Generalized resource theory of purity: one-shot purity distillation with local noisy operations and one way classical communication

Sayantan Chakraborty
Center for Quantum Technologies
National University of Singapore, Singapore
sayantc@nus.edu.sg

Aditya Nema and Francesco Buscemi
Department of Mathematical Informatics
Nagoya University, Japan
aditya.nema30@gmail.com, buscemi@i.nagoya-u.ac.jp

***Abstract*—We investigate the problem of producing local pure states by performing local noisy operations assisted by one-way classical communication on a given bipartite mixed state in the one-shot setting. We consider the following two scenarios:**

1) **Scenario I: A party, say Alice, is provided with a single copy of some quantum state ρ^A on system A. The task for Alice is to extract pure qubit states using only noisy operations on A. We call this task *purity concentration*.**
2) **Scenario II: Two parties, Alice and Bob possess the A and B sub-systems, respectively, of a given bipartite quantum state ρ^{AB}. They are allowed to perform any local noisy operations and communicate via a one-way dephasing (i.e., classical) channel. The task for them is to design a protocol using these resources such that together they can extract pure local qubit states from the shared state ρ^{AB}. We call this task *local purity distillation*.**

I. Introduction

Pure quantum states are ubiquitous in quantum information theory, both in practical applications (e.g., pure state registers provide a necessary building-block for most quantum algorithms [1]) and in the mathematical analysis of communication protocols, where pure ancillary systems are needed to consider Stinespring–Kraus dilations and model a quantum eavesdropper [1]. A clearer understanding of the role of pure states as a resource in QIT has been achieved within the framework of quantum resource theories of purity, viz. *nonuniformity* [2]–[4]. These are in turn intimately related with the formulation of quantum thermodynamics as a theory of statistical comparison [4]–[13], where pure states are naturally understood as states of full knowledge that can be used when discussing the notion of "work" in an information-theoretic language. This is, for example, the language in which Landauer's principle [14] is usually presented nowadays, see e.g. [8], [9].

Devetak [15] was the first to provide the quantum resource of "purity" with a full-fledged information-theoretic analysis. Devetak's analysis considers the i.i.d. asymptotic scenario and crucially relies on the method of types for the construction of a coding protocol. However, the theory of quantum thermodynamics is typically formulated in a one-shot setting. This makes a direct comparison between Devetak's results and analogous results, such as those in Refs. [4], [13], difficult. From this viewpoint, it is desirable to have a generalization of Devetak's work to the one-shot scenario, but various technical hurdles have impeded this project.

Crucially, the generalization of Devetak's results to the one-shot scenario requires a complete overhaul of the coding techniques. This is very different from what happens for other one-shot protocols, where the difficulty mostly lies in finding the appropriate entropic quantities characterizing the optimal rates, but the coding techniques remain otherwise the same as those used in the i.i.d. asymptotic setting. For example, when devising our one-shot purity distillation algorithm, a crucial step is the construction of the rank-one POVM needed in the two-party protocol. The choice of this POVM however, does not follow easily from the i.i.d. case, but it is nonetheless essential in determining the rate of classical communication. Our construction instead is inspired from a very recent one-shot version of Winter's measurement compression [16]–[18] and further some derandomization arguments. Further problems arise later in the protocol, when the parties use 2-universal hashing, with certain additional regularity conditions about the hash function. That such conditions are satisfied is easy to prove in the asymptotic i.i.d regime based on typicality arguments, which however completely fail in the one-shot regime. To get around this issue, we devise a new proof technique based on random permutations.

The ideas that we present in this work not only provide a way to circumvent the issues mentioned above, but also demonstrate an entirely new approach to designing purity distillation protocols, which does not rely on any regularity assumptions on the underlying resource state. We envision that these insights may also be helpful in other areas of QIT.

II. Preliminaries

Here we give only the definitions necessary to present the main results in Section III. In what follows, all Hilbert spaces, denoted $\mathscr{H}$, are assumed to be finite dimensional, so that the space of linear operators on $\mathscr{H}$, denoted $\mathsf{L}(\mathscr{H})$, is essentially the set of matrices acting on $\mathbb{C}^d$ with $d = \dim \mathscr{H}$. Logarithms are taken in base 2.

