AIP Journal of Mathematical Physics

Resource cost results for one-way entanglement distillation and state merging of compound and arbitrarily varying quantum sources

H. Boche and G. Janßen

Citation: Journal of Mathematical Physics **55**, 082208 (2014); doi: 10.1063/1.4893635
View online: http://dx.doi.org/10.1063/1.4893635
View Table of Contents: http://scitation.aip.org/content/aip/journal/jmp/55/8?ver=pdfcov
Published by the AIP Publishing

JOURNAL OF MATHEMATICAL PHYSICS **55**, 082208 (2014)

Resource cost results for one-way entanglement distillation and state merging of compound and arbitrarily varying quantum sources

H. Boche[a] and G. Janßen[b]
Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, 80290 München, Germany

(Received 19 February 2014; accepted 4 August 2014; published online 26 August 2014)

We consider one-way quantum state merging and entanglement distillation under compound and arbitrarily varying source models. Regarding quantum compound sources, where the source is memoryless, but the source state an unknown member of a certain set of density matrices, we continue investigations begun in the work of Bjelaković *et al.* ["Universal quantum state merging," J. Math. Phys. **54**, 032204 (2013)] and determine the classical as well as entanglement cost of state merging. We further investigate quantum state merging and entanglement distillation protocols for arbitrarily varying quantum sources (AVQS). In the AVQS model, the source state is assumed to vary in an arbitrary manner for each source output due to environmental fluctuations or adversarial manipulation. We determine the one-way entanglement distillation capacity for AVQS, where we invoke the famous robustification and elimination techniques introduced by Ahlswede. Regarding quantum state merging for AVQS we show by example that the robustification and elimination based approach generally leads to suboptimal entanglement as well as classical communication rates.
 [http://dx.doi.org/10.1063/1.4893635]

I. INTRODUCTION

Investigations on communication tasks involving bipartite (or multipartite) sources within the local operations and classical communications (LOCC) paradigm made a substantial contribution to the progress in quantum Shannon theory which took place over the past two decades.

Especially the role of shared pure entanglement as a communication resource was clarified and substantiated by establishment of LOCC protocols inter-converting shared entanglement with optimal rates.

Two prominent tasks, entanglement distillation and quantum state merging are considered in this work. Quantum state merging was introduced by Horodecki, Oppenheim, and Winter.[14] In this setting a bipartite quantum source described by a quantum state ρ_{AB} shared by communication parties A (sender) and B (receiver) is required to be merged at the receivers site by local operations and classical communication together with shared pure entanglement as resource, such that in the limit of large blocklengths, the source is approximately restored on B's site. The optimal asymptotic net entanglement cost was determined in Ref. 14 to be $S(A|B)$ ebits of shared entanglement per copy of the state, which was shown to be achievable with optimal classical cost $I(A; E)$ bits of $A \rightarrow B$ classical side communication per copy ($I(A; E)$ is the quantum mutual information of A with an environment E purifying ρ_{AB}). This result allows interpretation of the negative values of $S(A|B)$. For states with $S(A|B)$ being negative, quantum state merging is possible with net production of shared maximal entanglement which may serve as a credit for future quantum communication.

[a]Electronic mail: boche@tum.de
[b]Electronic mail: gisbert.janssen@tum.de

Entanglement distillation is in some sense a task subaltern to quantum state merging, since entanglement distillation protocols are readily derived from quantum state merging protocols.[14] In this task, a given bipartite quantum source has to be transformed into shared maximal entanglement by LOCC in the limit of large number of outputs. The optimal entanglement gain was determined in Ref. 12, where a connection to secret key distillation from bipartite quantum states was exhausted.

However, these results were shown under strong idealizations of the sources. It was assumed that the sources, where tasks are performed on, are memoryless and perfectly known. Since source uncertainties, may they be present due to hardware imperfections of the preparation devices and/or manipulation by adversarial communication parties, are inherent to all real-life communication settings, this assumption seems rather restrictive.

The contribution of this work is to partly drop these conditions. We consider entanglement distillation and quantum state merging in presence of compound and arbitrarily varying quantum sources. A compound memoryless source models a preparation device which emits systems, uncorrelated from output to output, all described by the same given density matrix, which in turn is not perfectly known to the communication parties, but identified as a member of a certain set $\mathcal{X}$ of quantum states. Consequently, the communication parties have to use protocols which are of sufficient fidelity for each member of the set of states generating the compound source.

