Moderate Deviation Analysis for Quantum State Transfer

Navneeth Ramakrishnan*†, Marco Tomamichel†‡ and Mario Berta*§¶
*Department of Computing, Imperial College London, nr2618@ic.ac.uk
†Centre for Quantum Technologies, National University of Singapore, Singapore
‡Department of Electrical and Computer Engineering, National University of Singapore, Singapore
§IQIM, California Institute of Technology, Pasadena, USA
¶AWS Center for Quantum Computing, Pasadena, USA

2021 IEEE Information Theory Workshop (ITW) | 978-1-6654-0312-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ITW48936.2021.9611459

***Abstract*—Quantum state transfer involves two parties who use pre-shared entanglement and noiseless communication in order to transfer parts of a quantum state. In this work, we quantify the communication cost of one-shot state splitting in terms of the partially smoothed max-information. We then give an analysis of state splitting in the moderate deviation regime, where the error in the protocol goes sub-exponentially fast to zero as a function of the number of i.i.d. copies. The main technical tool we derive is a tight relation between the partially smoothed max-information and the hypothesis testing relative entropy, which allows us to obtain the expansion of the partially smoothed max-information for i.i.d. states in the moderate deviation regime. This then also establishes the moderate deviation analysis for other variants of state transfer such as state merging and source coding.**

I. Introduction and motivation

Many fundamental tasks in quantum information theory can be accomplished using the primitive of transferring a quantum state from one party to another. The setting we consider involves two parties who have access to pre-shared entanglement and one-way communication. At the end of the primitive, one part of a quantum state must then be transferred from one party to the other. Note that this scenario covers in particular state splitting, state merging, and source coding [1].

Employing the convex-split lemma based protocol from [2], we first show that the communication cost of ε-error one-shot state splitting can be given in terms of the partially smoothed max-information. Having understood the one-shot results, we consider the i.i.d. setting of transferring a state of the form $\rho^{\otimes n}$ in the asymptotic limit $n \to \infty$ and vanishing error $\varepsilon \to 0$. The asymptotic equipartition property of the smoothed max-information then recovers a communication cost in terms of the mutual information. While this result is a useful theoretical bound, practical settings use finitely many i.i.d. copies of the state. It is therefore beneficial to also understand the quantitative trade-off between the error in a state transfer protocol and the communication cost as a function of the number of copies available. Following [3], a moderate sequence is defined as a sequence of positive numbers $\{a_n\}$ such that $\lim_{n\to\infty} a_n = 0$ and $\lim_{n\to\infty} a_n\sqrt{n} = \infty$. In the moderate deviation regime, the error in the protocol is chosen to be $\varepsilon_n = e^{-a_n^2 n}$. Our main result shows that the communication cost in this regime is given in terms of the mutual information and a correction term that depends on the mutual information variance with a pre-factor a_n.

Our state splitting results already cover other settings as well. In particular, it was shown in [4] that the communication cost of ε-error one-shot state merging is given by the same partially smoothed max-information. Our moderate deviation analysis therefore holds for state merging as well. We also investigate a special case of state transfer where one of the registers is trivial and obtain the moderate deviation expansion for source coding.

II. Notation

We use standard quantum information notation and refer the reader to the references herein for the detailed definitions of the following quantities.

The exponential and logarithm functions are taken with respect to an arbitrary base, where the same base is used to define communication costs. The set of positive semidefinite operators with $\mathrm{Tr}\,(\rho_A) \leq 1$ is denoted by $\mathcal{S}_{\leq}(A)$ and the set of positive semi-definite operators with unit trace is denoted by $\mathcal{S}(A)$. The purified distance between states $\rho, \sigma \in \mathcal{S}_{\leq}(H)$ is given by $P(\rho,\sigma) = \sqrt{1 - \bar{F}(\rho,\sigma)^2}$, where $\bar{F}(\rho,\sigma)$ is the generalized fidelity defined as $\bar{F}(\rho,\sigma) = \|\sqrt{\rho \oplus (1-\mathrm{Tr}\,(\rho))}\sqrt{\sigma \oplus (1-\mathrm{Tr}\,(\sigma))}\|_1$ [5]. Note that if either ρ or σ have unit trace, then the generalized fidelity is equal to the standard fidelity $F(\rho,\sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$. If $P(\rho,\sigma) \leq \varepsilon$, we denote this by $\rho \approx_\varepsilon \sigma$.

The von Neumann entropy is denoted by $S(\rho)$, the relative entropy by $D(\rho\|\sigma)$, and the relative entropy variance by $V(\rho\|\sigma)$ [6], [7]. The sandwiched Renyi relative entropy is denoted by $\widetilde{D}_\alpha(\rho\|\sigma)$ [8], [9] and the generalized sandwiched Renyi mutual information is denoted by $\widetilde{I}_\alpha(\rho_{AB}\|\tau_A)$ [10]. The special cases of α that we are interested in this work are $\alpha = 1$ to obtain the quantum relative entropy $D(\rho\|\sigma)$ and mutual information $I(A:B)_\rho$, $\alpha = \frac{1}{2}$ to obtain the min-relative entropy $D_{\min}(\rho\|\sigma)$ and min-information $I_{\min}(A;B)_\rho$, and $\alpha = \infty$ to obtain the max-relative entropy $D_{\max}(\rho\|\sigma)$ and max-information $I_{\max}(A;B)_\rho$. One defines smooth entropy measures by extremizing aforementioned quantities over a set of sub-normalized states in an ε neighbourhood in terms of the purified distance, denoted by $\mathcal{B}^\varepsilon(\rho)$ [11]. We have the smooth max- and min-relative entropies given