Definition II.1 (noisy operations [2], [4]): A completely positive trace-preserving (CPTP) linear map $\mathcal{E} : \mathsf{L}(\mathscr{H}_A) \to \mathsf{L}(\mathscr{H}_B)$ is said to be a *noisy operation* iff there exist

1) finite-dimensional Hilbert spaces $\mathscr{H}_X$ and $\mathscr{H}_Y$ with $\mathscr{H}_A \otimes \mathscr{H}_X \cong \mathscr{H}_{A'} \otimes \mathscr{H}_Y$;
2) a unitary operator $U : \mathscr{H}_A \otimes \mathscr{H}_X \to \mathscr{H}_{A'} \otimes \mathscr{H}_Y$

such that

$$\mathcal{E}(\bullet_A) = \mathrm{Tr}_Y\big[U(\bullet_A \otimes u_X)U^\dagger\big] \text{ , where } u_X = \frac{\mathbb{1}_X}{d_X} \ .$$

Unitary transformations $U : \mathscr{H}_A \to \mathscr{H}_A$ and dephasing channels $\mathcal{D}(\bullet) = \sum_i |\psi_i\rangle\langle\psi_i| \bullet |\psi_i\rangle\langle\psi_i|$, for $\{\psi_i\}_i$ an arbitrary orthonormal basis, both fall within the class of noisy operations: the former can be realized with a one-dimensional ancillary system, while the latter can be realized as uniform random mixtures of Weyl-Heisenberg unitary operators and are therefore noisy operations. Moreover, the composition of a finite number of noisy operations is also a noisy operation.

Definition II.2 (bipartite case with one-way cc): A CPTP linear map $\mathcal{E}^{AB\to A'B'} : \mathsf{L}(\mathscr{H}_A \otimes \mathscr{H}_B) \to \mathsf{L}(\mathscr{H}_{A'} \otimes \mathscr{H}_{B'})$ is said to be a *two-party noisy operation assisted by one-way classical communication* (or *one-way noisy operation*) whenever there exist

1) a noisy operation $\mathcal{N} : \mathsf{L}(\mathscr{H}_A) \to \mathsf{L}(\mathscr{H}_{A'} \otimes \mathscr{H}_C)$;
2) a dephasing channel $\mathcal{D} : \mathsf{L}(\mathscr{H}_C) \to \mathsf{L}(\mathscr{H}_C)$, representing the classical communication;
3) a noisy operation $\mathcal{N}' : \mathsf{L}(\mathscr{H}_B \otimes \mathscr{H}_C) \to \mathsf{L}(\mathscr{H}_{B'})$

such that

$$\mathcal{E}^{AB\to A'B'}(\bullet_{AB}) = \Big[(\mathcal{N}'_{BC}) \circ (\mathsf{id}_{A'} \otimes \mathcal{D}_C \otimes \mathsf{id}_B) \circ (\mathcal{N}_A)\Big](\bullet_{AB}) \ .$$

Using the above operational scenario, we define the rates of purity distillation (single-party and two-party) as follows.

Definition II.3 (single-party ε-distillable purity): Given a density matrix ρ_A on $\mathscr{H}_A$ and a value $\varepsilon \in [0,1]$, a purity concentration ε-code consists of a d_P-dimensional ancillary system and a noisy operation $\mathcal{N} : \mathsf{L}(\mathscr{H}_P \otimes \mathscr{H}_A) \to \mathsf{L}(\mathscr{H}_{A'})$, such that

$$\|\mathcal{N}(|0\rangle\langle 0|_P \otimes \rho_A) - |0\rangle\langle 0|_{A'}\|_1 \leq \varepsilon \ ,$$

in which case the rate $k := \log d_{A'} - \log d_P$ is said to be ε-achievable. The supremum over all ε-achievable rates is called the ε-purity of ρ and is denoted as $\kappa_\varepsilon(\rho_A)$. Notice also how we allow the protocol to borrow an initial pure ancillary state, whose dimension is however discounted from the final net distillation rate.

In the above definition, the pure state $|0\rangle$ is simply an arbitrarily fixed reference pure state, whose actual identity is immaterial for the protocol as unitary transformations are freely available noisy operations.

Definition II.4 (two-party one-way ε-distillable purity): Given a bipartite density matrix ρ_{AB} on $\mathscr{H}_A \otimes \mathscr{H}_B$ and a value $\varepsilon \in [0,1]$, a one-way purity distillation ε-code consists of a d_P-dimensional ancillary system and a one-way noisy operation $\mathcal{N}^{PA\to B} : \mathsf{L}(\mathscr{H}_P \otimes \mathscr{H}_A \otimes \mathscr{H}_B) \to \mathsf{L}(\mathscr{H}_{A'} \otimes \mathscr{H}_{B'})$ such that

$$\big\|\mathcal{N}^{PA\to B}(|0\rangle\langle 0|_P \otimes \rho_{AB}) - |0\rangle\langle 0|_{A'} \otimes |0\rangle\langle 0|_{B'}\big\|_1 \leq \varepsilon \ ,$$

in which case the rate $k := \log d_{A'} + \log d_{B'} - \log d_P$ is said to be one-way ε-achievable. The supremum over all one-way ε-achievable rates is called the one-way ε-purity of ρ_{AB} and is denoted as $\kappa_\varepsilon^{A\to B}(\rho_{AB})$.