In the arbitrarily varying source (AVQS) model, the source state can vary from output to output over a generating set of states. This variation can be understood as a natural fluctuation as well as a manipulation of an adversarial communication party changing the source state from output to output in an arbitrary manner. Consequently, the parties are forced to accomplish the tasks with protocols, which are robust in the sense, that work with sufficient fidelity for each possible state sequence. In this work, we contribute the following. Regarding one-way quantum state merging for compound sources, we answer a question left open in the preceding work.[7] We derive protocols which beside being optimal regarding their entanglement cost also approximate the lowest classical one-way communication requirements allowed by corresponding converse theorems[7,14] which lower bound the resource requirements for asymptotically faithful merging schemes.

We use the results on one-way entanglement distillation for compound sources established earlier[7] together with the famous elimination and robustification techniques introduced by Ahlswede[1,2] to determine the capacity for one-way entanglement distillation from AVQS generated by a set $\mathcal{X}$ of states. We show that the one-way entanglement distillation capacity in this case can be expressed by the capacity function of the compound source generated by the convex hull of the set generating the AVQS.

Considering quantum state merging under the AVQS model, we encounter unexpected behavior. Opposite to the intuition gathered by previous results from classical as well as quantum Shannon theory, the entanglement as well as classical communication resource costs for one-way merging of an AVQS do not match the costs known for the corresponding compound source generated by the convex hull of $\mathcal{X}$ in general. We demonstrate this fact giving a simple example.

A. Related work

The task of entanglement distillation was subject to several investigations in case of perfectly known memoryless quantum sources over the past 15 years. In this work, we generalize a result from Ref. 12, where the entanglement distillation capacity with one-way LOCC for perfectly known memoryless bipartite quantum sources was determined. Quantum state merging was first considered in Ref. 14, where the authors determined the entanglement as well as classical cost of quantum state merging for the scenario with perfectly known density matrix. Both results were partly generalized to the case of compound memoryless sources in Ref. 7 within the one-way LOCC scenario. In this work we continue and complete considerations made therein by determining the optimal classical cost of one-way merging for compound quantum sources.

Communication tasks involving arbitrarily varying channels and sources were considered in classical information theory from the late 1960s. Here we especially mention the robustification[2,3] and elimination[1] techniques developed by Ahlswede in the 1970s, which are crucial ingredients of our proof of the one-way entanglement distillation capacity for AVQS. Arbitrarily varying channels were also considered in quantum Shannon theory. The first result was by Ahlswede and Blinovsky,[5] who determined the capacity for transmission of classical messages over an arbitrarily varying channel with classical input and quantum output. A treatment of arbitrarily varying quantum channels was done by Ahlswede, Bjelaković, the first author, and Nötzel published in 2013.[4] There, they determined the quantum capacity of an arbitrarily varying quantum channel for entanglement transmission, entanglement generation as well as strong subspace transmission.

B. Outline

We set up the notation used in this paper in Sec. II, where we also state some conventions and preliminary facts we use freely in our considerations. The basic concepts relevant for this paper are concisely stated and explained in Sec. III.

In Sec. IV, we conclude the investigations on quantum state merging for compound sources begun in Ref. 7. Explicitly, we show existence of universal one-way LOCCs which are asymptotically optimal regarding the entanglement as well as classical $A \to B$ communication cost. For the proof, we use protocols derived in Ref. 7, which are optimal regarding their entanglement cost but require overmuch classical side communication in some cases. These are refined in a sufficient way by combination with an entropy estimating instrument used by the sender, where we utilize methods from representation theory of the symmetric groups from Refs. 15 and 10. Section V is devoted to determination of the capacity for entanglement distillation from an AVQS under restriction to one-way LOCC. We first prove an achievability result in case that the AVQS is generated by a finite set $\mathcal{X}$ of bipartite states. Here we use entanglement distillation schemes with fidelity going to one exponentially fast for the compound source generated by the convex hull of $\mathcal{X}$ from Ref. 7, together with Ahlswede's robustification and elimination techniques. Afterwards, we extend this result to the general case approximating the AVQS generating set by suitable finite AVQS. We also consider the issue of quantum state merging for AVQS and discover a strange feature of the quantum state merging task in this scenario. We show in Sec. VI that, in general, the entanglement as well as classical cost of merging an AVQS generated by a set $\mathcal{X}$ of bipartite state are strictly lower than the costs of merging the corresponding compound source generated by $\text{conv}(\mathcal{X})$. In Sec. VII, we discuss the results obtained.