978-1-6654-0312-2/21/$31.00 ©2021 IEEE

by $D^{\varepsilon}_{\max}(\rho_A\|\sigma_A)$ and $D^{\varepsilon}_{\min}(\rho_A\|\sigma_A)$ respectively [12]. The smoothed max-information is denoted by $I^{\varepsilon}_{\max}(A;B)_\rho$ [13] and the partially smoothed max-information is denoted by $I^{\varepsilon}_{\max}(\dot{A};B)_\rho$ [4]. The hypothesis testing relative entropy is denoted by $D^{\varepsilon}_h(\rho\|\sigma)$ [14]. Finally, the information spectrum relative entropy is denoted by $\underline{D}^{\varepsilon}_s(\rho\|\sigma)$ and the information spectrum entropy by $\bar{H}^{\varepsilon}_s(\rho)$ [15].

III. One-shot state splitting

In this section, we introduce one-shot quantum state splitting and the convex-split lemma from [2].

Definition III.1 (One-shot state splitting)**.** *Consider a pure state $\rho_{AA_1R} \in \mathcal{S}(AA_1R)$, where R denotes a reference system and Alice holds A and A_1. A state splitting protocol starts with some pre-shared pure state $\sigma_{K'K} \in \mathcal{S}(K'K)$, where Alice has access to K' and Bob has access to K. Then, a CPTP map $\mathcal{T} = \mathcal{D}\circ\mathcal{E}$ is applied that consists of a local encoding map $\mathcal{E}$ on Alice's registers, one round of communication from Alice to Bob, and a local decoding map $\mathcal{D}$ on Bob's registers. For $\varepsilon \in [0,1]$, we call $\mathcal{T}$ a $\{q,\varepsilon\}-$one-shot quantum state splitting protocol of ρ_{AA_1R} when the quantum communication cost is q and there exists $\sigma_{K'K}$ such that*[1]

$$(\mathcal{T}\otimes\mathcal{I}_R)\left(\rho_{AA_1R}\otimes\sigma_{K'K}\right) \approx_\varepsilon \rho_{ABR}, \tag{1}$$

where $\rho_{ABR} = (\mathcal{I}_{AR}\otimes\mathcal{I}_{A_1\to B})(\rho_{AA_1R})$ with Bob holding the B register. The minimal communication cost is defined as $q^\star_\varepsilon(\rho_{AA_1R}) = \min\{q \in \mathbb{N} : \exists\ \{q,\varepsilon\} -$ one-shot state splitting protocol of $\rho_{AA_1R}\}$.

We show the following characterization.

Theorem III.2. *For $\varepsilon \in (0,1]$ with $\delta \in (0,\varepsilon)$, the minimal communication cost of ε-error state splitting protocol of a pure state $\rho_{AA_1R} \in \mathcal{S}(AA_1R)$ is bounded as*

$$\frac{1}{2}I^{\varepsilon}_{\max}(\dot{R};A_1)_\rho \le q^\star_\varepsilon(\rho_{AA_1R}) \le \frac{1}{2}I^{\varepsilon-\delta}_{\max}(\dot{R};A_1)_\rho + \log\frac{1}{\delta}. \tag{2}$$

To prove this result, we start with the convex-split lemma.

Lemma III.3 (Convex-split lemma [2])**.** *For any finite non-empty set Σ, define the Hilbert space $A_\Sigma = \bigotimes_{s_i\in\Sigma} A_{s_i}$, where all A_{s_i} are Hilbert spaces of the same dimension. An extended state is a state $\rho_{A_\Sigma} = \rho_{A_{s_1}}\otimes\rho_{A_{s_2}}\otimes\ldots\otimes\rho_{A_{s_{|\Sigma|}}} \in \mathcal{S}_\le(A_\Sigma)$. Now, for $I = \{1,2,\ldots\ n\}$ consider the Hilbert space B_I and a Hilbert space R. Then, for $i \in I$, let $\rho_{B_iR} \in \mathcal{S}(B_iR)$ and $\sigma_{B_i} \in \mathcal{S}(B_i)$ and define the extended state $\tau_{B_IR} = \frac{1}{n}\sum_{i=1}^{n}\rho_{B_iR}\otimes\sigma_{B_{I\setminus i}}$. For $\delta > 0$ and n satisfying $\log n \ge D_{\max}(\rho_{B_iR}||\sigma_{B_i}\otimes\rho_R) + \log\frac{1}{\delta}$, it holds that*

$$F(\tau_{B_IR},\sigma_{B_I}\otimes\rho_R) \ge \frac{1}{\sqrt{(1+\delta)}}. \tag{3}$$

Next, we show the achievability bound in Theorem III.2. The following proposition closely follows the ideas in [2], with the sole difference that we employ the partially smoothed max-information instead of the smoothed max-information.

