimes \sigma_B$. Then

$$\frac{\mathbb{1}}{|M|} \otimes \tilde{\rho}_{AB} \quad \leq \quad 2^\lambda \, \frac{\mathbb{1}}{|M|} \otimes \tilde{\rho}_A \otimes \sigma_B \ . \tag{3.2}$$

By Proposition 2.3, there exists some extension ρ'_{MAB} of $\tilde{\rho}_{AB}$ such that $\rho'_{MAB} \in \mathsf{B}^\epsilon(\rho_{MAB})$. By construction, we have $\mathrm{Tr}_M(\rho'_{MAB}) = \tilde{\rho}_{AB}$. Consider the quantum-to-classical channel $\Phi : \mathsf{L}(\mathcal{M}) \to \mathsf{L}(\mathcal{M})$ defined by

$$\Phi(X) \quad = \quad \sum_i \langle e_i|X|e_i\rangle |e_i\rangle\langle e_i|$$

for all $X \in \mathsf{L}(\mathcal{M})$, where $\{|e_i\rangle\}$ is the standard orthonormal basis for the Hilbert space $\mathcal{M}$. The state $(\Phi \otimes \mathbb{1})(\rho'_{MAB})$ is classical on M, and is an extension of $\tilde{\rho}_{AB}$. Define $\tilde{\rho}_{MAB} \coloneqq (\Phi \otimes \mathbb{1})(\rho'_{MAB})$. Since $\rho'_{MAB} \in \mathsf{B}^\epsilon(\rho_{MAB})$, by monotonicity of fidelity under quantum channels and because ρ_{MAB} is classical on M, we have $\tilde{\rho}_{MAB} \in \mathsf{B}^\epsilon(\rho_{MAB})$. So $\tilde{\rho}_{MAB}$ may be written as

$$\tilde{\rho}_{MAB} \quad = \quad \sum_i \gamma_i |e_i\rangle\langle e_i| \otimes \sigma^i_{AB} \ ,$$

where all σ^i_{AB} are normalized and $\sum_i \gamma_i \leq 1$. We have $\tilde{\rho}_{MAB} \leq \mathbb{1}_M \otimes \tilde{\rho}_{AB}$. Combining this with Equation (3.2), we can conclude that

$$\tilde{\rho}_{MAB} \quad \leq \quad 2^\lambda |M| \left(\frac{\mathbb{1}_M}{|M|} \otimes \tilde{\rho}_A \otimes \sigma_B \right)$$

and consequently,

$$\mathrm{D}_{\max}\left(\tilde{\rho}_{MAB} \,\middle\|\, \frac{\mathbb{1}_M}{|M|} \otimes \tilde{\rho}_A \otimes \sigma_B \right) \quad \leq \quad \lambda + \log|M| \ .$$

By the definition of smooth max-information, this implies that

$$\mathrm{I}^\epsilon_{\max}(A:MB) \quad \leq \quad \lambda + \log|M| \ ,$$

as required. ■

Remark: The above lemma could alternatively be derived from an analogous inequality for *α-Rényi mutual information* [31, Equation (2.25)]). Taking the limit as $\alpha \to \infty$ gives us the inequality for max-information (i.e., for $\epsilon = 0$). We may extend this to any $\epsilon \geq 0$ by smoothing arguments similar to those in the above proof.

Using this lemma, we bound the average-case communication complexity of $\mathrm{RSP}(S, Q)$ from below.

Theorem 3.3. *For any finite set S, function $Q : S \to \mathsf{D}(\mathcal{H})$, and probability distribution p over S, let $\rho(p)$ be the bipartite quantum state*

$$\rho(p) \quad = \quad \sum_{x \in S} p_x |x\rangle\langle x|_A \otimes Q(x)_B \enspace .$$

For any $\epsilon \in [0, 1]$, we have

$$\mathsf{Q}^*_p(\mathrm{RSP}(S, Q), \epsilon) \quad \geq \quad \mathrm{I}^\epsilon_{\max}(A : B)_{\rho(p)} \enspace .$$

Proof: In this proof we follow the notation and convention described in Section 2.2. Consider a k-round LOCC protocol Π for $\mathrm{RSP}(S, Q)$ with average-case error ϵ. Suppose Bob sends the first message, and the joint state in Alice and Bob's registers (excluding the input register A) after the message is ϕ. As Bob receives no input, the joint state ϕ is known to both parties. Hence, the rest of the protocol can be considered as a *new* LOCC protocol, with the same output, in which the initial shared state of parties is ϕ, *and* Alice starts the protocol. The communication cost of this new protocol is less than the communication cost of the original one. Therefore, it suffices to show the lower bound for protocols in which Alice starts.

Let A be Alice's input register, and $Y_i \coloneqq P_i V_i$ and $Z_i \coloneqq Q_i W_i$ be Alice's and Bob's classical-quantum private registers, respectively, after the ith round of the protocol for $i \geq 0$. Initially, A and Z_0 are independent, and so

$$\mathrm{I}_{\max}(A : Z_0) \quad = \quad 0 \enspace . \tag{3.3}$$

Consider the ith round of a two-way LOCC protocol. The communication in each round is either from Alice to Bob (for odd i) or from Bob to Alice (for even i).

Odd round i: In this case, Alice measures her private qubits V_{i-1} controlled by P_{i-1} and A. She includes the outcome of her measurement M_i in the register P_i (recall that $P_i = P_{i-1} M_i$), and sends a copy of M_i to Bob using $m_i \coloneqq \lceil \log(|M_i| + 1) \rceil$ bits of communication. Then Bob includes the received message M_i in Q_i (recall that $Q_i = Q_{i-1} M_i$). Thus,

$$\begin{aligned} \mathrm{I}_{\max}(A : Z_i) \quad &\leq \quad \mathrm{I}_{\max}(A : Z_{i-1}) + \log |M_i| && \text{(by Lemma 3.2)} \\ &\leq \quad \mathrm{I}_{\max}(A : Z_{i-1}) + m_i \enspace . && (3.4) \end{aligned}$$

Even round i: In this case, Bob measures his private qubits W_{i-1} controlled by Q_{i-1}. He includes the outcome of his measurement M_i in Q_i, and sends a copy of M_i to Alice using $m_i = \lceil \log(|M_i| + 1) \rceil$ bits of communication. Alice includes the received message in P_i. Thus,

$$\mathrm{I}_{\max}(A : Z_i) \quad \leq \quad \mathrm{I}_{\max}(A : Z_{i-1}) \enspace . \qquad \text{(by Proposition 2.4)} \tag{3.5}$$

Combining Eqs. (3.4) and (3.5) recursively, we get

$$\mathrm{I}_{\max}(A : Z_k) \quad \leq \quad \mathrm{I}_{\max}(A : Z_0) + \sum_{\substack{1 \leq i \leq k \\ i \text{ odd}}} m_i \quad = \quad \sum_{\substack{1 \leq i \leq k \\ i \text{ odd}}} m_i \enspace ,$$

after k rounds of communication. Let $m \coloneqq \sum_{1 \le i \le k,\ \text{odd}} m_i$. At the end of the protocol, Bob applies a quantum channel on his register Z_k to get the output B. By monotonicity of max-information (Proposition 2.4), we have

$$\mathrm{I}_{\max}(A : B)_{\rho'(p)} \quad \le \quad m \ ,$$

where $\rho'(p) = \sum_x p_x |x\rangle\langle x| \otimes \sigma_x$ is the bipartite quantum state of registers AB, and m is the number of bits of communication from Alice to Bob. In addition, protocol Π guarantees that $\rho'(p)$ is within purified distance ϵ of $\rho(p)$. Therefore, we conclude the theorem. ■

4 Worst-case communication complexity

In this section, we characterize the worst-case communication complexity of remote state preparation, denoted as $\mathsf{Q}^*(\mathrm{RSP}(S, Q), \epsilon)$, in terms of smooth max-relative entropy.

4.1 An upper bound

We show that for some fixed $\epsilon \in (0, 1]$, the worst-case communication complexity of the approximate remote state preparation problem is bounded from above essentially by

$$\min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{x \in S} \mathrm{D}^{\delta}_{\max}(Q(x) \| \sigma) \ ,$$

where $\delta \in \Theta(\epsilon)$. As for the average case, we utilize the JRS protocol presented in Section 2.6.

Theorem 4.1. *Let S be a non-empty finite set, $Q : S \to \mathsf{D}(\mathcal{H})$ be a function from S to the set of density operators in the Hilbert space $\mathcal{H}$, and $\epsilon \in [0, 1]$. Then*

$$\mathsf{Q}^*(\mathrm{RSP}(S, Q), \epsilon) \quad \le \quad \min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{x \in S} \mathrm{D}^{\delta}_{\max}(Q(x) \| \sigma) + \log_2(1 + \epsilon^2) + \log_2 \ln \frac{2}{\epsilon^4} + 2 \ ,$$

where $\delta = \frac{\epsilon}{\sqrt{1+\epsilon^2}}$.

Proof: Let $\alpha \coloneqq \min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{x \in S} \mathrm{D}^{\delta}_{\max}(Q(x) \| \sigma)$ and σ' be the quantum state for which the minimum is achieved, i.e., $\alpha = \max_{x \in S} \mathrm{D}^{\delta}_{\max}(Q(x) \| \sigma')$. By definition, for all $x \in S$ there exists some $\sigma_x \in \mathsf{B}^{\delta}(Q(x))$ such that

$$\sigma' \quad \ge \quad 2^{-\alpha} \sigma_x \ .$$

Since $\mathrm{P}(\sigma_x, Q(x)) \le \delta$, we have $\mathrm{F}(\sigma_x, Q(x))^2 \ge 1 - \delta^2$. So, by Proposition 2.1, $\mathrm{Tr}(\sigma_x) \ge 1 - \delta^2 = \frac{1}{1+\epsilon^2}$ for all $x \in S$. For each $x \in S$, define $\rho_x \coloneqq \frac{\sigma_x}{\mathrm{Tr}(\sigma_x)}$. Then for all $x \in S$, ρ_x is a quantum state δ-close to $Q(x)$, i.e., $\rho_x \in \mathsf{B}^{\delta}(Q(x)) \cap \mathsf{D}(\mathcal{H})$, and

$$\sigma' \quad \ge \quad 2^{-\alpha}\ \mathrm{Tr}(\sigma_x)\ \rho_x \quad \ge \quad \frac{2^{-\alpha}}{1 + \epsilon^2}\ \rho_x \ .$$

This inequality is precisely in the form of inequality (2.6). Now we run the JRS protocol to approximate $Q(x)$, with $t = 2^{\alpha}(1 + \epsilon^2) \ln \frac{2}{\epsilon^4}$. At the end of this protocol, Bob's output is

$$\tilde{\sigma}_x \quad \coloneqq \quad \left(1 - \left(1 - 2^{-\kappa}\right)^t\right) \sigma_x + (1 - 2^{-\kappa})^t \frac{\mathbb{1}}{\dim(\mathcal{H})} \ ,$$

where $\kappa = \alpha + \log(1 + \epsilon^2)$.