II. NOTATION AND CONVENTIONS

All Hilbert spaces appearing in this work are considered to be finite dimensional complex vector spaces. $\mathcal{L}(\mathcal{H})$ is the set of linear maps and $\mathcal{S}(\mathcal{H})$ the set of states (density matrices) on a Hilbert space $\mathcal{H}$ in our notation. We denote the set of quantum channels, i.e., completely positive and trace preserving (c.p.t.p.) maps from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{K})$ by $\mathcal{C}(\mathcal{H}, \mathcal{K})$ and the set of trace-nonincreasing c.p. maps by $\mathcal{C}^{\downarrow}(\mathcal{H}, \mathcal{K})$ for two Hilbert spaces $\mathcal{H}$, $\mathcal{K}$.

Regarding states on multiparty systems, we freely make use of the following convention for a system consisting of some parties X, Y, Z, for instance, we denote $\mathcal{H}_{XYZ} := \mathcal{H}_X \otimes \mathcal{H}_Y \otimes \mathcal{H}_Z$, and denote the marginals by the letters assigned to subsystems, i.e., $\sigma_{XZ} := \text{tr}_{\mathcal{H}_Y}(\sigma)$ for $\sigma \in \mathcal{S}(\mathcal{H}_{XYZ})$ and so on. For a bipartite pure state $|\psi\rangle\langle\psi|$ on a Hilbert space $\mathcal{H}_{XY}$, we denote its Schmidt rank (i.e., number of nonzero coefficients in the Schmidt representation of ψ) by $\text{sr}(\psi)$. We define

$$F(a, b) := \left\| \sqrt{a}\sqrt{b} \right\|_1^2 \tag{1}$$

for any two positive semidefinite operators a, b on $\mathcal{H}$ (this is the quantum fidelity in case that a and b are density matrices). If one of the arguments is a pure state, the fidelity is linear in the remaining argument, explicitly F takes the form of an inner product,

$$F(\rho, |\psi\rangle\langle\psi|) = \langle\psi, \rho\psi\rangle . \tag{2}$$

Relations between F and the trace distance are well known, we will use the inequalities

$$F(a, \sigma) \geq \text{tr}(a) - \|a - \rho\|_1 \tag{3}$$

for a matrix $0 \leq a \leq \mathbf{1}$ and state ρ, and

$$\|\rho - \sigma\|_1 \leq 2\sqrt{1 - F(\rho, \sigma)} \tag{4}$$

for states ρ, σ. The von Neumann entropy of a quantum state ρ is defined

$$S(\rho) := -\mathrm{tr}(\rho \log \rho), \tag{5}$$

where we denote by $\log(\cdot)$ and $\exp(\cdot)$ the base two logarithms and exponentials throughout this paper. Given a quantum state ρ on $\mathcal{H}_{XY}$, we denote the conditional von Neumann entropy of ρ given Y by

$$S(X|Y, \rho) := S(\rho) - S(\rho_Y), \tag{6}$$

the quantum mutual information by

$$I(X; Y, \rho) := S(\rho_X) + S(\rho_Y) - S(\rho), \tag{7}$$

and the coherent information by

$$I_c(X\rangle Y, \rho) := S(\rho_Y) - S(\rho) = -S(X|Y, \rho). \tag{8}$$

A special class of channels mapping bipartite systems, which is of crucial importance for our considerations, are one-way LOCC channels, for which we give a concise definition in the following. For more detailed information, the reader is referred to the Appendix on one-way LOCCs given in Ref. 7 and references therein. A quantum instrument $\mathcal{T}$ on a Hilbert space $\mathcal{H}$ is given by a set $\{\mathcal{T}_k\}_{k=1}^K \subset \mathcal{C}^{\downarrow}(\mathcal{H}, \mathcal{K})$ of trace non-increasing cp maps, such that $\sum_{k=1}^K \mathcal{T}_k$ is a channel. In this paper, we will only admit instruments with $|K| < \infty$. With bipartite Hilbert spaces $\mathcal{H}_{AB}$ and $\mathcal{K}_{AB}$, a channel $\mathcal{N} \in \mathcal{C}(\mathcal{H}_{AB}, \mathcal{K}_{AB})$ is an $A \to B$ (one-way) LOCC channel, if it is a combination of an instrument $\{\mathcal{T}_k\}_{k=1}^K \subset \mathcal{C}^{\downarrow}(\mathcal{H}_A, \mathcal{K}_A)$ and a family $\{\mathcal{R}_k\}_{k=1}^K \subset \mathcal{C}(\mathcal{H}_B, \mathcal{K}_B)$ of channels in the sense that it can be written in the form

$$\mathcal{N}(a) = \sum_{k=1}^K (\mathcal{T}_k \otimes \mathcal{R}_k)(a) \qquad (a \in \mathcal{L}(\mathcal{H}_{AB})). \tag{9}$$

The cardinality of the message set for classical transmission from A to B within the application of $\mathcal{N}$ is K (the number of measurement outcomes of the instrument).