Proposition III.4 (Achievability bound for state splitting)**.** *Let $\rho_{AA_1R} \in \mathcal{S}(AA_1R)$ be a pure state, where R is inaccessible and Alice holds registers A, A_1. For $\varepsilon \in (0,1]$ with $\delta \in (0,\varepsilon)$ it holds that*

$$q^\star_\varepsilon(\rho_{AA_1R}) \le \frac{1}{2}I^{\varepsilon-\delta}_{\max}(\dot{R};A_1)_{\rho_{A_1R}} + \log\frac{1}{\delta}. \tag{4}$$

Proof. Let $I^{\varepsilon-\delta}_{\max}(\dot{R} : A_1)_{\rho_{A_1R}} = D_{\max}(\rho'_{A_1R}\|\sigma_{A_1}\otimes\rho'_R)$, where $\rho'_{A_1R} \in \mathcal{B}^{\varepsilon-\delta}(\rho_{A_1R})$, $\rho'_R = \rho_R$, and $\sigma_{A_1} \in \mathcal{S}(A_1)$. For $I = \{1,2\ldots\ n\}$, let $\sigma_{(K'K)_I} \in \mathcal{S}((K'K)_I)$ be an extended pure state pre-shared between Alice ($K'_{i\in I}$) and Bob ($K_{i\in I}$) with $K_i \cong A_1$. The joint state at the start of the protocol is

$$|\psi\rangle = |\rho_{AA_1R}\rangle\otimes|\sigma_{(K'K)_I}\rangle \tag{5}$$

$$\approx_{\varepsilon-\delta} |\rho'_{AA_1R}\rangle\otimes|\sigma_{(K'K)_I}\rangle = |\omega\rangle, \tag{6}$$

where $|\rho'_{AA_1R}\rangle$ is the purification of ρ'_{A_1R} on Alice's register A that achieves the minimal purified distance from the purification $|\rho_{AA_1R}\rangle$ of ρ_{AR}. Uhlmann's theorem guarantees that such a purification exists. The reduced state of $|\omega\rangle$ after tracing over Alice's registers A, A_1 and K'_I is $\rho'_R\otimes\sigma_{K_I}$ and these registers are with Bob and the inaccessible reference system. By Lemma III.3, we have that for $\delta \in (0,1]$ and $n \in \mathbb{N}$ satisfying

$$\log n \ge D_{\max}(\rho'_{A_1R}||\sigma_{A_1}\otimes\rho'_R) + \log\frac{1}{\delta^2}, \tag{7}$$

it holds that

$$F\left(\rho'_R\otimes\sigma_{K_I}, \frac{1}{n}\sum_{i=1}^{n}\rho'_{RK_i}\otimes\sigma_{K_{I\setminus i}}\right) \ge \frac{1}{\sqrt{(1+\delta^2)}}. \tag{8}$$

A purification of $\frac{1}{n}\sum_{i=1}^{n}\rho'_{RK_i}\otimes\sigma_{K_{I\setminus i}}$ is

$$|\tau\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}|i\rangle_X\otimes|\rho'_{ARK_i}\rangle\otimes|\sigma_{(K'K)_{I\setminus i}}\rangle, \tag{9}$$

where Alice holds classical register X. If $\log n$ satisfies (7), there exists an isometry U on Alice's registers such that $F(U(|\omega\rangle),|\tau\rangle) \ge \frac{1}{\sqrt{1+\delta^2}}$ and hence $P(U(|\omega\rangle),|\tau\rangle) \le \delta$. Moreover, since $\rho'_R = \rho_R$, there exists an isometry V on AA_1 such that $V|\psi\rangle = |\omega\rangle$. Both these isometries exist due to Uhlmann's theorem. If the joint state was $|\tau\rangle$, the following protocol achieves state splitting:

1) Alice measures the X register and sends the outcome i to Bob. Using superdense coding this has a quantum communication cost of $\frac{1}{2}\log n$.
2) Bob applies swap operations $K_i \leftrightarrow K_1$ to obtain $|\rho'_{ARK_1}\rangle$ which we relabel as $|\rho'_{ABR}\rangle$ and discards all other registers.

Let us denote the protocol above by $\mathcal{T}$. Then, since $UV|\psi\rangle \approx_\delta |\tau\rangle$ and $\mathcal{T}(|\tau\rangle) \approx_{\varepsilon-\delta} \rho_{ABR}$, we can use the triangle inequality and data processing for the purified distance to conclude that $\mathcal{T}(UV|\psi\rangle) \approx_\varepsilon \rho_{ABR}$. This $\varepsilon-$error one-shot quantum

[1] Since free entanglement is available, we equivalently have that the classical communication required is $2q$.

state splitting protocol then has a communication cost of $q^\star_\varepsilon(\rho_{AA_1R}) \leq \frac{1}{2}\log n$, which concludes the proof. □

A converse result for state splitting is now shown following [13], but with a tightened bound featuring the partially smoothed max-information.

Proposition III.5 (Converse bounds for state splitting)**.** *Consider a pure state $\rho_{AA_1R} \in \mathcal{S}(A_1R)$, where R is inaccessible and Alice holds registers A, A_1. The minimal communication cost of $\varepsilon-$error one-shot state splitting for ρ_{AA_1R} satisfies*

$$q^\star_\varepsilon(\rho_{AA_1R}) \geq \frac{1}{2} I^\varepsilon_{\max}(\dot{R}; A_1)_\rho. \tag{10}$$

Proof. Any state splitting protocol consists of the state $\rho_{AA_1R} \otimes \sigma_{K'K}$, where A, A_1, K' are with Alice, R is inaccessible and K is with Bob. The state $\sigma_{K'K}$ is some initial resource state. By the definition of state splitting, we have a local operation $\mathcal{E} : AA_1K' \to AQ$ on Alice's registers, a communication cost of $q = \log|Q|$ from Alice to Bob and local operation $\mathcal{D} : QK \to B$ on Bob's registers. Let $\omega_{AQKR} = (\mathcal{E} \otimes \mathcal{I}_{RK})(\rho_{AA_1R} \otimes \sigma_{K'K})$ and $\tau_{ABR} = (\mathcal{D} \otimes \mathcal{I}_{AR})(\omega_{AQKR}) \approx_\varepsilon \rho_{ABR}$. Note that $\tau_R = \rho_R$ and thus

$$\begin{aligned} I^\varepsilon_{\max}(\dot{R}; B)_\rho &\leq I_{\max}(R; B)_\tau & (11)\\ &\leq I_{\max}(R; QK)_\omega & (12)\\ &\leq I_{\max}(R; K)_\omega + 2\log|Q| & (13)\\ &\leq I_{\max}(R; K)_{\rho\otimes\sigma} + 2\log|Q|, & (14)\end{aligned}$$

where we have used the data processing inequality on the max-information and the non-lockability of the max-information [16, Corollary A.14]. Since $I_{\max}(R; K)_{\rho\otimes\sigma} = 0$ as Bob is not correlated with the reference at the start of the protocol, the result follows. □