By joint concavity of fidelity, and because σ_x is $\frac{\epsilon}{\sqrt{1+\epsilon^2}}$-close to $Q(x)$, we have

$$\mathrm{F}(Q(x), \tilde{\sigma}_x) \quad \geq \quad \left(1 - (1-2^{-\kappa})^t\right) \mathrm{F}(Q(x), \sigma_x) \quad \geq \quad \frac{1-(1-2^{-\kappa})^t}{\sqrt{1+\epsilon^2}} \quad \geq \quad \sqrt{1-\epsilon^2} \ .$$

Here we appealed to the inequalities $\ln(1-x) \leq -x$ and $\sqrt{1-x} \leq 1 - \frac{x}{2}$ (for $x \in [0,1)$), and the definition of κ and t. Thus, the purified distance of $Q(x)$ and $\tilde{\sigma}_x$ is at most ϵ, and the protocol performs remote state preparation with worst-case error ϵ. The communication cost of this protocol is $\lceil \log(t+1) \rceil$. Hence, we have

$$\mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \quad \leq \quad \lceil \log(t+1) \rceil \quad \leq \quad \alpha + \log_2(1+\epsilon^2) + \log_2 \ln \frac{2}{\epsilon^4} + 2 \ ,$$

the stated upper bound. ■

4.2 A lower bound

By definition, any protocol with worst-case error at most ϵ is also a protocol with average-case error at most ϵ. As a consequence, any lower bound for average-case communication complexity is also a lower bound for worst-case communication complexity. In particular, by Theorem 3.3, for each probability distribution p, $\mathrm{I}^{\epsilon}_{\max}(A:B)_{\rho(p)}$ is a lower bound for the worst-case communication complexity of remote state preparation. Therefore,

$$\max_p \ \mathrm{I}^{\epsilon}_{\max}(A:B)_{\rho(p)} \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \ , \tag{4.1}$$

where the maximum is over all probability distributions p on the set S. In the following theorem, we give a lower bound for $\mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon)$ in terms of max-relative entropy using Equation (4.1).

Theorem 4.2. *Let S be a non-empty finite set, $Q : S \to \mathsf{D}(\mathcal{H})$ be a function from S to the set of density operators in Hilbert space $\mathcal{H}$, $\epsilon \in (0,1]$, and $\delta \in (0, 1-\epsilon^2)$. Then*

$$\min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{x \in S} \ \mathrm{D}^{\gamma}_{\max}(Q(x) \| \sigma) - \log \frac{(1-\epsilon^2)(\epsilon^2+\delta)}{\delta^3} - 3\log 3 \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \ ,$$

where $\gamma = \sqrt{2(\epsilon^2+\delta)}$.

Proof: By definition of the smooth max-information, Eq. (4.1) implies that

$$\max_p \min_{\sigma \in \mathsf{D}(\mathcal{H})} \ \mathrm{D}^{\epsilon}_{\max}(\rho_{AB}(p) \ \| \ \rho_A(p) \otimes \sigma) \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \ , \tag{4.2}$$

whereas the upper bound shown in Theorem 4.1 is

$$\min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{x \in S} \ \mathrm{D}^{\delta}_{\max}(Q(x) \| \sigma) \ .$$

If the minimax theorem held for the above expression, the theorem would follow. However, smooth max-relative entropy $\mathrm{D}^{\epsilon}_{\max}$ is neither convex nor concave in its arguments, and the minimax theorem does not apply directly. Instead, we appeal to Theorem 2.10, and approximate it with hypothesis testing relative entropy $\mathrm{D}^{\epsilon}_{\mathrm{h}}$, and write it in terms of the hypothesis testing error β^{ϵ}. This measure satisfies the hypotheses of the minimax theorem (cf. Proposition 2.9 and 2.8). We then apply the minimax theorem, and finally return to $\mathrm{D}^{\epsilon}_{\max}$ to derive the lower bound.

By Theorem 2.10, we have

$$\begin{aligned}
\max_{p} \min_{\sigma \in \mathsf{D}(\mathcal{H})} \; \mathrm{D}^{\epsilon}_{\max}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) \quad &\geq \quad \max_{p} \min_{\sigma \in \mathsf{D}(\mathcal{H})} \; \mathrm{D}^{\lambda}_{\mathrm{h}}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) - f(\epsilon, \delta) \\
&= \quad \max_{p} \min_{\sigma \in \mathsf{D}(\mathcal{H})} \; \left(-\log \; \beta^{\lambda}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) \right) \\
&\qquad + \log(1-\lambda) - f(\epsilon, \delta) \\
&= \quad -\log \left(\min_{p} \max_{\sigma \in \mathsf{D}(\mathcal{H})} \; \beta^{\lambda}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) \right) \\
&\qquad + \log(1-\lambda) - f(\epsilon, \delta) \;\; ,
\end{aligned}$$

where $f(\epsilon, \delta) = \log \frac{(1-\epsilon^2)(\epsilon^2+\delta)}{\delta^3} + 3 \log 3$ and $\lambda = 1 - \epsilon^2 - \delta$.

Let A_1 be the set of all probability distributions p over S, and A_2 be the set of all quantum states $\sigma \in \mathsf{D}(\mathcal{H})$. Viewing σ as an element of the real vector space of Hermitian operators in $\mathsf{L}(\mathcal{H})$, A_1 and A_2 are non-empty, convex and compact subsets of $\mathbb{R}^n$ for some positive integer n. The quantity $\beta^{\lambda}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma)$ is a continuous function of its arguments. Moreover, by Proposition 2.8 and Proposition 2.9, it satisfies both conditions of the minimax theorem, Theorem 2.11. Thus, we conclude that

$$\begin{aligned}
\max_{p} \min_{\sigma \in \mathsf{D}(\mathcal{H})} \; \mathrm{D}^{\epsilon}_{\max}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) \quad &\geq \quad -\log \left(\max_{\sigma \in \mathsf{D}(\mathcal{H})} \min_{p} \; \beta^{\lambda}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) \right) \\
&\qquad + \log(1-\lambda) - f(\epsilon, \delta) \\
&= \quad \min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{p} \; \mathrm{D}^{\lambda}_{\mathrm{h}}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) - f(\epsilon, \delta) \\
&\geq \quad \min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{p} \; \mathrm{D}^{\gamma}_{\max}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma) - f(\epsilon, \delta) \\
&\geq \quad \min_{\sigma \in \mathsf{D}(\mathcal{H})} \max_{x \in S} \; \mathrm{D}^{\gamma}_{\max}(Q(x) \| \sigma) - f(\epsilon, \delta) \;\; , \qquad (4.3)
\end{aligned}$$

where $\gamma = \sqrt{2(1-\lambda)} = \sqrt{2(\epsilon^2 + \delta)}$. In the second inequality above, we use Theorem 2.10 to move between hypothesis testing relative entropy and max-relative entropy. Combining Eqs. (4.3) and (4.2), we get the lower bound for the worst-case communication complexity of ARSP. ■

5 Some observations

In earlier sections, we characterized the communication complexity of the approximate remote state preparation problem (ARSP) for both worst-case error and average-case error. We now discuss the results, especially in light of previous work.

5.1 A comparison with previous works

In Section 4, we derived bounds on the worst-case communication complexity of ARSP. Jain [25] showed that the worst-case communication complexity of ARSP of a sequence of quantum states $(Q(x) : x \in S)$ is bounded from above in terms of the "maximum possible information" $\mathsf{T}(Q)$ as:

$$\frac{8(4\mathsf{T}(Q) + 7)}{\left(1 - \sqrt{1-\epsilon^2}\right)^2} \;\; , \qquad (5.1)$$

where ϵ is the approximation error. (See Section 2.4 for a definition of $\mathsf{T}(Q)$.)

We observe that for certain sets of states there is a large separation between the bound established in Theorem 4.1, and Equation (5.1). Specifically, the upper bound in Theorem 4.1 may be asymptotically smaller than the bound in Equation (5.1).

The separation follows from a combination of two pieces of work. The first is an information-theoretic result, the *Substate theorem* due to Jain, Radhakrishnan, and Sen [30], which relates the smooth max-relative entropy of two states to their observational divergence. The precise form of the statement below is due to Jain and Nayak [26].

Theorem 5.1 (Substate theorem [30, 26]). *Let $\mathcal{H}$ be a Hilbert space, and let $\rho, \sigma \in \mathsf{D}(\mathcal{H})$ be quantum states such that* $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$. *For any $\epsilon \in (0,1)$,*

$$\mathrm{D}^{\epsilon}_{\max}(\rho\|\sigma) \quad \leq \quad \frac{\mathrm{D}_{\mathrm{obs}}(\rho\|\sigma)}{\epsilon^2} + \log\frac{1}{1-\epsilon^2} \ .$$

The second result is due to Jain, Nayak, and Su [27], who constructed an ensemble of quantum states for which there is a large separation between its Holevo and Divergence information. (See Section 2.4 for a definition of these two information quantities.)

Theorem 5.2. *Let n be a positive integer, and $\mathcal{H}$ be a Hilbert space of dimension n. For every positive real number $k \geq 1$ such that $\log_2 n > 36k^2$, there is a finite set S and an ensemble $\mathscr{E} = \{(\lambda_x, \xi_x) : x \in S\}$ of quantum states $\xi_x \in \mathsf{D}(\mathcal{H})$ with $\xi \coloneqq \sum_{x\in S} \lambda_x \xi_x = \frac{\mathbb{1}}{n}$, such that $\mathrm{D}_{\mathrm{obs}}(Q(x)\|\xi) = \mathrm{D}_{\mathrm{obs}}(\mathscr{E}) = k$ for all $x \in S$ and $\chi(\mathscr{E}) \in \Theta(k \log\log n)$.*

Jain *et al.* [27] also showed that this is the best separation possible for an ensemble of quantum states with a completely mixed ensemble average.