One-shot inner bounds for the task of purity distillation, both in the single-party and two-party cases, can be mathematically characterized in terms of two generalized entropic quantities.

Definition II.5 (smoothed support max-entropy): Let ρ_A be a d-dimensional density matrix on $\mathscr{H}_A$ with eigenvalues $(\lambda_i)_{i=1}^d$, ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. For any arbitrarily fixed value $\varepsilon \in [0,1]$, let $k \equiv k(\rho,\varepsilon)$ be such that $\sum_{i=k+1}^d \lambda_i \leq \varepsilon$. Then, the *$\varepsilon$-smoothed support max-entropy* of ρ_A is defined as

$$\widetilde{H}^\varepsilon_{\max}(A)_\rho := \log k \ .$$

Definition II.6 (hypothesis testing relative entropy and mutual information): Let ρ and σ be two density matrices, and $\varepsilon \in [0,1]$ arbitrarily fixed. The hypothesis testing relative entropy [19] between ρ and σ of order ε is defined as

$$D^\varepsilon_H(\rho\|\sigma) := \sup_{\substack{0\leq\Lambda\leq I:\\ \mathrm{Tr}\,\Lambda\rho\geq 1-\varepsilon}} -\log \mathrm{Tr}\,\Lambda\ \sigma \ ,$$

and the corresponding mutual information for a bipartite density matrix ρ_{AB} is defined as

$$I^\varepsilon_H(A:B)_\rho := D^\varepsilon_H(\rho_{AB}\|\rho_A \otimes \rho_B) \ ,$$

where $\rho_{A(B)}$ are the marginals $\mathrm{Tr}_{B(A)}[\rho_{AB}]$.

III. Results

Theorem III.1 (Purity Concentration): Given a density matrix ρ_A on $\mathscr{H}_A$ and a value $\varepsilon \in [0,1]$, its one-shot ε-purity can be lower bounded as

$$\kappa_{3\sqrt{\varepsilon}}(\rho_A) \geq \log d_A - \widetilde{H}^\varepsilon_{\max}(\rho_A) \ .$$

Theorem III.2 (Purity Distillation): Given a bipartite density matrix ρ_{AB} on $\mathscr{H}_A \otimes \mathscr{H}_B$, its one-way ε-purity can be bounded as

$$\kappa^{A\to B}_{32\sqrt{\varepsilon}}(\rho_{AB}) \geq \log d_A - \widetilde{H}^\varepsilon_{\max}(\rho_A) + \log d_B - \widetilde{H}^\varepsilon_{\max}(\rho_B) + J^{8\sqrt{\varepsilon}}_H(A\to B)_\rho + O(\log\varepsilon) \ ,$$

where $J^\varepsilon_H(A\to B)_\rho := \max_{\Lambda:A\to X} I^\varepsilon_H(X:B)_\sigma$, and maximum is taken over all qc-channels $\Lambda : \mathsf{L}(\mathscr{H}_A) \to \mathsf{L}(\mathscr{H}_X)$ of the form $\Lambda(\bullet_A) = \sum_x \mathrm{Tr}\,\Lambda_x\ \rho_A\, |x\rangle\langle x|_X$, for rank-one POVM elements $\{\Lambda_x\}$, orthonormal label states $\{|x\rangle\}$, and $\sigma_{XBR} := (\Lambda_A \otimes \mathsf{id}_B)(\varphi_{ABR})$, for some purification φ_{ABR} of the state ρ_{AB}.

Moreover, the classical communication needed to implement the two-party protocol can be upper bounded by $I^{\sqrt{\varepsilon}}_{\max}(X;BR)_{\sigma_{XBR}} - O(\log\varepsilon)$.

The above one-shot bounds recover the ones given by Devetak [15] in the asymptotic i.i.d. setting, even though the coding techniques used here are quite different from Devetak's. Another interesting by-product is the appearance of a one-shot generalization of the one-way distillable common randomness, namely, the quantity $J^\varepsilon_H(A\to B)_\rho$: its relations with quantum cryptography and the theory of quantum correlations (including entanglement and nonlocality) remain to be explored.