IV. Moderate deviation analysis

We can extend the results of the one-shot protocol of Section III to recover known results for asymptotic state splitting of i.i.d. states. By [4, Theorem 2], one can bound the partially smoothed max-information with the smoothed max-information. The asymptotic equipartition property of the smoothed max-information [13, Corollary B.22] then gives that the minimal communication cost $q^\star_\varepsilon(\rho^{\otimes n}_{AA_1R})$ of $\varepsilon-$error one-shot quantum state splitting for $\rho^{\otimes n}_{AA_1R}$ obeys

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} \frac{q^\star_\varepsilon(\rho^{\otimes n}_{AA_1R})}{n} = \frac{1}{2} I(R : A_1)_\rho. \tag{15}$$

In the following, we investigate the minimal communication cost of one-shot state transfer protocols for i.i.d. states in the moderate deviation regime. That is, for $\varepsilon_n = e^{-a_n^2 n}$ for some moderate sequence $\{a_n\}$, we look at ε_n-one-shot state transfer protocols of $\rho^{\otimes n}_{AA_1R}$ and investigate how the minimal $q^\star_{\varepsilon_n}(\rho^{\otimes n}_{AA_1R})$ behaves as a function of n. The main idea here is to obtain an expansion for the partially smoothed max-information of i.i.d. states (Proposition IV.7). This is done by bounding the partially smoothed max-information with the hypothesis testing relative entropy and then using the expansion of the hypothesis testing relative entropy in the moderate deviation regime (Proposition IV.6). Our main result is the following theorem.

Theorem IV.1. *The minimal communication cost of ε_n-error state splitting of $\rho^{\otimes n}_{AA_1R}$ for $\rho_{AA_1R} \in \mathcal{S}(AA_1R)$ pure and $\varepsilon_n = e^{-na_n^2}$ for a moderate sequence a_n is*

$$\begin{aligned}\frac{1}{n} q^\star_{\varepsilon_n}(\rho^{\otimes n}_{AA_1R}) = \frac{1}{2} I(R : A_1)_\rho + a_n\sqrt{V(\rho_{A_1R}\|\rho_{A_1}\otimes\rho_R)}\\ + o(a_n). \end{aligned}\tag{16}$$

The following lemma was already known for the conditional entropy [5, Remark 5.6]. The tight bounds on the triangle inequality for the purified distance in that remark are necessary for us and we clarify when exactly the tighter bounds hold.

Lemma IV.2. *For states $\rho, \sigma, \tau \in \mathcal{S}(H)$ with $P(\rho,\sigma)^2 + P(\sigma,\tau)^2 \leq 1$ we have*

$$P(\rho,\tau) \leq P(\rho,\sigma)F(\sigma,\tau) + P(\sigma,\tau)F(\rho,\sigma). \tag{17}$$

Proof. A careful reading of the proof of [14, Proposition 3.16] shows that this tighter triangle inequality holds when $\sin^{-1}(P(\rho,\sigma)) + \sin^{-1}(P(\sigma,\tau)) \leq \pi/2$. We show that this condition is equivalent to $P(\rho,\sigma)^2 + P(\sigma,\tau)^2 \leq 1$. By the monotonicity of the $\cos(\cdot)$ function on $[0,\pi]$, we have

$$\sin^{-1}(P(\rho,\sigma)) + \sin^{-1}(P(\sigma,\tau)) \leq \frac{\pi}{2} \tag{18}$$

$$\iff \cos(\sin^{-1}(P(\rho,\sigma)) + \sin^{-1}(P(\sigma,\tau))) \geq \cos\left(\frac{\pi}{2}\right) \tag{19}$$

$$\iff \sqrt{1-P(\rho,\sigma)^2}\sqrt{1-P(\sigma,\tau)^2} - P(\rho,\sigma)P(\sigma,\tau) \geq 0 \tag{20}$$

$$\iff P(\rho,\sigma)^2 + P(\sigma,\tau)^2 \leq 1, \tag{21}$$

where the third line follows by using the addition formula for cosine and using that $\cos^2(\theta) = 1 - \sin^2(\theta)$. □

Lemma IV.3. *For $\varepsilon, \varepsilon' \in [0,1]$ with $\varepsilon^2 + \varepsilon'^2 \leq 1$, the function $\varepsilon\sqrt{1-\varepsilon'^2} + \varepsilon'\sqrt{1-\varepsilon^2}$ is monotonically increasing in both ε and ε'.*

Proof. From the proof of Lemma IV.2, we have that $\sin^{-1}(\varepsilon) + \sin^{-1}(\varepsilon') \leq \pi/2$. Since the $\sin(\cdot)$ function is monotone increasing in $[0,\pi/2]$ and $\varepsilon\sqrt{1-\varepsilon'^2} + \varepsilon'\sqrt{1-\varepsilon^2} = \sin(\sin^{-1}(\varepsilon) + \sin^{-1}(\varepsilon'))$, the result follows. □