Putting these together, we get:

Theorem 5.3. *Let $\delta \in (0,1]$ and $\mathcal{H}$ be Hilbert space with dimension n. Then, for every positive real number $k \geq 1$ such that $\log_2 n > 36k^2$, there is a finite set S and a function $Q : S \to \mathsf{D}(\mathcal{H})$ such that $\mathsf{T}(Q) \in \Omega(k \log\log n)$ while*

$$\min_{\sigma\in\mathsf{D}(\mathcal{H})} \max_{x\in S} \mathrm{D}^{\delta}_{\max}(Q(x)\|\sigma) \quad \leq \quad \frac{k}{\delta^2} + \log\frac{1}{1-\delta^2} \ .$$

Proof: Let S be the set S and $\mathscr{E} = \{(\lambda_x, \xi_x) : x \in S\}$ the ensemble given by Theorem 5.2. Let $Q : S \to \mathsf{D}(\mathcal{H})$ be the function such that $Q(x) = \xi_x$ for all $x \in S$. Suppose that $\xi \coloneqq \sum_{x\in S} \lambda_x \xi_x$ is the ensemble average. Then we have

$$\begin{aligned} \min_{\sigma\in\,\mathsf{D}(\mathcal{H})} \max_{x\in S} \mathrm{D}^{\delta}_{\max}(Q(x)\|\sigma) \quad &\leq \quad \max_{x\in S} \mathrm{D}^{\delta}_{\max}(Q(x)\|\xi) \\ &\leq \quad \frac{\max_x \mathrm{D}_{\mathrm{obs}}(Q(x)\|\xi)}{\delta^2} + \log\frac{1}{1-\delta^2} \\ &= \quad \frac{k}{\delta^2} + \log\frac{1}{1-\delta^2} \ , \end{aligned}$$

where the second inequality is derived using the Substate theorem (Theorem 5.1). Moreover, by definition of the maximum possible information $\mathsf{T}(Q)$, we have $\mathsf{T}(Q) \geq \chi(\mathscr{E})$. This gives us the existence of the required function Q. ■

Jain [25] also gave a lower bound of $\mathsf{T}(Q)/2$ for exact remote state preparation. The above observation also implies that allowing remote state preparation with non-zero error in approximating the state may decrease the communication cost asymptotically. By Theorem 5.3, we get a function Q for which the worst-case complexity with zero error $\mathsf{Q}^*(\mathrm{RSP}(S,Q), 0) \in \Omega(k \log\log n)$, while for any $\epsilon \in (0,1]$, the complexity with error ϵ is

$$\mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \quad \leq \quad \frac{k}{\delta^2} + \log\frac{1}{1-\delta^2} \ ,$$

where $\delta \coloneqq \frac{\epsilon}{2\sqrt{1+\epsilon^2}}$.

5.2 Average-case error vs. worst-case error

Requiring bounded worst-case error in approximating states in remote state preparation is more demanding, and potentially requires more communication, as compared to the average case. Here we quantify how much more expensive it could be.

For the rest of this subsection, we let n be a positive integer, fix $S = \{1, 2, \ldots, 2^n\}$, $\mathcal{H} = \mathrm{span}\,\{|x\rangle : x \in S\}$, and define $Q : S \to \mathsf{D}(\mathcal{H})$ by $Q(x) = |x\rangle\langle x|$ for all $x \in S$.

Proposition 5.4. *For every $\epsilon \in [0, 1/\sqrt{2}\,)$, there is a probability distribution p_ϵ over the set S such that $\mathsf{Q}^*_{p_\epsilon}(\mathrm{RSP}(S,Q), \epsilon) = 0$, while $\mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \geq n$.*

Using quantum teleportation, any set of quantum states in space $\mathcal{H}$ can be prepared with zero error with communication cost $2n$. Thus, the above separation is maximal, up to the factor of 2.

To prove Proposition 5.4, we first analyze worst-case error protocols.

Lemma 5.5. *For any $\epsilon \in [0, 1/\sqrt{2}\,)$, $\mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \quad \geq \quad n$.*

Proof: Given any ARSP protocol Π for the given set of states Q, we construct an LOCC protocol Π' for transmitting n bits:

Protocol Π'

1. Alice, with input $x \in S$, and Bob (with no input) simulate the protocol Π.
2. Let σ_x be the output of Π, obtained by Bob. Bob measures σ_x according to the projective measurement $(|y\rangle\langle y| : y \in S)$.

The communication complexity of Π' equals that of Π.

Suppose Alice is given a uniformly random input, and let X be the corresponding random variable. Let Y be the random variable corresponding to Bob's output in Π'. Then, by the monotonicity of fidelity under quantum channels, the success probability of Π' is

$$\Pr[Y = X] \quad \geq \quad \frac{1}{2^n} \sum_x \mathrm{F}(\sigma_x, Q(x))^2 \quad \geq \quad 1 - \epsilon^2 \ .$$

By Theorem 1.2, the communication cost of Π', and therefore of Π, is at least $n + \log(1-\epsilon^2)$. Since $\epsilon \in [0, \frac{1}{\sqrt{2}})$, we have $\log(1-\epsilon^2) > -1$. So $\mathsf{Q}^*(\mathrm{RSP}(S,Q), \epsilon) \geq n$. ■

We show that the complexity of the task drops drastically, if average-case error is considered.

Lemma 5.6. *For every $\epsilon \in [0, 1/\sqrt{2}\,)$, There is a probability distribution p_ϵ over the set S such that $\mathsf{Q}^*_{p_\epsilon}(\mathrm{RSP}(S,Q), \epsilon) = 0$.*

Proof: Fix some $x_0 \in S$. Let p_ϵ be the probability distribution defined by

$$p_{\epsilon,x} \quad = \quad \begin{cases} \sqrt{1-\epsilon^2} & x = x_0 \\ \frac{1-\sqrt{1-\epsilon^2}}{2^n - 1} & x \neq x_0 \ . \end{cases}$$

Consider the protocol Π in which Alice does not send any message to Bob, and Bob always prepares the state $Q(x_0) = |x_0\rangle\langle x_0|$. The final joint state of the input-output registers in the protocol Π is

$$\rho'_{AB} \quad = \quad \sum_{x \in S} p_{\epsilon,x} |x\rangle\langle x| \otimes Q(x_0)$$

and the communication cost is zero. Denoting by ρ_{AB} the ideal input-output state, we have

$$\mathrm{F}(\rho_{AB}, \rho'_{AB}) \quad \geq \quad \sqrt{1-\epsilon^2} \ .$$

So $\mathsf{Q}^*_p(\mathrm{RSP}(S,Q),\epsilon) = 0$. ■

Thus we conclude Proposition 5.4. In fact we can construct an ensemble independent of ϵ, which exhibits a similar disparity between worst and average-case ARSP.

Proposition 5.7. *There is a probability distribution p over S such that for every $\epsilon \in [0, 1/\sqrt{2}\,)$, we have*

$$\mathsf{Q}^*_p(\mathrm{RSP}(S,Q),\epsilon) \quad \leq \quad \log\left(\min\left\{2^n, \log_2 \frac{2}{\epsilon^2}\right\}\right) + 2 \ .$$

Proof: Let $m \coloneqq 2^n$. Define p as the geometrically decreasing probability distribution

$$p_x \quad = \quad \begin{cases} \frac{1}{2^x} & x \in \{1, \ldots, m-1\} \\ \frac{1}{2^{m-1}} & x = m \end{cases} \ .$$

Now consider the following protocol Π for ARSP. If Alice's input x belongs to the set $\{1, \ldots, t\}$ with $t = \min\{\lceil \log \frac{2}{\epsilon^2} \rceil, m\}$, then she sends x to Bob. Otherwise, she sends a random number chosen from the set $\{1, \ldots, t\}$ to Bob. After receiving Alice's message y, Bob outputs the state $Q(y)$.

In protocol Π, the final state of Alice and Bob is of the form

$$\rho'_{AB} \quad \coloneqq \quad \sum_{x=1}^{m} p_x |x\rangle\langle x| \otimes \sigma_x \ ,$$

where $\sigma_x = Q(x)$ for $x \leq t$. Consequently

$$\mathrm{F}(\rho_{AB}, \rho'_{AB}) \quad = \quad \sum_{x=1}^{m} p_x \, \mathrm{F}(Q(x), \sigma_x) \quad \geq \quad \sum_{x=1}^{t} p_x \quad \geq \quad \sqrt{1-\epsilon^2} \ .$$

Therefore, the average-case error is at most ϵ, and the communication is $\lceil \log t \rceil$. This implies that

$$\mathsf{Q}^*_p(\mathrm{RSP}(S,Q),\epsilon) \quad \leq \quad \log\left(\min\left\{2^n, \log \frac{2}{\epsilon^2}\right\}\right) + 2 \ ,$$

as claimed. ■

This example illustrates that the more sharply skewed the probability distribution over Q, the bigger the gap between the worst-case and the average-case is. The example in Lemma 5.6 is a limiting case of such a distribution.

5.3 Connection to the asymptotic case

It is worth mentioning that our bounds for the average-case communication complexity of ARSP in the one-shot scenario also gives the optimal bounds in the asymptotic scenario established earlier by Berry and Sanders [7]. This can be derived using the *Quantum Asymptotic Equipartition Property* of max-information, i.e., Theorem 2.6. In the asymptotic scenario, Alice is given n independent and identically distributed inputs. Using the notation from Section 3, the target joint state of Alice's input and Bob's output is $\rho(p)^{\otimes n}$, and the goal is to prepare it approximately on Bob's side with average error ϵ.

Let $\mathsf{q}_p^*(\mathrm{RSP}(S,Q),\epsilon)$ denote the asymptotic rate of communication complexity of ARSP with average error ϵ. This is the limit of the communication complexity of preparing $\rho(p)^{\otimes n}$ with average-case error ϵ, divided by n, as $n \to \infty$. By Theorems 3.1 and 3.3, we have

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{I}_{\max}^{\epsilon}(A:B)_{\rho(p)^{\otimes n}} \quad \leq \quad \mathsf{q}_p^*(\mathrm{RSP}(S,Q),\epsilon) \quad \leq \quad \lim_{n\to\infty} \frac{1}{n} \left(\mathrm{I}_{\max}^{\delta}(A:B)_{\rho(p)^{\otimes n}} + \log_2 \ln \frac{8}{\epsilon^2} + 2 \right) \ ,$$

where $\delta = \frac{\epsilon}{2\sqrt{2}}$. So by inequalities (2.1) and (2.2) in Theorem 2.6, we get the following bounds:

$$\mathrm{I}(A:B)_{\rho(p)} - 2\epsilon \log(|A|\,|B|) \quad \leq \quad \mathsf{q}_p^*(\mathrm{RSP}(S,Q),\epsilon) \quad \leq \quad \mathrm{I}(A:B)_{\rho(p)} \ .$$

6 On LOCC protocols for transmitting bits

In this section, we digress from the main theme of this article; we characterize the communication required to convey classical bits through LOCC protocols as in Theorem 1.2. We have used this in Section 5 to highlight a key difference between worst-case and average-case protocols for remote state preparation.