IV. Purity Concentration

We now give a detailed discussion and proof of Theorem III.1. Recall from Definition II.3 that purity concentration involves some noisy operation on the input state ρ^A and then discarding a part of the output system, such that the register which is left, i.e. A_p, contains a state which is close to a pure state $|0\rangle^{A_p}$. To do this, the main idea is to discard the smallest eigenvalues of ρ^A, which add up to at most ε. The eigenvectors which are left then span a space of dimension $2^{\widetilde{H}^{\varepsilon}_{\max}(A)}$, but are embedded in the larger system A. Let us refer to these eigenvectors as 'good'. This embedding necessarily requires that the eigenvectors be padded with 0's on the extra coordinates which are not required to specify them. Thus, we can relabel each of these good eigenvectors with a vector of dimension $2^{\widetilde{H}^{\varepsilon}_{\max}(A)}$ tensored with the unit vector $|0\rangle$. These unit vectors then necessarily belongs to a space of dimension $|A|/2^{\widetilde{H}^{\varepsilon}_{\max}(A)}$. We make these ideas rigorous below.

We will first require the following fact [15, Lemma 1]:

Fact 1: Consider a vector space $A \cong A_g \otimes A_p$, with $\dim A_g = d_1$ and $\dim A_p = d_2$, a state ρ on A, and a projector Π with rank equal to d_1. If $\mathrm{Tr}[\Pi\rho] \geq 1-\varepsilon$, then there exists a unitary U on A, a (normalized) state $\tilde{\rho}$ on A_g, and a pure state $|0\rangle \in A_p$ such that

$$\left\| U\rho U^\dagger - \tilde{\rho} \otimes |0\rangle\langle 0| \right\|_1 \leq 3\sqrt{\varepsilon} \; .$$

Using Fact 1 above, it is easy now to prove Theorem III.1 mentioned in Section III as follows. Given ρ^A and $\varepsilon \in [0,1]$, let us introduce an ancillary system P such that $d_P = 2^{\widetilde{H}^{\varepsilon}_{\max}(A)_\rho}$. This step is necessary in order to make dimensions (which are integer numbers) factorize nicely, so that the protocol is deterministic (i.e., unitary).

Denote by Π^A the projector onto the support of ρ'^A, namely, the sub-normalized state obtained by zeroing out the smallest eigenvalues of ρ^A which add to less than or equal to ε. Notice that $\mathrm{Tr}[\Pi^A \rho^A] \geq 1-\varepsilon$ and $\mathrm{Tr}[\Pi^A] = d_P$. We now define the extended state $\rho^{AP} := \rho^A \otimes |0\rangle\langle 0|^P$.

Clearly, $\widetilde{H}^{\varepsilon}_{\max}(AP)_\rho = \widetilde{H}^{\varepsilon}_{\max}(A)_\rho$. Analogously, let us define the extended projector $\Pi^{AP} = \Pi^A \otimes |0\rangle\langle 0|^P$. Then, $\mathrm{Tr}[\Pi^{AP}\ \rho^{AP}] = \mathrm{Tr}[\Pi^A\ \rho^A] \geq 1-\varepsilon$.

Thus, by invoking Fact 1, we see that there exists a unitary operator from $A \otimes P$ to $A_g \otimes A_p \cong A \otimes P$, satisfying

$$\left\| U(\rho^A \otimes |0\rangle\langle 0|^P)U^\dagger - \tilde{\rho}^{A_g} \otimes |0\rangle\langle 0|^{A_p} \right\|_1 \leq 3\sqrt{\varepsilon} \; ,$$
$$\Rightarrow \kappa_{3\sqrt{\varepsilon}}(\rho^A \otimes |0\rangle\langle 0|^P) \geq \log d_{A_p} = \log(d_A d_P) - \widetilde{H}^{\varepsilon}_{\max}(A),$$

which reduces to the statement of the theorem once we discount the amount $\log d_P$ of purity that we borrowed.

V. Bipartite purity distillation

A. Overview

We will make use of the purity concentration protocol described in Section IV as a subroutine.

The setup: Alice and Bob share the A and B parts, respectively, of a bipartite state ρ^{AB}. They are allowed to use local unitaries and a one-way dephasing (i.e., classical) channel to communicate. They can also borrow local pure ancilla, but these will be discounted from the final rate. An obvious protocol, requiring no communication, is obtained if Alice and Bob simply enact the concentration protocol locally on A and B systems respectively. In this way, they can extract local pure states at the rate

$$\log d_A - \widetilde{H}^{\varepsilon}_{\max}(A) + \log d_B - \widetilde{H}^{\varepsilon}_{\max}(B) \; . \qquad (1)$$

However, the above is not optimal, as shown in the following example.