Lemma IV.4. *For $\rho \in \mathcal{S}(H)$, $\sigma \in \mathcal{P}(H)$, and $\varepsilon, \varepsilon' \in [0,1]$ with $\varepsilon^2 + \varepsilon'^2 \leq 1$ we have that*

$$\begin{aligned} D^\varepsilon_{\min}(\rho\|\sigma) \leq D^{\varepsilon'}_{\max}(\rho\|\sigma)\\ - \log\left(1 - \left(\varepsilon\sqrt{1-\varepsilon'^2} + \varepsilon'\sqrt{1-\varepsilon^2}\right)^2\right). \end{aligned}\tag{22}$$

Proof. Let $\lambda = D^{\varepsilon'}_{\max}(\rho\|\sigma) = D_{\max}(\tilde{\rho}\|\sigma)$ for $\tilde{\rho} \approx_{\varepsilon'} \rho$. It then holds that $\tilde{\rho} \leq 2^{\lambda}\sigma$. We also have that for some state $\bar{\rho} \approx_{\varepsilon} \rho$

$$D^{\varepsilon}_{\min}(\rho\|\sigma) = D_{\min}(\bar{\rho}\|\sigma) = -\log F(\bar{\rho},\sigma)^2 \tag{23}$$
$$\leq -\log F(\bar{\rho}, 2^{-\lambda}\tilde{\rho})^2 \tag{24}$$
$$= \lambda - \log F(\bar{\rho},\tilde{\rho})^2 \tag{25}$$
$$\leq D^{\varepsilon'}_{\max}(\rho\|\sigma) - \log(1 - P(\bar{\rho},\tilde{\rho})^2) \tag{26}$$
$$\leq D^{\varepsilon'}_{\max}(\rho\|\sigma) - \log\left(1 - \left(\varepsilon\sqrt{1-\varepsilon'^2} + \varepsilon'\sqrt{1-\varepsilon^2}\right)^2\right), \tag{27}$$

where the last step is via Lemma IV.2 and Lemma IV.3. □

Lemma IV.5. *For $\rho,\sigma \in \mathcal{P}(H)$ and $\varepsilon \in (0, \frac{1}{\sqrt{2}})$, we have*

$$D^{\varepsilon}_{\min}(\rho\|\sigma) \leq D^{2\varepsilon^2}_{h}(\rho\|\sigma) - \log f(\varepsilon), \tag{28}$$

where $f(\varepsilon) = \left(1 - \left(\varepsilon^2\sqrt{2} + \sqrt{1-2\varepsilon^2}\sqrt{1-\varepsilon^2}\right)^2\right)$

Proof. We choose $\varepsilon' = \sqrt{1-2\varepsilon^2}$ in Lemma IV.4 to obtain

$$D^{\varepsilon}_{\min}(\rho\|\sigma) \leq D^{\sqrt{1-2\varepsilon^2}}_{\max}(\rho\|\sigma) - \log f(\varepsilon). \tag{29}$$

Note that choice of ε' would not be possible without the tighter triangle inequality. By [17, Theorem 4], we have for $\delta \in (0,1)$

$$D^{\sqrt{\delta}}_{\max}(\rho\|\sigma) \leq D^{1-\delta}_{h}(\rho\|\sigma) - \log(1-\delta). \tag{30}$$

Combining the two results, we have

$$D^{\varepsilon}_{\min}(\rho\|\sigma) \leq D^{\sqrt{1-2\varepsilon^2}}_{\max}(\rho\|\sigma) - \log f(\varepsilon) \tag{31}$$
$$\leq D^{1-(\sqrt{1-2\varepsilon^2})^2}_{h}(\rho\|\sigma) - \log f(\varepsilon) - \log(2\varepsilon^2) \tag{32}$$
$$\leq D^{2\varepsilon^2}_{h}(\rho\|\sigma) - \log(f(\varepsilon)) - \log 2\varepsilon^2. \tag{33}$$

□

Next, we recall the moderate deviation expansion for the hypothesis testing relative entropy with i.i.d. states [3], [18].

Proposition IV.6. *For any moderate sequence $\{a_n\}$, let $\varepsilon_n = e^{-a_n^2 n}$. For states $\rho \ll \sigma$, the hypothesis testing relative entropy scales as*

$$\frac{1}{n}D^{\varepsilon_n}_{\mathrm{h}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = D(\rho\|\sigma) - \sqrt{2V(\rho\|\sigma)}a_n + o\left(a_n\right). \tag{34}$$

$$\frac{1}{n}D^{1-\varepsilon_n}_{\mathrm{h}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = D(\rho\|\sigma) + \sqrt{2V(\rho\|\sigma)}a_n + o\left(a_n\right). \tag{35}$$

We now have the main technical result needed for our moderate deviation analysis — an expansion for the partially smoothed max-information of i.i.d. states in the moderate deviation regime.