Consider the following communication task $\mathcal{T}$:

> Two physically separated parties, Alice and Bob, have unlimited computational power and can communicate with each other. Alice is given a uniformly random n-bit string X unknown to Bob, that is independent of their initial state. Alice and Bob communicate with each other so that Bob learns X with probability at least $p \in (0,1]$.

Consider a classical communication protocol in which Alice sends exactly $\lceil n - \log \frac{1}{p} \rceil$ bits of X, and Bob chooses uniformly random bits as his guess for the remaining bits. Then the probability that Bob correctly decodes Alice's message is at least p. In this section, we show that even if we allow Alice and Bob to use LOCC protocols, the classical communication complexity of the task $\mathcal{T}$ does not decrease. In other words, in any (potentially two-way) LOCC protocol for this task, Alice sends at least $n + \log p$ bits in order to achieve success probability at least p (Theorem 1.2). Nayak and Salzman [35] showed that in any two-way *quantum* communication protocol with shared entanglement for the task $\mathcal{T}$, Alice sends at least $\frac{1}{2}(n + \log p)$ qubits to Bob. We obtain Theorem 1.2 by strengthening their proof.

6.1 Preparation

In LOCC protocols we assume that Alice and Bob each have access to an arbitrarily large but finite supply of qubits in some fixed basis state, say $|\bar{0}\rangle$. Without loss of generality, we further assume that during a protocol, each party performs some unitary operation followed by the measurement of a subset of qubits in the standard basis. Note that any measurement can be implemented in this manner [37, Sec 2.2.8]. Further, if the subset of qubits measured is of size k, we may assume that it consists of the leftmost k qubits.

We state some properties of protocols and states from Ref. [35] which are used later in this section. For completeness we include their proofs here.

Proposition 6.1 ([35]). *In any communication protocol with prior entanglement and local quantum channels, we may assume that the initial shared quantum state is of the form*

$$(\mathbb{1}_A \otimes \Lambda) \sum_{r\in\{0,1\}^e} |r\rangle_A |r\rangle_B \ ,$$

for some $\Lambda := \sum_{r\in\{0,1\}^e} \sqrt{\lambda_r} |r\rangle\langle r|$ *with* $\lambda_r \geq 0$, $\sum_{r\in\{0,1\}^e} \lambda_r = 1$, *and for some integer* $e \geq 1$.

Proof: Without loss of generality, assume that Alice and Bob hold e_A and e_B qubits of the initial state, respectively, where $e_B \geq e_A$. Let $|\phi\rangle = \sum_{i \in \{0,1\}^{e_A}} \sqrt{\gamma_i} |a_i\rangle_A |b_i\rangle_B$ be a Schmidt decomposition of the initial shared state.

We define a new protocol in which Alice and Bob start with the shared state $|\psi\rangle \coloneqq \sum_{r \in \{0,1\}^{e_B}} \sqrt{\lambda_r} |r\rangle_A |r\rangle_B$, where $\lambda_{\bar{0}s} = \gamma_s$ for $s \in \{0,1\}^{e_A}$ and is zero otherwise. The state simplifies to

$$\sum_{i \in \{0,1\}^{e_A}} \sqrt{\gamma_i} |\bar{0}, i\rangle_A |\bar{0}, i\rangle_B \ .$$

Using appropriate local unitary operators, Alice and Bob produce the state $|\phi\rangle$ (tensored with some fixed pure state), and then run the original protocol. ■

Proposition 6.2 ([35]). *For any linear transformation T on e qubits and any orthonormal set $\{|\phi_a\rangle : a \in \{0,1\}^e\}$ over $e' \geq e$ qubits,*

$$\sum_{a \in \{0,1\}^e} T|a\rangle \otimes |\phi_a\rangle \quad = \quad \sum_{a \in \{0,1\}^e} |a\rangle \otimes \tilde{T}|\phi_a\rangle \ ,$$

where $\tilde{T}$ is any transformation on e' qubits such that for all $a' \in \{0,1\}^e$, $\tilde{T}|\phi_{a'}\rangle = \sum_{a \in \{0,1\}^e} \langle a'|T|a\rangle |\phi_a\rangle$. If T is a unitary operation, then we may take $\tilde{T}$ to be a unitary operation on e' qubits.

Proof: Since the set $\{|a\rangle : a \in \{0,1\}^e\}$ is an orthonormal basis for the Hilbert space of e qubits, we have

$$\begin{aligned} \sum_{a \in \{0,1\}^e} T|a\rangle|\phi_a\rangle &= \sum_a \sum_{a'} \langle a'|T|a\rangle |a'\rangle |\phi_a\rangle \\ &= \sum_{a'} |a'\rangle \sum_a \langle a'|T|a\rangle |\phi_a\rangle \\ &= \sum_{a'} |a'\rangle \tilde{T} |\phi_{a'}\rangle \ , \end{aligned}$$

as claimed. The second part of the proposition is straightforward. ■

We also use this property in the following form in our analysis. The proof is straightforward, and is omitted.

Corollary 6.3. *For any controlled unitary operation $T \coloneqq \sum_{z \in \{0,1\}^m} |z\rangle\langle z| \otimes T_z$ on a classical-quantum register with m bits and e qubits, and collections of orthonormal sets $\{|\psi_{za}\rangle : a \in \{0,1\}^e\}$ over e' qubits with $e' \geq e$ and $z \in \{0,1\}^m$,*

$$\sum_{z \in \{0,1\}^m} \sum_{a \in \{0,1\}^e} T|za\rangle \otimes |z\rangle|\psi_{za}\rangle \quad = \quad \sum_{z \in \{0,1\}^m} \sum_{a \in \{0,1\}^e} |za\rangle \otimes \tilde{T}(|z\rangle|\psi_{za}\rangle) \ ,$$

where $\tilde{T} \coloneqq \sum_{z \in \{0,1\}^m} |z\rangle\langle z| \otimes \tilde{T}_z$, and $(\tilde{T}_z)$ is a sequence of unitary transformations on e' qubits such that for all $z \in \{0,1\}^m$ and $a' \in \{0,1\}^e$, $\tilde{T}_z|\psi_{za'}\rangle = \sum_{a \in \{0,1\}^e} \langle a'|T_z|a\rangle |\psi_{za}\rangle$.

6.2 One-way LOCC protocols

As a warm-up, we prove the analogue of Theorem 1.2 for one-way LOCC protocols.

Theorem 6.4. *Let Y be Bob's output in any one-way LOCC protocol for task $\mathcal{T}$ when Alice receives uniformly distributed n-bit input X. Let $p \coloneqq \Pr[Y = X]$ be the probability that Bob gets the output X. Then*

$$m \quad \geq \quad n - \log \frac{1}{p} \ ,$$

where m is the number of classical bits Alice sends to Bob in the protocol.

Proof: Using Proposition 6.1, we assume that the initial shared entangled state is $\sum_{r\in\{0,1\}^e}|r\rangle\Lambda|r\rangle$ for some $\Lambda := \sum_{r\in\{0,1\}^e}\sqrt{\lambda_r}|r\rangle\langle r|$ with $\lambda_r \geq 0$ and $\sum_{r\in\{0,1\}^e}\lambda_r = 1$, and some $e \geq 1$. As explained in Section 6.1, first Alice performs a unitary transformation on her part of the initial state depending on her input X and measures the left-most m qubits in the standard basis. Let U_x be the unitary operation Alice uses when she is given x as input. After the unitary operation U_x is performed, the joint state is

$$\begin{aligned}(U_x\otimes\mathbb{1})(\mathbb{1}\otimes\Lambda)\sum_{r\in\{0,1\}^e}|r\rangle\otimes|r\rangle &= (\mathbb{1}\otimes\Lambda)(U_x\otimes\mathbb{1})\sum_{r\in\{0,1\}^e}|r\rangle\otimes|r\rangle \\ &= (\mathbb{1}\otimes\Lambda)(\mathbb{1}\otimes U_x^{\mathsf{T}})\sum_{r\in\{0,1\}^e}|r\rangle\otimes|r\rangle \qquad \text{(By Proposition 6.2)} \\ &= \sum_{r\in\{0,1\}^e}|r\rangle\Lambda U_x^{\mathsf{T}}|r\rangle\ .\end{aligned}$$

Then Alice measures the state as described above and sends Bob the outcome of her measurement. Bob's state after this step is

$$\xi_x = \sum_{z\in\{0,1\}^m}|z\rangle\langle z|\otimes\Lambda U_x^{\mathsf{T}}(|z\rangle\langle z|\otimes\mathbb{1})\overline{U}_x\Lambda^*\ .$$

Note that

$$\xi_x = (\mathbb{1}\otimes\Lambda)\left(\sum_{z\in\{0,1\}^m}|z\rangle\langle z|\otimes U_x^{\mathsf{T}}(|z\rangle\langle z|\otimes\mathbb{1})\overline{U}_x\right)(\mathbb{1}\otimes\Lambda^*) \leq (\mathbb{1}\otimes\Lambda\Lambda^*)\ , \tag{6.1}$$

where the identity operator acts on a 2^m dimensional space. Finally, Bob performs a projective measurement $\{P_y\}_{y\in\{0,1\}^n}$ on his qubits, and gets as outcome the random variable Y. The success probability p of the protocol is

$$\begin{aligned}\Pr[X=Y] &= \sum_{x\in\{0,1\}^n}\Pr[X=x]\Pr[Y=x|X=x] \\ &= \sum_{x\in\{0,1\}^n}\frac{1}{2^n}\operatorname{Tr}(P_x\xi_x) \\ &\leq \frac{1}{2^n}\sum_x\operatorname{Tr}\big(P_x(\mathbb{1}\otimes\Lambda\Lambda^*)\big) \qquad \text{(By equation (6.1))} \\ &= \frac{1}{2^n}\operatorname{Tr}(\mathbb{1}\otimes\Lambda\Lambda^*) \\ &= \frac{2^m}{2^n}\ .\end{aligned}$$

We conclude that $m \geq n + \log p$. ■

6.3 The extension to two-way LOCC protocols

We now extend the above result to any two-way LOCC protocol. In particular we prove Theorem 1.2, which we restate here for convenience.