Consider the maximally correlated state

$$\rho^{AB} := \frac{1}{2}\,|0\rangle\langle 0|^A \otimes |0\rangle\langle 0|^B + \frac{1}{2}\,|1\rangle\langle 1|^A \otimes |1\rangle\langle 1|^B$$

and set $\varepsilon = \frac{1}{4}$. Clearly, we cannot discard any eigenvalues from the marginals ρ^A and ρ^B, and hence the two concentration protocols on the A and B systems together produce no pure states. However, if Alice were to send the system A to Bob via a dephasing channel with operational elements $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$, then Bob could apply the CNOT gate(unitary $\equiv |0\rangle\langle 0|^A \otimes \mathbb{I}^B + |1\rangle\langle 1|^A \otimes X^B$) where X is the quantum NOT (i.e., Pauli X) operator. This allows Bob to extract one qubit pure state. Thus, this example demonstrates that introducing classical communication between the two parties can lead to strictly better rates.

The key idea in the above example is to leverage the *classical* correlations between the systems A and B. However in general the A and B systems shared by Alice and Bob will share quantum correlations. Alice will thus measure her system using a POVM to create a classical-quantum (cq) state, and then send the contents of the classical register created by this measurement to Bob. The hope is that by doing some measurement on his system, Bob should be able to distinguish among the contents of the classical register. If he is able to do this, then he can appropriately map the contents of the classical register to a pure state $|0\rangle$. However, there are several subtle issues that needs to be addressed.

B. New approach towards one-shot purity distillation

Our approach is different from [15], in the sense that we do not rely on concentration arguments to show that there exists a good compressed POVM $\Gamma(k)$ with a small number of good outcomes, which also preserves the classical correlations. Firstly, we require a one-shot measurement compression theorem, recently proved in [18]. Next, we show in Lemma 1 that there exists at least one sub-normalized POVM, which preserves the classical correlations between the two systems, as measured in terms of the smoothed hypothesis testing mutual information. This step is hard since chain rules, readily available for mutual information, are not known for this quantity.

Next, we extend this sub-POVM to a full rank-one POVM $\Gamma(k)$ by extending the set of outcomes using the eigendecomposition of the POVM element $\Gamma_\perp(k)$. Note that this blows up the set of outcomes to a set which contains at least as many indices as the dimension of the underlying space. This is because of the additional outcomes which together correspond to the bad outcome $\perp$. However, we mitigate this issue by using the fact that all these bad outcomes together have probability at most ε. The key idea is that instead of using the set of indices with the lowest probabilities which add up to ε for the extracting the pure states locally at Alice's end, we instead use the set of *bad* outcomes of our POVM. This allows us to distill local purity at Alice's end at the rate $\log d_A - I^{\varepsilon}_{\max}(X:RB)$, where

the quantity $I^{\varepsilon}_{\max}(X:RB)$ can be bounded from above by $\widetilde{H}^{\varepsilon}_{\max}(A)$, as required.

VI. Technical lemmas

In what follows, for the sake of readability, we will adopt the shorthand notation for which a dot placed between two operators denotes the adjoint map, that is, $X\cdot Y := XYX^{\dagger}$.

Lemma 1 (Choosing a POVM): Given a bipartite state ρ^{AB} and a rank-one POVM $\left\{\Lambda^A_x\right\}$ with outcomes in the set $\mathcal{X}$, consider the post measurement state

$$\rho^{XRB} := \sum_x |x\rangle\langle x|^X \otimes \mathrm{Tr}_A\left[(\Lambda^A_x \otimes \mathbb{I}^{RB})\ |\varphi_\rho\rangle\langle\varphi_\rho|^{ARB}\right]$$

where $|\varphi_\rho\rangle^{ARB}$ is a purification of ρ^{AB}. Then, there exists a rank-one POVM $\left\{\tilde{\Lambda}^A_y\right\}$ with outcomes in the set $\mathcal{Y}$ such that:

1) for any $\varepsilon > 0$, there exists a subset $\mathcal{S} \subset \mathcal{Y}$ such that
$$|\mathcal{S}| \leq 2^{I^{\varepsilon}_{\max}(X:RB)_{\rho^{XRB}}} \text{ and } \Pr_{P_Y}[\mathcal{S}] \geq 1-\varepsilon^{1/4}\ ,$$
where P_Y is induced by $\tilde{\Lambda}$ on $\mathcal{Y}$ upon measuring ρ^A;
2) denoting $\Pi^Y_{\mathcal{S}}$ the projector onto the space spanned by the vectors corresponding to the elements in $\mathcal{S}$ and defining the corresponding projected state as
$$\sigma^{YRB} := \frac{\Pi^Y_{\mathcal{S}}\cdot\tilde{\Lambda}^A(\varphi^{ARB}_\rho)}{\mathrm{Tr}[\Pi_{\mathcal{S}}\ \tilde{\Lambda}(\varphi_\rho)]}\ ,$$
we get, with $\varepsilon' := \varepsilon^{1/8}$,
$$I^{\varepsilon'}_H(Y:B)_\sigma \geq I^{\varepsilon'}_H(X:B)_\rho - O(1) + O(\log(1-\varepsilon^2))\ .$$