We denote the set of classical probability distributions on a set $\mathbf{S}$ by $\mathfrak{P}(\mathbf{S})$. The l-fold Cartesian product of $\mathbf{S}$ will be denoted $\mathbf{S}^l$ and $s^l := (s_1, \ldots, s_l)$ will be a notation for elements of $\mathbf{S}^l$. For positive integer n, the shortcut $[n]$ is used to abbreviate the set $\{1, \ldots, n\}$. For two probability distributions $p, q \in \mathfrak{P}(\mathbf{S})$ on a finite set $\mathbf{S}$, the relative entropy of p with respect to q is defined

$$D(p||q) := \begin{cases} \sum_{s \in \mathbf{S}} p(s) \log \frac{p(s)}{q(s)} & \text{if } p \ll q \\ \infty & \text{else} \end{cases}, \tag{10}$$

where $p \ll q$ means $\forall s \in \mathbf{S} : q(s) = 0 \Rightarrow p(s) = 0$. We denote the Shannon entropy of a probability distribution p by $H(p)$. For a set A we denote the convex hull of A by $\mathrm{conv}(A)$. If $\mathcal{X} := \{\rho_s\}_{s \in \mathbf{S}}$ is a finite set of states on a Hilbert space $\mathcal{H}$, it holds

$$\mathrm{conv}(\mathcal{X}) = \left\{ \rho_p \in \mathcal{S}(\mathcal{H}) : \ \rho_p = \sum_{s \in \mathbf{S}} p(s)\, \rho_s, \ q \in \mathfrak{P}(\mathbf{S}) \right\}. \tag{11}$$

By $\mathfrak{S}_l$, we denote the group of permutations on l elements, in this way $\sigma(s^l) = (s_{\sigma(1)}, \ldots, s_{\sigma(l)})$ for each $s^l = (s_1, \ldots, s_l) \in \mathbf{S}^l$ and permutation $\sigma \in \mathfrak{S}_l$.

For any two nonempty sets $\mathcal{X}, \mathcal{X}'$ of states on a Hilbert space $\mathcal{H}$, the Hausdorff distance between $\mathcal{X}$ and $\mathcal{X}'$ (induced by the trace norm $\|\cdot\|_1$) is defined by

$$d_H(\mathcal{X}, \mathcal{X}') := \max \left\{ \sup_{\sigma \in \mathcal{X}} \inf_{\sigma' \in \mathcal{X}'} \|\sigma - \sigma'\|_1, \ \sup_{\sigma' \in \mathcal{X}'} \inf_{\sigma \in \mathcal{X}} \|\sigma - \sigma'\|_1 \right\}. \tag{12}$$

III. BASIC DEFINITIONS

In this section, we define the underlying scenarios, considered in the rest of this paper. Given any set $\mathcal{X} := \{\rho_s\}_{s\in\mathbf{S}} \subset \mathcal{S}(\mathcal{H})$ of states on a Hilbert space $\mathcal{H}$, the *compound source generated by* $\mathcal{X}$ (or *the compound source* $\mathcal{X}$, for short) is given by the family $\{\{\rho_s^{\otimes l}\}_{s\in\mathbf{S}}\}_{l\in\mathbb{N}}$ of states. The above definition models a memoryless quantum source under uncertainty of the statistical parameters. The source outputs each system according to a constant density matrix, while the density matrix itself is not known perfectly by the communication parties. It only can be identified as a member of $\mathcal{X}$. The *AVQS generated by* $\mathcal{X}$ (or *the AVQS* $\mathcal{X}$) is given by the family $\{\{\rho_{s^l}\}_{s^l\in S^l}\}_{l\in\mathbb{N}}$, where we use the definition

$$\rho_{s^l} := \rho_{s_1} \otimes \ldots \otimes \rho_{s_l} \tag{13}$$

for each member $s^l = (s_1, \ldots, s_l)$ of $\mathbf{S}^l$. In the AVQS model, the source density matrix can be chosen from the set $\mathcal{X}$ independently for each output. The variation in the source state models hardware imperfections, where the source is subject to fluctuations in the state on one hand. On the other hand, this definition also can be understood as a powerful communication attack, where the statistical parameters of the source are, to some extend, perpetually manipulated by an adversarial communication party.