Theorem 6.5. *Let Y be Bob's output in any two-way LOCC protocol for task $\mathcal{T}$ when Alice receives uniformly distributed n-bit input X. Let m_A be the total number of bits Alice sends to Bob, and $p := \Pr[Y=X]$ be the probability that Bob produces output X. Then*

$$m_A \geq n - \log\frac{1}{p}\ .$$

To prove the theorem, we characterise the joint state of Alice and Bob at the end of a bounded round LOCC protocol.

Lemma 6.6. *Let Π be a bounded round LOCC protocol. Let e be the initial number of qubits with each of Alice and Bob, q be the total number of bits sent by Alice to Bob, q' be the total number of bits sent by Bob to Alice, and m be the total number of bits exchanged in Π (so $m = q + q'$). Then Alice and Bob's joint state at the end of the protocol (before the measurement for producing the output) can be written as*

$$\sum_{z\in\{0,1\}^m} \sum_{r,s\in\{0,1\}^{e-q}} |z,r\rangle\langle z,s|_A \otimes \Lambda|\phi_{z,r}\rangle\langle\phi_{z,s}|_B \Lambda^* \ ,$$

where

1. *A and B are classical-quantum registers with m-bit classical parts that contain the transcript of the protocol; register A is with Alice, and B with Bob,*
2. *Λ is a linear transformation that maps classical-quantum states with m bits and e qubits to classical-quantum states of the same form, depends only on the initial joint state and the unitary transformations applied by Bob, and satisfies $\mathrm{Tr}(\Lambda\Lambda^*) = 2^q$; and*
3. *$\{|\phi_{z,r}\rangle\}$ is an orthonormal set of classical-quantum states of the form $|\phi_{z,r}\rangle \coloneqq |z\rangle|\psi_{z,r}\rangle$ over m-bits and e qubits, and depends only on the initial joint state and the unitary transformations applied by Alice.*

Proof: Suppose that Π is a t-round LOCC protocol. Let ρ_i be the joint state of Alice and Bob after i-th round, and m_i be the total number of bits exchanged by Alice and Bob in the first i rounds, of which q_i bits are sent by Alice, for $1 \le i \le t$. Let ρ_0 be their initial state.

We prove the lemma by induction on t.

Base Case: Suppose that Π is a zero communication LOCC protocol, i.e., $t = 0$. By Proposition 6.1, we have

$$\rho_0 \quad = \quad \sum_{r,s\in\{0,1\}^e} |r\rangle\langle s| \otimes \Lambda|r\rangle\langle s|\Lambda^* \ ,$$

where $\Lambda = \sum_{r\in\{0,1\}^e} \sqrt{\lambda_r}|r\rangle\langle r|$ for some $\lambda_r \ge 0$ and $\sum_r \lambda_r = 1$. Since $\mathrm{Tr}(\Lambda\Lambda^*) = 1$, the state ρ_0 satisfies the claimed properties.

Induction Hypothesis: Suppose the lemma holds for any l-round LOCC protocol, for some $l \ge 0$.

Inductive Step: Suppose that Π is an $(l+1)$-round protocol. By the induction hypothesis, after the first l rounds of communication we have

$$\rho_l \quad = \quad \sum_{z\in\{0,1\}^{m_l}} \sum_{r,s\in\{0,1\}^{e-q_l}} |z,r\rangle\langle z,s| \otimes \Lambda_l|\phi_{z,r}\rangle\langle\phi_{z,s}|\Lambda_l^* \ ,$$

where Λ_l and $|\phi_{z,r}\rangle$ satisfy the properties stated in the lemma. In particular, suppose $|\phi_{z,r}\rangle \coloneqq |z\rangle|\psi_{z,r}\rangle$ for each z, r. We show that at the end of the protocol ρ_{l+1} is in the required form as well. Consider the $(l+1)$-th round of Π.

Case (1): Suppose that the communication in the last round is from Alice to Bob. Alice applies a unitary transformation $U \coloneqq \sum_z |z\rangle\langle z| \otimes U_z$, which acts on the quantum part of her register, controlled by the classical part of her register. She then measures the k leftmost qubits in the standard basis, appends the outcome to the message transcript in her classical register, and sends the outcome a of her measurement to Bob. The joint state after applying U is

$$
\begin{aligned}
& (U \otimes \mathbb{1})(\mathbb{1} \otimes \Lambda_l) \left[\sum_{\substack{r,s \in \{0,1\}^{e-q_l} \\ z \in \{0,1\}^{m_l}}} |z,r\rangle\langle z,s| \otimes |\phi_{z,r}\rangle\langle\phi_{z,s}| \right] (\mathbb{1} \otimes \Lambda_l^*)(U^* \otimes \mathbb{1}) \\
= \quad & (\mathbb{1} \otimes \Lambda_l) \left[\sum_{r,s,z} U|z,r\rangle\langle z,s|U^* \otimes |\phi_{z,r}\rangle\langle\phi_{z,s}| \right] (\mathbb{1} \otimes \Lambda_l^*) \\
= \quad & (\mathbb{1} \otimes \Lambda_l) \left[\sum_{r,s,z} |z,r\rangle\langle z,s| \otimes \tilde{U}|\phi_{z,r}\rangle\langle\phi_{z,s}|\tilde{U}^* \right] (\mathbb{1} \otimes \Lambda_l^*) \ ,
\end{aligned}
$$

where $\tilde{U} := \sum_z |z\rangle\langle z| \otimes \tilde{U}_z$ is the unitary operation given by Corollary 6.3. After Alice performs her measurement and sends the measurement outcome a to Bob, say he stores the message in register M. Denote by $\mathbb{1}_M \otimes \Lambda_l \tilde{U}$ the operator $\Lambda_l \tilde{U}$ on the registers originally with Bob, extended to include the register M. (The order of the operators in tensor product does not represent the order of the registers.) The joint state then may be expressed as below.

$$
\rho_{l+1} \quad = \sum_{\substack{r',s' \in \{0,1\}^{e-(q_l+k)} \\ a \in \{0,1\}^k \\ z \in \{0,1\}^{m_l}}} |za,r'\rangle\langle za,s'| \otimes (\mathbb{1}_M \otimes \Lambda_l \tilde{U})(|z\rangle\langle z| \otimes |a\rangle\langle a|_M \otimes |\psi_{z,ar'}\rangle\langle\psi_{z,as'}|)(\mathbb{1}_M \otimes \tilde{U}^* \Lambda_l^*) \ ,
$$

where $\Lambda_l \tilde{U}$ acts on the classical-quantum register with Bob before the message was sent. We define $\Lambda_{l+1} := \mathbb{1}_M \otimes \Lambda_l$, and $|\phi_{z',r'}\rangle := |za\rangle \otimes \tilde{U}_z |\psi_{z,ar'}\rangle$, where $z' := za$. Noting that $m_{l+1} = m_l + k$ and $q_{l+1} = q_l + k$, we have

$$
\rho_{l+1} \quad = \sum_{\substack{r',s' \in \{0,1\}^{e-q_{l+1}} \\ z' \in \{0,1\}^{m_{l+1}}}} |z',r'\rangle\langle z',s'| \otimes \Lambda_{l+1} |\phi_{z',r'}\rangle\langle\phi_{z',s'}| \Lambda_{l+1}^* \ .
$$

Further note that $\mathrm{Tr}(\Lambda_{l+1}\Lambda_{l+1}^*) = 2^{q_{l+1}}$ and $\{|\phi_{z',r'}\rangle\}$ is an orthonormal set of the claimed form.

Case (2): Suppose that the communication in the last round is from Bob to Alice. Bob applies a unitary transformation $V := \sum_z |z\rangle\langle z| \otimes V_z$ to the quantum part of his register, controlled by the classical part of his register. Then he measures the k leftmost qubits (say in sub-register L) in the standard basis, and appends the outcome b to the message transcript, in classical register M. Finally, he sends the outcome b of the measurement to Alice. Denote by $\mathbb{1}_M \otimes (\langle b|_L \otimes \mathbb{1}) V \Lambda_l$, the extension of the operator $(\langle b|_L \otimes \mathbb{1}) V \Lambda_l$ to include the register M. (Here, the order of the operators in tensor product does not represent the order of the registers on which they act. The same applies to the operator Λ_{l+1} defined below.) The joint state then is as follows.

$$
\rho_{l+1} \quad = \sum_{\substack{r,s \in \{0,1\}^{e-q_l} \\ b \in \{0,1\}^k \\ z \in \{0,1\}^{m_l}}} |zb,r\rangle\langle zb,s| \otimes (\mathbb{1}_M \otimes (\langle b|_L \otimes \mathbb{1}) V \Lambda_l)(|zb\rangle\langle zb| \otimes |\psi_{z,r}\rangle\langle\psi_{z,s}|)(\mathbb{1}_M \otimes \Lambda_l^* V^* (|b\rangle_L \otimes \mathbb{1})) \ .
$$

Note that $q_{l+1} = q_l$, and $m_{l+1} = m_l + k$. Define $\Lambda_{l+1} := \sum_b |b\rangle\langle b|_M \otimes (\langle b|_L \otimes \mathbb{1}) V \Lambda_l$ and $|\phi_{z',r'}\rangle = |zb\rangle \otimes |\psi_{z,br'}\rangle$, where $z' := zb$. It is straightforward to verify that $\mathrm{Tr}(\Lambda_{l+1}\Lambda_{l+1}^*) = 2^{q_{l+1}}$, the set $\{|\phi_{z',r'}\rangle\}$ is of the claimed form, and

$$
\rho_{l+1} \quad = \sum_{\substack{r',s' \in \{0,1\}^{e-q_{l+1}} \\ z' \in \{0,1\}^{m_{l+1}}}} |z',r'\rangle\langle z',s'| \otimes \Lambda_{l+1} |\phi_{z',r'}\rangle\langle\phi_{z',s'}| \Lambda_{l+1}^* \ .
$$

This completes the proof. ■

We are ready to prove Theorem 1.2, restated in this section as Theorem 6.5.