Lemma 2 (Dividing the domain): Given the control state

$$\rho^{XB} = \sum_x P_X(x)\,|x\rangle\langle x|^X \otimes \rho^B_x$$

and a value $\varepsilon \in (0,1)$, there exists a bijection $\sigma : \mathcal{X} \to [M]\times[N]$, such that:

1) $M \times N = |\mathcal{X}|$;
2) $\log N < I^{\varepsilon}_H(X:B) + 2\log\varepsilon$;
3) let the state after applying the bijection is given by
$$\sigma^{MNB} := \sum_{m,n} P_{MN}(m,n)\,|m,n\rangle\langle m,n|^{MN} \otimes \rho^B_{mn}\ ;$$
then there exists, for all $m \in [M]$, a POVM $\{\Theta_n(m)\}$ with outcomes labeled by $n \in [N]$, such that
$$\sum_{m,n} P_{MN}(m,n)\left\|\rho_{mn} - \sqrt{\Theta_n(m)}\rho_{mn}\sqrt{\Theta_n(m)}\right\|_1 \leq \varepsilon^{1/4}\ .$$

The following corollary that ensures the reduced state after extraction from X is almost unperturbed.

Corollary 1: Given the state $\sigma^{MNB} = \sum_{m,n} P_{MN}(m,n)\,|m,n\rangle\langle m,n|^{MN} \otimes \rho^B_{mn}$ as in Lemma 2, there exists a unitary W^{MNB} such that both the following conditions hold simultaneously:

$$\left\|\mathrm{Tr}_{MB}\big(W^{MNB}\cdot\sigma^{MNB}\big) - |0\rangle\langle 0|^N\right\|_1 \leq \sqrt{\varepsilon^{1/4}}\ ,$$

and

$$\left\|\mathrm{Tr}_{MN}\big(W^{MNB}\cdot\sigma^{MNB}\big) - \sum_{m,n} P_{MN}(m,n)\rho^B_{mn}\right\|_1 \leq \varepsilon^{1/4}\ .$$

Lemma 3: The following relation holds between $I^{\varepsilon}_{\max}(X:RB)$ and $\widetilde{H}^{(\varepsilon)}_{\max}(A)$:

$$\begin{aligned} &I^{2\varepsilon}_{\max}(X:RB)_{\sigma^{XR}} \\ \leq &\widetilde{H}^{O(\varepsilon^2)}_{\max}(A)_\rho - O(\log\varepsilon) + O(\log\frac{\varepsilon^2}{12})\ . \end{aligned}$$

VII. The two-party purity distillation protocol

We now describe the main purity distillation protocol and prove the achievable one-shot rate as stated in Theorem III.2. The proof consists in the following protocol.

1) Alice and Bob start with the A and B parts of the state ρ^{AB} in their possession respectively. In the first step, Alice applies the rank-one POVM $\left\{\tilde{\Lambda}^A_y\right\}$ given by Lemma 1 on her system A coherently. This means that Alice borrows $\log|\mathcal{Y}|$ amount of ancilla and applies the isometry
$$V_1^{A\to YA} := \sum_y |y\rangle^Y \sqrt{\tilde{\Lambda}_y}^A$$
on the system A. Let $|\varphi_\rho\rangle^{ARB}$ be a purification of ρ^{AB}, the global state is
$$\begin{aligned}&(V_1\otimes\mathbb{I}^{RB})\,|\varphi_\rho\rangle^{ARB}\\ =&\sum_y \sqrt{P_Y(y)}\,|y\rangle^Y\,|\psi_y\rangle^A\,|\phi_y\rangle^{RB}\ .\end{aligned}$$
Notice that $\left\{\tilde{\Lambda}^A_y\right\}$ is the compressed POVM, whereas Devetak's original protocol does not involve measurement compression. Alice then applies the unitary
$$U_1^{AY\to AY} := \sum_y |y\rangle\langle y|^Y \otimes U^{A\to A}_y$$
such that $U_y\,|\psi_y\rangle = |0\rangle$. This step yields $\log d_A$ amount of purity while using $\log|\mathcal{Y}|$ amount of ancilla.
2) Next Alice measures the Y system in the computational basis to create $\tau^{YB} := \sum_y P_Y(y)\,|y\rangle\langle y|^Y \otimes \rho^B_y$. In this case we define ρ^B_y for every $y \in \mathcal{Y}$ as the reduced state on B conditioned on y. Let $\mathcal{S}\subset\mathcal{Y}$ be the set of high probability given by Lemma 1 and let $\Pi^Y_{\mathcal{S}}$ be the projector onto the span of the computational basis vectors corresponding to the elements in $\mathcal{S}$. Then, by Fact 1 there exists a local purity concentration protocol with error at most $O(\varepsilon^{1/8})$ with rate
$$\begin{aligned}\log|Y_2| &\geq \log|\mathcal{Y}| - \log|\mathcal{S}|\\ &\geq \log|\mathcal{Y}| - I^{\varepsilon}_{\max}(X:RB)_{\rho^{XRB}}\end{aligned}$$
The net purity at the end of this step is then
$$\begin{aligned}&\log d_A - \log|\mathcal{Y}| + \log|\mathcal{Y}| - I^{\varepsilon}_{\max}(X:RB)_{\rho^{XRB}}\\ =&\log d_A - I^{\varepsilon}_{\max}(X:RB)_{\rho^{XRB}}\ .\end{aligned}$$