Proposition IV.7. *Let $\varepsilon_n = e^{-na_n^2}$ for a moderate sequence a_n. For a quantum state $\rho_{A_1R} \in \mathcal{S}(A_1R)$ we have that*

$$\frac{1}{n}I^{\varepsilon_n}_{\max}(\dot{R}^n;A_1^n)_{\rho^{\otimes n}_{A_1R}} = D(\rho_{A_1R}\|\rho_{A_1}\otimes\rho_R) + a_n\sqrt{4V(\rho_{A_1R}\|\rho_{A_1}\otimes\rho_R)} + o(a_n). \tag{36}$$

Proof. We start by showing that the LHS is upper bounded by the RHS. We have

$$I^{\varepsilon_n}_{\max}(\dot{R}^n;A_1^n)_{\rho^{\otimes n}_{A_1R}} \leq I^{\frac{\varepsilon_n}{4}}_{\max}(R^n;A_1^n)_{\rho^{\otimes n}_{A_1R}} + \log\frac{8+(\frac{\varepsilon_n}{2})^2}{(\frac{\varepsilon_n}{2})^2} \tag{37}$$
$$\leq D^{\frac{\varepsilon_n}{4}}_{\max}(\rho^{\otimes n}_{A_1R}\|\rho^{\otimes n}_{A_1}\otimes\rho^{\otimes n}_R) + \log\frac{8+(\frac{\varepsilon_n}{2})^2}{(\frac{\varepsilon_n}{2})^2} \tag{38}$$
$$\leq D^{1-(\frac{\varepsilon_n}{4})^2}_{h}(\rho^{\otimes n}_{A_1R}\|\rho^{\otimes n}_{A_1}\otimes\rho^{\otimes n}_R) + \log\frac{8+(\frac{\varepsilon_n}{2})^2}{(\frac{\varepsilon_n}{2})^2} \tag{39}$$
$$\leq D^{1-\frac{\varepsilon_n^2}{16}}_{h}(\rho^{\otimes n}_{A_1R}\|\rho^{\otimes n}_{A_1}\otimes\rho^{\otimes n}_R) + n\cdot o(a_n) \tag{40}$$
$$= nD(\rho_{A_1R}\|\rho_{A_1}\otimes\rho_R) + n\cdot a_n\sqrt{4V(\rho_{A_1R}\|\rho_{A_1}\otimes\rho_R)} + n\cdot o(a_n), \tag{41}$$

where the first inequality follows from the approximate equivalence of the smoothed max-information and partially smoothed max-information [4, Theorem 2], the third inequality follows from the bounds between the max relative entropy and the hypothesis testing relative entropy [17, Theorem 4]) and the last equality follows by Proposition IV.6. To prove the reverse inequality, let $\rho^{\otimes n}_{A_1RR'}$ be a purification of $\rho^{\otimes n}_{A_1R}$ with $R' \cong A_1R$. For any $\tilde{\rho}_{A_1^nR^n} \approx_{\varepsilon_n} \rho^{\otimes n}_{A_1R}$, there exists a purification $\tilde{\rho}_{A_1^nR^nR'^n} \approx_{\varepsilon_n} \rho^{\otimes n}_{A_1RR'}$. We have

$$I^{\varepsilon_n}_{\max}(\dot{R}^n;A_1^n)_{\rho^{\otimes n}} = \min_{\sigma_{A_1^n}} \min_{\substack{\tilde{\rho}\in\mathcal{B}^{\varepsilon_n}(\rho^{\otimes n})\\ \tilde{\rho}_{R^n}=\rho_R^{\otimes n}}} D_{\max}(\tilde{\rho}_{A_1^nR^n}\|\rho^{\otimes n}_R\otimes\sigma_{A_1^n}) \tag{42}$$
$$\geq \min_{\sigma_{A_1^n}} \min_{\tilde{\rho}\in\mathcal{B}^{\varepsilon_n}(\rho)} D_{\max}(\tilde{\rho}_{A_1^nR^n}\|\rho^{\otimes n}_R\otimes\sigma_{A_1^n}) \tag{43}$$
$$= \min_{\tilde{\rho}\in\mathcal{B}^{\varepsilon_n}(\rho)} \widetilde{I}_{\infty}(\tilde{\rho}_{A_1^nR^n}\|\rho^{\otimes n}_R) \tag{44}$$
$$= \min_{\tilde{\rho}\in\mathcal{B}^{\varepsilon_n}(\rho)} -\widetilde{I}_{\frac{1}{2}}(\tilde{\rho}_{R^nR'^n}\|(\rho^{\otimes n}_R)^{-1}) \tag{45}$$
$$\geq \min_{\tilde{\rho}\in\mathcal{B}^{\varepsilon_n}(\rho)} -\widetilde{D}_{\frac{1}{2}}(\tilde{\rho}_{R^nR'^n}\|(\rho^{\otimes n}_R)^{-1}\otimes\rho^{\otimes n}_{R'}) \tag{46}$$
$$= -D^{\varepsilon_n}_{\min}(\rho^{\otimes n}_{RR'}\|(\rho^{\otimes n}_R)^{-1}\otimes\rho^{\otimes n}_{R'}) \tag{47}$$
$$\geq -D^{2\varepsilon_n^2}_{h}(\rho^{\otimes n}_{RR'}\|(\rho^{\otimes n}_R)^{-1}\otimes\rho^{\otimes n}_{R'}) + o(a_n) \tag{48}$$
$$= -nD(\rho_{RR'}\|\rho_R^{-1}\otimes\rho_{R'}) + n\cdot a_n\sqrt{4V(\rho_{RR'}\|\rho_R^{-1}\otimes\rho_{R'})} - n\cdot o(a_n) \tag{49}$$
$$= nD(\rho_{A_1R}\|\rho_R\otimes\rho_{A_1}) + n\cdot a_n\sqrt{4V(\rho_{A_1R}\|\rho_R\otimes\rho_{A_1})} + n\cdot o(a_n), \tag{50}$$

where the second inequality follows from relaxing the constraint on $\tilde{\rho}$, the second equality follows from the duality of the sandwiched Rényi mutual information [10, Lemma 6], the third inequality follows by choosing a $\rho^{\otimes n}_{R'}$ instead of a minimization, the fourth inequality follows due to Lemma IV.5, the fourth equality due to Proposition IV.6, and the final equality using duality of the relative entropy and relative entropy variance [10, Lemma 6 & Eq. 3.15]. □

Proposition IV.7 is strengthened straightforwardly by noting that constant multiplicative factors on the error ε_n do not affect the moderate deviation analysis.