Proof of Theorem 6.5: By Lemma 6.6, at the end of any two-way LOCC protocol, when Alice has input $x \in \{0,1\}^n$, Bob's state before performing his final measurement to get Y is

$$\xi_x \quad = \sum_{\substack{r\in\{0,1\}^{e-m_A} \\ z\in\{0,1\}^m}} \Lambda|\phi_{z,r}(x)\rangle\langle\phi_{z,r}(x)|\Lambda^* \ ,$$

for some linear transformation Λ with $\mathrm{Tr}(\Lambda\Lambda^*) = 2^{m_A}$ and orthonormal set $\{|\phi_{z,r}(x)\rangle\}_{z,r}$. The transformation Λ only depends on Bob's unitary operations and the initial state, and is therefore independent of Alice's input x. Note that

$$\xi_x \quad \leq \quad \Lambda\Lambda^* \ . \tag{6.2}$$

After Bob performs his final projective measurement $\{P_y\}_{y\in\{0,1\}^n}$ and gets the output Y, the probability of correctly recovering an input X chosen uniformly at random is

$$\begin{aligned} p \quad := \quad \Pr[Y = X] \quad &= \quad \frac{1}{2^n} \sum_{x\in\{0,1\}^n} \mathrm{Tr}(P_x \xi_x) \\ &\leq \quad \frac{1}{2^n} \sum_x \mathrm{Tr}(P_x \Lambda\Lambda^*) && \text{(Equation (6.2))} \\ &= \quad \frac{1}{2^n} \mathrm{Tr}(\Lambda\Lambda^*) \quad = \quad \frac{2^{m_A}}{2^n} \ . \end{aligned}$$

Therefore, we have $m_A \geq n - \log \frac{1}{p}$, as required. ■

7 Conclusion

In this article, we studied the communication complexity of remote state preparation in the one-shot scenario. Our main results can be summarized as follows:

- The communication complexity of remote state preparation with bounded average-case error ϵ can be characterized tightly in terms of the smooth max-information Bob's output has about Alice's input.
- The communication complexity of remote state preparation with bounded worst-case error ϵ can be characterized in terms of a similar natural expression involving smooth max-relative entropy.

The bounds we derive for the worst-case communication complexity are provably tighter than earlier ones. We also show out how protocols that guarantee low worst-case error necessarily use more communication than those that require low error on average. In the process, we strengthen a lower bound on LOCC protocols for transmitting classical bits.

In this work, we focused on the remote preparation of a possibly mixed quantum state. However, often the quantum state to be remotely prepared is entangled with other systems ("the environment"). We can consider the problem of preparing an approximation of the quantum state such that its entanglement with other systems does not change significantly. This problem has been studied in asymptotic scenario [6, 7]. Berta [8] implicitly studied this problem in the one-shot scenario by considering the *quantum state merging* problem, and showed that the minimal entanglement cost needed for this problem is equal to minus the ϵ-smooth conditional min-entropy of Alice's register conditioned on the environment, while classical communication is allowed for free. Note that the entanglement cost is defined as the difference between the number of bits of pure entanglement at the beginning and at the end of the process. It would be interesting to characterize the minimum classical communication of such "faithful" ARSP in terms of non-asymptotic information theoretic quantities.

References

[1] Scott Aaronson. Guest column: NP-complete problems and physical reality. *SIGACT News*, 36(1):30–52, March 2005.

[2] Shima Bab Hadiashar. *Communication Complexity of Remote State Preparation.* Master's Thesis, University of Waterloo, Waterloo, Canada, May 2014.

[3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70(13):1895–1899, March 1993.

[4] Charles H. Bennett, David P. DiVincenzo, Peter W. Shor, John A. Smolin, Barbara M. Terhal, and William K. Wootters. Remote state preparation. *Physical Review Letters*, 87(7):077902, July 2001.

[5] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54:3824–3851, Nov 1996.

[6] Charles H. Bennett, Patrick Hayden, Debbie W. Leung, Peter W. Shor, and Andreas Winter. Remote preparation of quantum states. *IEEE Transactions on Information Theory*, 51(1):56–74, January 2005.

[7] Dominic W. Berry and Barry C. Sanders. Optimal remote state preparation. *Physical Review Letters*, 90:057901, February 2003.

[8] Mario Berta. *Single-Shot Quantum State Merging.* Diploma thesis, ETH, Zurich, February 2008.

[9] Mario Berta, Matthias Christandl, and Renato Renner. The Quantum Reverse Shannon theorem based on one-shot information theory. *Communications in Mathematical Physics*, 306(3):579–615, September 2011.

[10] Igor Bjelakovic and Rainer Siegmund-Schultze. Quantum Stein's lemma revisited, inequalities for quantum entropies, and a concavity theorem of Lieb. Technical Report arXiv:quant-ph/0307170v2, arXiv.org, 2003.

[11] Lenore Blum, Mike Shub, and Steve Smale. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bulletin of the American Mathematical Society*, 21:1–46, 1989.

[12] Fernando G. S. L. Brandão and Nilanjana Datta. One-shot rates for entanglement manipulation under non-entangling maps. *IEEE Transactions on Information Theory*, 57(3):1754–1760, March 2011.

[13] Francesco Buscemi and Nilanjana Datta. The quantum capacity of channels with arbitrarily correlated noise. *IEEE Transactions on Information Theory*, 56(3):1447–1460, March 2010.

[14] Eric Chitambar, Debbie Leung, Laura Mancinska, Maris Ozols, and Andreas Winter. Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask). *Communications in Mathematical Physics*, 328(1):303–326, 2014.

[15] Nikola Ciganovic, Normand J. Beaudry, and Renato Renner. Smooth max-information as one-shot generalization for mutual information. *Information Theory, IEEE Transactions on*, 60(3):1573–1581, March 2014.

[16] Nilanjana Datta. Min- and max-relative entropies and a new entanglement monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, June 2009.

[17] Igor Devetak and Toby Berger. Low-entanglement remote state preparation. *Physical Review Letters*, 87(19), October 2001.

[18] Dawei Ding and Mark M. Wilde. Strong converse exponents for the feedback-assisted classical capacity of entanglement-breaking channels. Technical Report arXiv:1506.02228v4 [quant-ph], arXiv.org, 2017.

[19] Frédéric Dupuis, Lea Kraemer, Philippe Faist, Joseph M. Renes, and Renato Renner. Generalized entropies. In *Proceedings of the XVIIth International Congress on Mathematical Physics, Aalborg, Denmark, 2012*, pages 134–153, August 2012.

[20] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71:062310, Jun 2005.

[21] Akihisa Hayashi, Takeji Hashimoto, and Minoru Horibe. Remote state preparation without oblivious conditions. *Physical Review A*, 67:052302, May 2003.

[22] Masahito Hayashi. Role of hypothesis testing in quantum information. Technical Report arXiv:1709.07701 [quant-ph], arXiv.org, 2017.

[23] Masato Hayashi and Hiroshi Nagaoka. A general formula for the classical capacity of a general quantum channel. In *Proceedings of the IEEE International Symposium on Information Theory*. IEEE, 2002.

[24] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, 1991.

[25] Rahul Jain. Communication complexity of remote state preparation with entanglement. *Quantum Information & Computation*, 6(4):461–464, July 2006.

[26] Rahul Jain and Ashwin Nayak. Short proofs of the quantum Substate theorem. *IEEE Transactions on Information Theory*, 58(6):3664–3669, June 2012.

[27] Rahul Jain, Ashwin Nayak, and Yi Su. A separation between divergence and Holevo information for ensembles. *Mathematical Structures in Computer Science*, 20(5):977–993, 2010.

[28] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 429–438, 2002.

[29] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Prior entanglement, message compression and privacy in quantum communication. In *Proceedings of the Twentieth Annual IEEE Conference on Computational Complexity, 2005*, pages 285–296, June 2005.

[30] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A property of quantum relative entropy with an application to privacy in quantum communication. *Journal of the ACM*, 56(6):33:1–33:32, September 2009.

[31] Felix Leditzky, Mark M. Wilde, and Nilanjana Datta. Strong converse theorems using Rényi entropies. *Journal of Mathematical Physics*, 57(8):082202, 2016.

[32] Debbie W. Leung and Peter W. Shor. Oblivious remote state preparation. *Physical Review Letters*, 90(12):127905, March 2003.

[33] Hoi-Kwong Lo. Classical-communication cost in distributed quantum-information processing: A generalization of quantum-communication complexity. *Physical Review A*, 62(1):012313, June 2000.

[34] William Matthews and Stephanie Wehner. Finite blocklength converse bounds for quantum channels. *IEEE Transactions on Information Theory*, 60(11):7317–7329, 2014.

[35] Ashwin Nayak and Julia Salzman. Limits on the ability of quantum states to convey classical messages. *Journal of the ACM*, 53(1):184 – 206, January 2006.

[36] Ashwin Nayak and Peter Shor. Bit-commitment-based quantum coin flipping. *Physical Review A*, 67:012304, January 2003.

[37] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge university press, 2010.

[38] Tomohiro Ogawa and Hiroshi Nagaoka. Strong converse and Stein's lemma in quantum hypothesis testing. *IEEE Transactions on Information Theory*, 46(7):2428–2433, 2000.

[39] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory.* MIT Press, 1994.

[40] Alexey E. Rastegin. Relative error of state-dependent cloning. *Physical Review A*, 66(4):042304, 2002.

[41] Alexey E. Rastegin. A lower bound on the relative error of mixed-state cloning and related operations. *Journal of Optics B: Quantum and Semiclassical Optics*, 5(6):S647, 2003.

[42] Alexey E. Rastegin. Sine distance for quantum states. Technical Report arxiv:quant-ph/0602112, arXiv.org, 2006.

[43] Renato Renner. *Security of Quantum Key Distribution.* PHD thesis, ETH, Zurich, December 2005.

[44] Renato Renner and Stefan Wolf. Smooth Rényi entropy and applications. In *IEEE International Symposium on Information Theory*, page 233, 2004.

[45] Marco Tomamichel, Roger Colbeck, and Renato Renner. Duality between smooth min- and max-entropies. *IEEE Transactions on Information Theory*, 56(9):4674–4681, September 2010.

[46] Marco Tomamichel and Masahito Hayashi. A hierarchy of information quantities for finite block length analysis of quantum tasks. *IEEE Transactions on Information Theory*, 59(11):7693–7710, November 2013.

[47] Dave Touchette. Quantum information complexity. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 317–326. ACM, 2015.