3) Alice and Bob are now left with the state

$$\sigma^{Y_1 B} := \frac{1}{\mathrm{Tr}[\Pi_{\mathcal{S}}\tau]}\left(\mathbb{I}^B \otimes \Pi_{\mathcal{S}}^Y \cdot \tau^{YB}\right)^{Y_1} = \frac{1}{\mathrm{Tr}[\Pi_{\mathcal{S}}\tau]} \sum_{y\in\mathcal{S}} P_Y(y)\, |y\rangle\langle y|^{Y_1} \otimes \rho_y^B \ ,$$

Alice then applies the bijection given by Lemma 2 to create the state

$$\sigma^{MNB} := \sum_{m,n} P_{MN}(m,n)\, |m,n\rangle\langle m,n|^{MN} \otimes \rho_{mn}^B \ ,$$

where $MN = |Y_1| = |\mathcal{S}|$ and

$$\log N \leq I_H^{\varepsilon^{1/8}}(Y_1 : B)_{\sigma^{Y_1 B}} + \log \varepsilon^{1/4} \ .$$

Alice sends the systems MN to Bob through the dephasing channel, which requires at most $\log|\mathcal{S}| \leq I_{\max}^{\varepsilon}(X:RB)_{\rho^{XRB}}$ number of bits. This quantity can be further bounded by $H_{\max}^{O(\varepsilon^2)}(A)$, see Lemma 3.

4) Finally, after receiving the system MN, Bob applies the unitary W^{MNB} given by Corollary 1 such that

$$\left\| \mathrm{Tr}_{MB}\left(W^{MNB} \cdot \sigma^{MNB}\right) - |0\rangle\langle 0|^N \right\|_1 \leq 2\varepsilon^{1/32}$$

to distill $\log N$ amount of purity.

5) Corollary 1 also tells us that the state on system B after Bob applies the unitary W^{MNB} is $2\varepsilon^{1/32}$ away from $\frac{1}{\mathrm{Tr}[\Pi_{\mathcal{S}}\tau]}\sum_{y\in\mathcal{S}} P_Y(y)\rho_y^B$.
However, since $\mathcal{S}$ is a set of high probability under P_Y, this implies that

$$\left\| \sum_y P_Y(y)\rho_y^B - \frac{\sum_{y\in\mathcal{S}} P_Y(y)\rho_y^B}{\mathrm{Tr}[\Pi_{\mathcal{S}}\tau]} \right\|_1 \leq O(\varepsilon^{1/64}) \ ,$$

that is, in the end, Bob has a state which is not far from his original state.
Thus, Bob applies protocol of Theorem III.1 on B to recover $\log d_B - \widetilde{H}_{\max}^{\varepsilon}(B)$ amount of purity with error $O(\varepsilon^{1/32})$. Also, $I_H^{\varepsilon^{1/8}}(Y_1:B)_{\sigma^{Y_1B}} \geq I_H^{\varepsilon^{1/8}}(X:B)_{\rho^{XB}} - O(1)O(\log(1-\varepsilon^{1/4}))$.

Hence, the total amount of purity recovered is:

$$\kappa_{\varepsilon^{1/32}}(\rho_A) \underset{O(1)}{\geq} \log d_A d_B - I_{\max}^{\sqrt{\varepsilon}}(X:RB)_{\rho^{XRB}} - \widetilde{H}_{\max}^{\varepsilon}(B) + I_H^{\varepsilon^{1/8}}(X:B)_{\rho^{XB}} + \log\varepsilon^{1/32}) \quad (2)$$

which by Lemma 3 is further lower bounded by:

$$\log d_A d_B - \widetilde{H}_{\max}^{\varepsilon}(A)_{\rho^A} - \widetilde{H}_{\max}^{\varepsilon}(B) + I_H^{\varepsilon^{1/8}}(X:B)_{\rho^{XB}} + O(\log\varepsilon^{1/32}) - O(1) \ . \quad (3)$$

At each step the protocol makes an additive error of at most $O(\varepsilon^{1/32})$, so that the total error of the protocol is still bounded by $O(\varepsilon^{1/32})$.