Remark IV.8. *Proposition IV.7 holds for $0 < \varepsilon_n' < 1$ where $\varepsilon_n' = \Theta(e^{-a_n^2 n})$ for moderate sequence a_n. That is, we have*

$$\frac{1}{n} I_{\max}^{\varepsilon_n}(\dot{R}^n; A_1^n)_{\rho^{\otimes n}} = D(\rho_{A_1R} \| \rho_{A_1} \otimes \rho_R) + a_n\sqrt{4V(\rho_{A_1R}\|\rho_{A_1} \otimes \rho_R)} + o(a_n). \tag{51}$$

This holds because for all constants $k > 0$, $k\varepsilon_n = \exp\left(-n\left(a_n^2 - \frac{\log k}{n}\right)\right) = \exp(-nb_n^2)$ for a moderate sequence $b_n = a_n + o(a_n)$.

Our main result Theorem IV.1 follows by combining Proposition IV.7 and Theorem III.2 and using Remark IV.8.

V. Extensions

Quantum state merging can be understood as a time reversed version of a quantum state splitting protocol [1]. Alice and Bob start with a state ρ_{ABR}, where Alice holds the A register, Bob holds the B register and R is inaccessible. The goal is to use pre-shared entanglement and one-way communication to send the B register to Alice. Note that an equivalent statement to that of Theorem III.2 holds for state merging [4, Theorem 6]. Hence, Theorem IV.1 also holds for state merging and we may henceforth simply state that it applies to quantum state transfer.

Another application of our results is the case where the A register in Definition III.1 is trivial. State splitting in this case is equivalent to entanglement-assisted source coding. To simplify the cost function, we follow the classical definitions [19] to define the quantum varentropy of states $\rho_A \in \mathcal{S}(A)$ as

$$V(A)_\rho = \mathrm{Tr}\left(\rho(\log\rho)^2\right) - S(A)_\rho^2. \tag{52}$$

Lemma V.1. *For bipartite pure state $\rho_{AB} \in \mathcal{S}(AB)$ with marginals ρ_A and ρ_B, the mutual information variance and conditional entropy variance take the form*

$$V(\rho_{AB}\|\rho_A \otimes \rho_B) = 4V(A)_\rho,\ V(\rho_{AB}\|I_A \otimes \rho_B) = V(A)_\rho, \tag{53}$$

respectively.

Proof. From the definition of the relative entropy variance, we have

$$\begin{aligned}
&V(\rho_{AB}\|\rho_A \otimes \rho_B) = \mathrm{Tr}\left(\rho_{AB}(\log\rho_{AB})^2\right) + \mathrm{Tr}\left(\rho_A(\log\rho_A)^2\right) \\
&+ \mathrm{Tr}\left(\rho_B(\log\rho_B)^2\right) + 2\mathrm{Tr}\left(\rho_{AB}\log\rho_A \otimes \log\rho_B\right) \\
&- 2\mathrm{Tr}\left(\rho_{AB}\log\rho_{AB}\log\rho_A\right) - 2\mathrm{Tr}\left(\rho_{AB}\log\rho_{AB}\log\rho_B\right) \\
&- D(\rho_{AB}\|\rho_A \otimes \rho_B)^2 \qquad (54) \\
&= \mathrm{Tr}\left(\rho_A(\log\rho_A)^2\right) + \mathrm{Tr}\left(\rho_B(\log\rho_B)^2\right) \\
&+ 2\mathrm{Tr}\left(\rho_{AB}\log\rho_A \otimes \log\rho_B\right) - (S(A)_\rho + S(B)_\rho)^2 \qquad (55) \\
&= 4V(A)_\rho, \qquad (56)
\end{aligned}$$

where in the second equality we have removed various terms which are zero due to ρ_{AB} being pure, and in the third equality we have used that the marginals ρ_A and ρ_B share the same eigenvalues and that $\mathrm{Tr}\left(\rho_{AB}\log\rho_A \otimes \log\rho_B\right) = \mathrm{Tr}\left(\rho_A(\log\rho_A)^2\right)$ for pure ρ_{AB}. The last claim can be verified by expanding the left hand side in the Schmidt basis of ρ_{AB}. A similar argument shows the second equality. □

Corollary V.2. *The minimal quantum communication cost of one-shot ε_n-error entanglement-assisted source coding of $\rho_{A_1R}^{\otimes n}$ for $\rho_{A_1R} \in \mathcal{S}(A_1R)$ pure with $\varepsilon_n = e^{-na_n^2}$ for a moderate sequence a_n is given as*

$$\frac{1}{n} q_{\varepsilon_n}^{\star}(\rho_{A_1R}^{\otimes n}) = S(A_1)_\rho + 2a_n\sqrt{V(A_1)_\rho} + o(a_n). \tag{57}$$