[48] Ligong Wang and Renato Renner. One-shot classical-quantum capacity and hypothesis testing. *Physical Review Letters*, 108:200501, May 2012.

[49] John Watrous. *Theory of Quantum Information.* 2016. Book draft, September 2016. Available at `https://cs.uwaterloo.ca/~watrous/TQI/` .

[50] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched Réenyi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, October 2014.

[51] Andrew Chi-Chih Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science, 1993*, pages 352–361, November 1993.

A Some properties of entropic quantities

In this section, we present the proofs of some properties of information-theoretic quantities stated in Section 2.4. For convenience, we restate the properties here.

Proposition A.1 (Proposition 2.5)**.** *Let $\rho_{AB} \in \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ be a bipartite quantum state that is classical on A. For any $\epsilon \geq 0$, there exists $\rho'_{AB} \in \mathsf{B}^{\epsilon}(\rho_{AB}) \cap \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ classical on A such that*

$$\mathrm{I}^{\epsilon}_{\max}(A:B)_{\rho} \quad = \quad \mathrm{I}_{\max}(A:B)_{\rho'} \ .$$

Proof: Let $\lambda = \mathrm{I}^{\epsilon}_{\max}(A:B)_{\rho}$, and $\tilde{\rho}_{AB} \in \mathsf{B}^{\epsilon}(\rho_{AB})$ and $\sigma_B \in \mathsf{D}(\mathcal{H})$ be two quantum states for which

$$\tilde{\rho}_{AB} \quad \leq \quad 2^{\lambda}\, \tilde{\rho}_A \otimes \sigma_B \ .$$

Without loss of generality, we assume that $\tilde{\rho}_{AB}$ has trace equal to one, i.e., $\tilde{\rho}_{AB} \in \mathsf{B}^{\epsilon}(\rho_{AB}) \cap \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$. If not, we consider the state $\omega_{AB} := \frac{\tilde{\rho}_{AB}}{\mathrm{Tr}(\tilde{\rho}_{AB})}$ instead of $\tilde{\rho}_{AB}$. Since ρ has trace 1, $\mathrm{P}(\omega, \rho) \leq \mathrm{P}(\tilde{\rho}, \rho)$. Further, $\omega_{AB} \leq 2^{\lambda}\, \omega_A \otimes \sigma_B$.

Let $\Phi_A : \mathrm{L}(\mathcal{H}) \to \mathrm{L}(\mathcal{H})$ be a quantum-to-classical channel such that:

$$\Phi_A(X) = \sum_i \langle e_i|X|e_i\rangle |e_i\rangle\langle e_i|$$

for all $X \in \mathrm{L}(\mathcal{H})$, where $\{|e_i\rangle\}$ is the standard basis for $\mathrm{L}(\mathcal{H})$. Let $\rho'_{AB} = (\Phi_A \otimes \mathbb{1}_B)(\tilde{\rho}_{AB})$. By the definition of ρ'_{AB} and the monotonicity of purified distance $\rho'_{AB} \in \mathsf{B}^{\epsilon}(\rho_{AB}) \cap \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$.

By optimality of $\tilde{\rho}_{AB}$, we have

$$\mathrm{I}^{\epsilon}_{\max}(A:B)_{\rho} \quad = \quad \mathrm{I}_{\max}(A:B)_{\tilde{\rho}} \quad \leq \quad \mathrm{I}_{\max}(A:B)_{\rho'} \ ,$$

and by Proposition 2.4, monotonicity of smooth max-information, we have

$$\mathrm{I}_{\max}(A:B)_{\rho'} \quad \leq \quad \mathrm{I}_{\max}(A:B)_{\tilde{\rho}} \ .$$

Therefore, we conclude that

$$\mathrm{I}^{\epsilon}_{\max}(A:B)_{\rho} \quad = \quad \mathrm{I}_{\max}(A:B)_{\rho'} \ ,$$

where $\rho'_{AB} \in \mathsf{B}^{\epsilon}(\rho_{AB}) \cap \mathsf{D}(\mathcal{H}' \otimes \mathcal{H})$ and is classical on A. ■

Proposition A.2 (Proposition 2.8)**.** *Let $\rho_{AB}(p) \in \mathsf{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a state classical on A such that the distribution on A is given by the probability vector p. Let $\rho_A(p) = \mathrm{Tr}_B(\rho_{AB}(p))$, and $\sigma \in \mathsf{D}(\mathcal{H}_B)$ be a quantum state on Hilbert space $\mathcal{H}_B$. Then the function $\beta^{\epsilon}(\rho_{AB}(p) \,\|\, \rho_A(p) \otimes \sigma)$ is convex with respect to p.*

Proof: Let p_0 and p_1 be two arbitrary probability distributions on the standard basis of $\mathcal{H}_A$. For $\lambda \in [0,1]$, let $q = \lambda p_0 + (1-\lambda) p_1$. We show that

$$\beta^{\epsilon}(\rho_{AB}(q) \,\|\, \rho_A(q) \otimes \sigma) \quad \leq \quad \lambda\, \beta^{\epsilon}(\rho_{AB}(p_0) \,\|\, \rho_A(p_0) \otimes \sigma) + (1-\lambda)\, \beta^{\epsilon}(\rho_{AB}(p_1) \,\|\, \rho_A(p_1) \otimes \sigma) \ ,$$

which proves the claim.

Let $\Phi : \mathsf{L}(\mathcal{H}_A) \to \mathsf{L}(\mathbb{C}^2 \otimes \mathcal{H}_A)$ be the quantum channel with Kraus operators $A_{a,x} = \sqrt{\alpha^a_x}|x\rangle \otimes |a\rangle\langle a|$ for all a and $x \in \{0,1\}$, where $\alpha^a_0 := \lambda \frac{p_0(a)}{q(a)}$ and $\alpha^a_1 = (1-\lambda)\frac{p_1(a)}{q(a)}$. Then we have

$$\rho_{XAB}(q) \quad = \quad (\Phi \otimes \mathbb{1}_B)(\rho_{AB}(q)) \quad = \quad \lambda\, |0\rangle\langle 0| \otimes \rho_{AB}(p_0) + (1-\lambda)\, |1\rangle\langle 1| \otimes \rho_{AB}(p_1) \ .$$

Since $\rho_{XAB}(q)$ is an extension of $\rho_{AB}(q)$, using Proposition 2.7 twice, we get

$$\beta^{\epsilon}(\rho_{AB}(q) \,\|\, \rho_A(q) \otimes \sigma) \quad = \quad \beta^{\epsilon}(\rho_{XAB}(q) \,\|\, \rho_{XA}(q) \otimes \sigma) \ . \tag{A.1}$$

For each $x \in \{0,1\}$, let Q_x be the measurement operator that achieves $\beta^{\epsilon}(\rho_{AB}(p_x) \,\|\, \rho_A(p_x) \otimes \sigma)$. Consider the measurement operator $Q := \sum_{x \in \{0,1\}} |x\rangle\langle x| \otimes Q_x$. This satisfies

$$\langle Q, \rho_{XAB}(q)\rangle \quad = \quad \lambda\, \langle Q_0, \rho_{AB}(p_0)\rangle + (1-\lambda)\, \langle Q_1, \rho_{AB}(p_1)\rangle \quad \geq \quad 1 - \epsilon \ ,$$

by definition of Q_0, Q_1. By Eq. (A.1) and the definition of β^{ϵ}, we get

$$\begin{aligned}
\beta^{\epsilon}(\rho_{AB}(q) \,\|\, \rho_A(q) \otimes \sigma) \quad &= \quad \beta^{\epsilon}(\rho_{XAB}(q) \,\|\, \rho_{XA}(q) \otimes \sigma) \\
&\leq \quad \langle Q, \rho_{XA}(q) \otimes \sigma\rangle \\
&= \quad \lambda\, \langle Q_0, \rho_A(p_0) \otimes \sigma\rangle + (1-\lambda)\, \langle Q_1, \rho_A(p_1) \otimes \sigma\rangle \\
&= \quad \lambda\, \beta^{\epsilon}(\rho_{AB}(p_0) \,\|\, \rho_A(p_0) \otimes \sigma) + (1-\lambda)\, \beta^{\epsilon}(\rho_{AB}(p_1) \,\|\, \rho_A(p_1) \otimes \sigma) \ ,
\end{aligned}$$

as we set out to prove. ■

Proposition A.3 (Proposition 2.9)**.** *For any fixed quantum state $\rho \in \mathsf{D}(\mathcal{H})$, the function $\beta^\epsilon(\rho\|\sigma)$ is a concave function with respect to σ.*

Proof: For any choice of $\sigma_0, \sigma_1 \in \mathsf{D}(\mathcal{H})$ and $\lambda \in [0,1]$, let Q be the measurement operator that achieves hypothesis testing error $\beta^\epsilon(\rho \,\|\, \lambda\sigma_0 + (1-\lambda)\sigma_1)$. Then

$$\begin{aligned}\beta^\epsilon(\rho \,\|\, \lambda\sigma_0 + (1-\lambda)\sigma_1) &= \langle Q,\ \lambda\sigma_0 + (1-\lambda)\sigma_1\rangle \\ &= \lambda \langle Q, \sigma_0\rangle + (1-\lambda) \langle Q, \sigma_1\rangle \\ &\geq \lambda\, \beta^\epsilon(\rho\|\sigma_0) + (1-\lambda)\, \beta^\epsilon(\rho\|\sigma_1)\ ,\end{aligned}$$

since $\langle Q, \rho\rangle \geq 1-\epsilon$. ■

B Preparing states from an infinite set

In this section, we discuss remote state preparation of states drawn from an infinite set. This scenario has been studied by Lo [33] and in later works on the topic.

In remote state preparation, Alice's input is supposed to provide a complete description of the state to be prepared at Bob's end. In any physically realistic model of computation, the description necessarily has finite bit-length (see, e.g., Ref. [1]). For instance, if a d-dimensional quantum state is described by specifying $\Theta(d^2)$ complex entries in the corresponding $d \times d$ matrix, the complex numbers would have to be specified with finite precision. This implies that the input set S (following the notation in Section 2.6) is necessarily countable. This point has not been addressed in previous works.

To meaningfully consider the preparation states drawn from an uncountable set, we may instead consider *approximations* drawn from a suitable countable set. For example, instead of the set $\mathsf{D}(\mathcal{H})$ of all quantum states over a d-dimensional space $\mathcal{H}$, we may instead study the countably dense set of states whose matrix representations only have complex entries with rational real and imaginary parts. Such states have unique finite-length representations. (Similar approximation is also implicit in the case of RSP of a finite set of states, when the corresponding matrices involve irrational numbers.)