ACKNOWLEDGMENT

SC would like to acknowledge support from the National Research Foundation, including under NRF RF Award No. NRF-NRFF2013-13 and NRF2021-QEP2-02-P05 and the Prime Minister's Office, Singapore and the Ministry of Education, Singapore, under the Research Centres of Excellence program. SC would also like to acknowledge support from the Google Late PhD Fellowship grant.

AN and FB acknowledge support from MEXT Quantum Leap Flagship Program (MEXT QLEAP) Grant No. JPMXS0120319794. FB acknowledges support also from MEXT-JSPS Grant-in-Aid for Transformative Research Areas (A) "Extreme Universe", No. 21H05183, and from JSPS KAKENHI Grants No. 20K03746 and No. 23K03230.

REFERENCES

[1] M. M. Wilde, *Quantum Information Theory, 2nd edition*. Cambridge University Press, 2017.

[2] M. Horodecki, K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), and U. Sen, "Local information as a resource in distributed quantum systems," *Phys. Rev. Lett.*, vol. 90, p. 100402, Mar 2003. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.90.100402

[3] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, "Thermodynamical approach to quantifying quantum correlations," *Physical Review Letters*, vol. 89, no. 18, oct 2002. [Online]. Available: https://doi.org/10.1103$\%$2Fphysrevlett.89.180402

[4] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Yunger Halpern, "The resource theory of informational nonequilibrium in thermodynamics," *Physics Reports*, vol. 583, pp. 1–58, 2015, the resource theory of informational nonequilibrium in thermodynamics. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037015731500229X

[5] M. Horodecki and J. Oppenheim, "Fundamental limitations for quantum and nanoscale thermodynamics," *Nature Communications*, vol. 4, p. 2059, 2013.

[6] F. G. Brandão, M. Horodecki, J. Oppenheim, J. Renes, and R. W. Spekkens, "The resource theory of quantum states out of thermal equilibrium," *arXiv:1111.3882*, 2011.

[7] F. G. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," *Proceedings of the National Academy of Sciences*, vol. 112, no. 11, pp. 3275–3279, feb 2015.

[8] D. Reeb and M. M. Wolf, "An improved Landauer principle with finite-size corrections," *New Journal of Physics*, vol. 16, no. 10, p. 103011, 2014.

[9] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, "The Minimal Work Cost of Information Processing," *Nat. Comm.*, vol. 6, p. 7669, jul 2015.

[10] F. Buscemi, "Fully quantum second-law–like statements from the theory of statistical comparisons," 2015. [Online]. Available: https://arxiv.org/abs/1505.00535

[11] F. Buscemi and G. Gour, "Quantum relative lorenz curves," *Physical Review A*, vol. 95, no. 1, jan 2017. [Online]. Available: https://doi.org/10.1103%2Fphysreva.95.012110

[12] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, "Quantum majorization and a complete set of entropic conditions for quantum thermodynamics," *Nature Communications*, vol. 9, no. 1, dec 2018. [Online]. Available: https://doi.org/10.1038%2Fs41467-018-06261-7

[13] B. Morris, L. Lami, and G. Adesso, "Assisted work distillation," *Physical Review Letters*, vol. 122, no. 13, apr 2019. [Online]. Available: https://doi.org/10.1103%2Fphysrevlett.122.130601

[14] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM Journal of Research and Development*, vol. 5, no. 3, pp. 183–191, 1961.

[15] I. Devetak, "Distillation of local purity from quantum states," *Phys. Rev. A*, vol. 71, p. 062303, Jun 2005. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.71.062303

[16] A. Winter, "Extrinsic and intrinsic data in quantum measurements: Asymptotic convex decomposition of positive operator valued measures," *Communications in Mathematical Physics*, vol. 244, no. 1, pp. 157–185, jan 2004. [Online]. Available: https://doi.org/10.1007%2Fs00220-003-0989-z

[17] M. M. Wilde, P. Hayden, F. Buscemi, and M.-H. Hsieh, "The information-theoretic costs of simulating quantum measurements," *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 45, p. 453001, oct 2012. [Online]. Available: https://doi.org/10.1088$\%$2F1751-8113$\%$2F45$\%$2F45$\%$2F453001

[18] S. Chakraborty, A. Padakandla, and P. Sen, "Centralised multi link measurement compression with side information," 2022. [Online]. Available: https://arxiv.org/abs/2203.16157

[19] F. Buscemi and N. Datta, "The quantum capacity of channels with arbitrarily correlated noise," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1447–1460, mar 2010. [Online]. Available: https://doi.org/10.1109\%2Ftit.2009.2039166