Next, we show that this result also holds without the presence of entanglement. For the case with no entanglement, the converse proof in Proposition III.5 remains valid. For the achievability we use the bounds shown in [15], for which the error is quantified in terms of the fidelity. For our setting, we modify their result so that source coding is ε_n-close in purified distance. We get an ε_n-error state transfer protocol for pure $\rho_{A_1R}^{\otimes n}$ with $\varepsilon_n = e^{-a_n^2 n}$ for a moderate sequence a_n with the quantum communication cost

$$\begin{aligned}
&\frac{1}{n} q_{\varepsilon_n}^{\star}(\rho_{A_1R}^{\otimes n}) \le \frac{1}{n}\bar{H}_s^{\frac{\varepsilon_n'}{2}}(\rho_{A_1}^{\otimes n}) = -\frac{1}{n}\underline{D}_s^{\frac{\varepsilon_n'}{2}}(\rho_{A_1}^{\otimes n}\|I_{A_1^n}) \qquad (58) \\
&\le -\frac{1}{n} D_h^{\varepsilon_n'/4}(\rho_{A_1}^{\otimes n}\|I_{A_1^n}) - \frac{1}{n}\log\frac{\varepsilon_n'}{4} \qquad (59) \\
&= -D(\rho_{A_1}\|I_{A_1}) + a_n\sqrt{4V(\rho_{A_1}\|I_{A_1})} + o(a_n) \qquad (60) \\
&= D(\rho_{A_1R}\|I_{A_1} \otimes \rho_R) + a_n\sqrt{4V(\rho_{A_1R}\|I_{A_1} \otimes \rho_R)} + o(a_n) \qquad (61) \\
&= S(A_1)_\rho + 2a_n\sqrt{V(A_1)_\rho} + o(a_n), \qquad (62)
\end{aligned}$$

where $\varepsilon_n' = \Theta(\varepsilon_n^2)$. The first inequality is the bound [15, Theorem 5.5 (ii)], the first equality is from the definition of the information spectrum entropy, the second inequality is from the bound between the hypothesis testing relative entropy and the information spectrum relative entropy [15, Proposition 4.7], the second equality is due to Proposition IV.6 and Remark IV.8, and the third equality is through the duality of the relative entropy [10, Lemma 6] and Lemma V.1.

VI. Conclusion

We characterized the communication cost of state transfer in the moderate deviation regime. We note that state transfer can be seen as the underlying primitive of other tasks in quantum information — such as channel simulation or noisy channel coding. For example, one can use de Finetti reductions to reduce the channel simulation problem to a state splitting question [13], [20]. While asymptotic results are understood for this setting, moderate deviation analysis for channel simulation is a natural next application of our results.

VII. Acknowledgements

This work was completed prior to MB joining the AWS Center for Quantum Computing. We thank Michael Cao for useful discussions. MT is supported by the National Research Foundation, Prime Minister's Office, Singapore and the Ministry of Education, Singapore under the Research Centres of Excellence programme, as well as NUS startup grants (R-263-000-E32-133 and R-263-000-E32-731).

References

[1] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, "The mother of all protocols: restructuring quantum information's family tree," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, vol. 465, no. 2108, pp. 2537–2563, 2009.

[2] A. Anshu, V. K. Devabathini, and R. Jain, "Quantum communication using coherent rejection sampling," *Physical Review Letters*, vol. 119, no. 12, p. 120506, 2017.

[3] C. T. Chubb, V. Y. Tan, and M. Tomamichel, "Moderate deviation analysis for classical communication over quantum channels," *Communications in Mathematical Physics*, vol. 355, no. 3, pp. 1283–1315, 2017.

[4] A. Anshu, M. Berta, R. Jain, and M. Tomamichel, "Partially smoothed information measures," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 5022–5036, 2020.

[5] M. Tomamichel, "A framework for non-asymptotic quantum information theory," *arXiv preprint arXiv:1203.2142*, 2012.

[6] K. Li *et al.*, "Second-order asymptotics for quantum hypothesis testing," *Annals of Statistics*, vol. 42, no. 1, pp. 171–189, 2014.

[7] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7693–7710, 2013.

[8] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *Journal of Mathematical Physics*, vol. 54, no. 12, p. 122203, 2013.

[9] M. M. Wilde, A. Winter, and D. Yang, "Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched Rényi relative entropy," *Communications in Mathematical Physics*, vol. 331, no. 2, pp. 593–622, 2014.

[10] M. Hayashi and M. Tomamichel, "Correlation detection and an operational interpretation of the renyi mutual information," *Journal of Mathematical Physics*, vol. 57, no. 10, p. 102201, 2016.

[11] R. Renner, "Security of quantum key distribution," *arXiv preprint arXiv:quant-ph/0512258*, 2005.

[12] F. Dupuis, L. Kraemer, P. Faist, J. M. Renes, and R. Renner, "Generalized entropies," in *XVIIth International Congress on Mathematical Physics*. World Scientific, 2014, pp. 134–153.

[13] M. Berta, M. Christandl, and R. Renner, "The quantum reverse Shannon theorem based on one-shot information theory," *Communications in Mathematical Physics*, vol. 306, no. 3, p. 579, 2011.

[14] M. Tomamichel, *Quantum information processing with finite resources: mathematical foundations*. Springer, 2015, vol. 5.

[15] N. Datta and F. Leditzky, "Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 582–608, 2014.

[16] M. Berta, "Quantum side information: Uncertainty relations, extractors, channel simulations," *arXiv preprint arXiv:1310.4581*, 2013.

[17] A. Anshu, M. Berta, R. Jain, and M. Tomamichel, "A minimax approach to one-shot entropy inequalities," *Journal of Mathematical Physics*, vol. 60, no. 12, p. 122201, 2019.

[18] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 1385–1403, 2017.

[19] I. Kontoyiannis and S. Verdú, "Optimal lossless compression: Source varentropy and dispersion," in *2013 IEEE International Symposium on Information Theory*. IEEE, 2013, pp. 1739–1743.

[20] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, "Quantum reverse Shannon theorem," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2926–2959, 2014.