Another approach, perhaps only of theoretical interest, would be to allow the local operations in an LOCC protocol to be defined on a suitable generalization of the Real RAM model due to Blum, Shub, and Smale [11]. We do not attempt to define such a model of computation here. For our purposes, it would suffice to assume a model which enables the implementation of quantum operations such as unitary operations controlled by the registers holding real numbers in finite time.

We assume that we take one of the abovementioned approaches in the analysis in this section. The underlying idea, that of approximating states from an infinite set with those from a *net*, probably applies in other reasonable approaches as well.

As before, we restrict ourselves to states over a finite dimensional Hilbert space $\mathcal{H}$.

Definition B.1. *Let $\nu \in (0,1]$ and $D \subseteq \mathsf{D}(\mathcal{H})$ be any set of quantum states. A ν-*net *N in D is a subset of D such that for any state $\rho \in D$, there is a state $\sigma \in N$ such that $\mathrm{P}(\rho,\sigma) < \nu$.*

We argue that every subset of finite-dimensional states admits a finite net.

Proposition B.1. *For every $\nu \in (0,1]$, and every set $D \subseteq \mathsf{D}(\mathcal{H})$ of quantum states, there is a finite ν-net in D.*

Proof: Since $\mathsf{D}(\mathcal{H})$ is compact, it has a finite cover (B_i) consisting of open balls of radius $\nu/2$. This is also a cover for any subset D of quantum states. Let N be a subset of D constructed by taking one point from $B_i \cap D$, whenever this intersection is non-empty. We claim that this is a finite ν-net in D.

Consider a state $\rho \in D$. Since (B_i) is a cover for the set of all quantum states, $\rho \in B_j$ for some j. By construction, there is a state $\sigma \in N$ from $B_j \cap D$. Since ρ, σ both belong to the same ball B_j of radius $\nu/2$, we have $\mathrm{P}(\rho, \sigma) < \nu$. So N is a ν-net in D. ■

Suppose S is an infinite set, and $Q : S \to \mathsf{D}(\mathcal{H})$ is a one-to-one function mapping each element of S to a quantum state. (We view an element $x \in S$ as a description, i.e., unique encoding, of the quantum state $Q(x)$.) Define $R \coloneqq Q(S)$ as the image of S under Q; this is the set of quantum states under consideration. We fix an approximation parameter $\nu > 0$ of our choice, and a finite ν-net N in R, and let $T \coloneqq Q^{-1}(N)$ be the set of inputs corresponding to N. We bound the communication required for remote state preparation of states from R with that for states from N. We may then appeal to Theorem 1.1 to infer bounds on $\mathrm{RSP}(S, Q)$.

Worst-case error. We first consider the simpler case, that of worst-case error $\epsilon > 0$. Any protocol for $\mathrm{RSP}(S, Q)$ with worst-case error ϵ is also a protocol for $\mathrm{RSP}(T, Q)$ as T is a subset of S. So we have

$$\mathsf{Q}^*(\mathrm{RSP}(T, Q), \epsilon) \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(S, Q), \epsilon) \ .$$

Now suppose Π is a protocol for $\mathrm{RSP}(T, Q)$ with communication cost c and worst case error ϵ. We design a protocol Π' for $\mathrm{RSP}(S, Q)$ as follows. Given an $x \in S$, Alice chooses $y \in T$ such that $\mathrm{P}(Q(x), Q(y)) \leq \nu$, and prepares an approximation of $Q(y)$ on Bob's side using protocol Π. Suppose Bob's output is σ_y. Then

$$\mathrm{P}(Q(x), \sigma_y) \quad \leq \quad \mathrm{P}(Q(x), Q(y)) + \mathrm{P}(Q(y), \sigma_y) \quad \leq \quad \nu + \epsilon \ .$$

So Π' is a protocol for $\mathrm{RSP}(S, Q)$ with communication cost c, and worst case error $\epsilon + \nu$. Therefore,

$$\mathsf{Q}^*(\mathrm{RSP}(S, Q), \epsilon + \nu) \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(T, Q), \epsilon) \ .$$

Putting the two together, for ν, ϵ such that $0 < \nu < \epsilon$, we get

$$\mathsf{Q}^*(\mathrm{RSP}(T, Q), \epsilon) \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(S, Q), \epsilon) \quad \leq \quad \mathsf{Q}^*(\mathrm{RSP}(T, Q), \epsilon - \nu) \ .$$

Average-case error. Next we consider approximate RSP with average error at most $\epsilon \in (0, 1]$ with respect to a probability measure μ on the set of states R. For simplicity, we only consider the case when the open sets in R generated by the metric P are measurable. Since Q is injective, we may equivalently consider μ as a probability measure on S.

Let (ρ_i) be an enumeration of the states in N, and (B_i) be open balls of radius ν centred at ρ_i with respect to the metric P. Since N is a ν-net in R, we have $R \subseteq \cup_i B_i$. Define the function $f : R \to N$ as $f(\sigma) \coloneqq \rho_i$ for all states $\sigma \in (B_i \cap R) \setminus (\cup_{j<i} B_j)$. The function f maps each quantum state $\rho \in R$ to a quantum state in N such that $\mathrm{P}(\rho, f(\rho)) < \nu$. Moreover, it is measurable.

The function f induces a probability distribution p on N in the natural way:

$$p_{\rho_i} \quad \coloneqq \quad \mu(f^{-1}(\rho_i))$$

for $\rho_i \in N$. We may view the distribution p as being over the corresponding set T of inputs: for $y \in T$ such that $Q(y) = \rho_i$, we define $p_y \coloneqq p_{\rho_i}$.

We relate protocols for $\mathrm{RSP}(S, Q)$ with average error ϵ with respect to μ to protocols for $\mathrm{RSP}(T, Q)$ with average error "close" to ϵ with respect to p.

Lemma B.2. *Suppose* Π *is a protocol for* $\mathrm{RSP}(S, Q)$ *with communication cost* c *and average error* ϵ *with respect to* μ. *Then there is a protocol* Π' *for* $\mathrm{RSP}(T, Q)$ *with communication cost* c *and average error at most* $\nu + \epsilon$ *with respect to* p.

Proof: For $y \in T$, define $R_y \coloneqq f^{-1}(Q(y))$, the set of states in R that are mapped to $Q(y) \in N$. Define $S_y \coloneqq Q^{-1}(R_y)$, the set of inputs corresponding to the states in R_y. Note that (R_y) is a partition of R and (S_y) of S. Since f is measurable, R_y is a measurable set. When R_y has non-zero measure, we define a probability measure μ_y on R_y as $\mu_y(W) \coloneqq \mu(W)/\mu(R_y)$ for all measurable sets $W \subseteq R_y$. We also view μ_y as a probability measure on S_y.

We now construct the protocol Π' for $\mathrm{RSP}(T, Q)$ as follows. Given $y \in T$, Alice selects an input $x \in S_y$ randomly with respect to the probability measure μ_y and runs the protocol Π on this input.

The communication in Π' is also c. Suppose σ_x is the output of the protocol Π when the input is x. Then the average error of the protocol Π' is

$$\begin{aligned}
\sum_{y \in T} p_y \int_{x \in S_y} \mathrm{P}(Q(y), \sigma_x) \; \mathrm{d}\mu_y(x) \quad &= \quad \sum_{y \in T} \int_{x \in S_y} \mathrm{P}(Q(y), \sigma_x) \; \mathrm{d}\mu(x) \\
&\leq \quad \sum_{y \in T} \int_{x \in S_y} \mathrm{P}(Q(y), Q(x)) \; \mathrm{d}\mu(x) \\
&\quad\quad + \sum_{y \in T} \int_{x \in S_y} \mathrm{P}(Q(x), \sigma_x) \; \mathrm{d}\mu(x) \\
&\leq \quad \nu + \epsilon \;,
\end{aligned}$$

as claimed. ■

Conversely, we can also derive a protocol for $\mathrm{RSP}(S, Q)$ from one for $\mathrm{RSP}(T, Q)$.

Lemma B.3. *Suppose* Π *is a protocol for* $\mathrm{RSP}(T, Q)$ *with communication cost* c *and average error* ϵ *with respect to the distribution* p. *There exists a protocol* Π' *for* $\mathrm{RSP}(S, Q)$ *with communication cost* c *and average error at most* $\epsilon + \nu$ *with respect to* μ.

Proof: In the protocol Π', given input $x \in S$, Alice runs the protocol Π on input y defined as $y \coloneqq Q^{-1}(f(Q(x)))$. This is the input corresponding to the state in the ν-net to which f maps $Q(x)$. The communication cost of Π' is also c.

Suppose the output of Π' on input x is σ_x. Note that f maps all states $Q(x)$ for $x \in S_y$ to the same value $Q(y)$, and therefore the outputs σ_x for all inputs $x \in S_y$ are equal to σ_y.

The average error of the protocol with respect to μ is

$$\begin{aligned}
\int_{x \in S} \mathrm{P}(Q(x), \sigma_x) \; \mathrm{d}\mu(x) \quad &\leq \quad \int_{x \in S} \mathrm{P}(Q(x), f(Q(x))) \; \mathrm{d}\mu(x) + \int_{x \in S} \mathrm{P}(f(Q(x)), \sigma_x) \; \mathrm{d}\mu(x) \\
&\leq \quad \nu + \sum_{y \in T} \int_{x \in S_y} \mathrm{P}(f(Q(x)), \sigma_x) \; \mathrm{d}\mu(x) \\
&= \quad \nu + \sum_{y \in T} \int_{x \in S_y} \mathrm{P}(Q(y), \sigma_y) \; \mathrm{d}\mu(x) \\
&= \quad \nu + \sum_{y \in T} p_y \, \mathrm{P}(Q(y), \sigma_y) \\
&\leq \quad \nu + \epsilon \;,
\end{aligned}$$

where in the third step we have used the abovementioned property that f is constant on S_y. ■

For ν, ϵ such that $\nu < \epsilon$, the above two lemmata imply that

$$\mathsf{Q}^*_p(\mathrm{RSP}(T,Q), \epsilon + \nu) \quad \leq \quad \mathsf{Q}^*_\mu(\mathrm{RSP}(S,Q), \epsilon) \quad \leq \quad \mathsf{Q}^*_p(\mathrm{RSP}(T,Q), \epsilon - \nu) \ .$$