A. Quantum state merging

We first give a concise notion of the protocols we admit for quantum state merging. We are interested in the entanglement as well as classical resource costs of quantum state merging. A quantum channel $\mathcal{M}$ is an (l, k_l, D_l) $A \to B$ *merging* for bipartite sources on $\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B$, if it is an $A \to B$ LOCC channel (according to the definition from (9))

$$\mathcal{M} : \mathcal{L}(\mathcal{K}_{0,AB}^l \otimes \mathcal{H}_{AB}^{\otimes l}) \to \mathcal{L}(\mathcal{K}_{1,AB}^l \otimes \mathcal{H}_{B'B}^{\otimes l}), \tag{14}$$

with $k_l := \dim \mathcal{K}_{A,0}^l / \dim \mathcal{K}_{A,1}^l$, where we assume $\mathcal{K}_{A,i} \simeq \mathcal{K}_{B,i}$ $(i = 1, 2)$, and

$$\mathcal{M}(x) = \sum_{k=1}^{D_l} \mathcal{A}_k \otimes \mathcal{B}_k(x). \qquad (x \in \mathcal{L}(\mathcal{K}_{0,AB}^l \otimes \mathcal{H}_{AB}^{\otimes l})), \tag{15}$$

where $\{\mathcal{A}_k\}_{k=1}^{D_l} \subset \mathcal{C}^{\downarrow}(\mathcal{K}_{0,A}^l \otimes \mathcal{H}_A^{\otimes l}, \mathcal{K}_{1,A}^l)$ constitutes an instrument and $\{\mathcal{B}_k\}_{k=1}^{D_l} \subset \mathcal{C}(\mathcal{K}_{B,0}^l \otimes \mathcal{H}_B^{\otimes l}, \mathcal{K}_{B,1}^l \otimes \mathcal{H}_{B'B}^{\otimes l})$ is a set of channels depending on the parameter $k \in [D_l]$. The spaces $\mathcal{K}_{AB,0}^l, \mathcal{K}_{AB,1}^l$ are understood to represent bipartite systems shared by A and B, which carry the input and output entanglement resources used in the process. As a convention, we will incorporate the maximally entangled states $\phi_i^l \in \mathcal{S}(\mathcal{K}_{AB,i}^l)$, $i = 0, 1$ into the definition of the protocol, it holds

$$k_l := \frac{\dim \mathcal{K}_{0,A}^l}{\dim \mathcal{K}_{1,A}^l} = \frac{\dim \mathcal{K}_{0,B}^l}{\dim \mathcal{K}_{1,B}^l} = \frac{\mathrm{sr}(\phi_0^l)}{\mathrm{sr}(\phi_1^l)}. \tag{16}$$

We define the *merging fidelity* of $\mathcal{M}_l$ given a state $\rho^l \in \mathcal{S}(\mathcal{H}_{AB}^{\otimes l})$ by

$$F_m(\rho^l, \mathcal{M}_l) := F\left(\mathcal{M}_l \otimes \mathrm{id}_{\mathcal{H}_E^l}(\phi_0^l \otimes \psi^l), \phi_1^l \otimes \psi'^l\right). \tag{17}$$

Here, ψ^l is a purification of ρ^l with an environmental system described on an additional Hilbert space $\mathcal{H}_E^l$ (usually $\mathcal{H}_E^l = \mathcal{H}_E^{\otimes l}$ with some space $\mathcal{H}_E$), and ψ'^l is a state identical to ψ^l but defined on $\mathcal{H}_{B'B}^{\otimes l}$ completely under control of B. It was shown in Ref. 7 (Lemma 1) that the r.h.s. of (17) does not depend on the chosen purification (which justifies the definition of F_m), and that the function F_m is convex in the first and linear in the second argument. For the rest of this section, we assume $\mathcal{X} := \{\rho_s\}_{s\in\mathbf{S}}$ to be any set of bipartite states on $\mathcal{H}_{AB}$.

Definition 1. A number $R_q \in \mathbb{R}$ *is called an* achievable entanglement cost for $A \to B$ merging of the compound source $\mathcal{X}$ with classical communication rate R_c, *if there exists a sequence* $\{\mathcal{M}_l\}_{l\in\mathbb{N}}$ *of* (l, k_l, D_l) $A \to B$ *mergings, such that the conditions*

1. $\lim_{l\to\infty} \inf_{\rho\in\mathcal{X}} F_m(\rho^{\otimes l}, \mathcal{M}_l) = 1$,

2. $\limsup_{l\to\infty} \frac{1}{l}\log k_l \leq R_q$,

3. $\limsup_{l\to\infty} \frac{1}{l}\log D_l \leq R_c$

are satisfied.

In the following definition, priority lies on the optimal entanglement consumption (or gain) of merging processes, while the classical communication requirements are of subordinate priority. However, the classical communication is required to be rate bounded in the asymptotic limit. Since the classical communication requirements are of interest as well, we also determine the optimal classical communication cost in Sec. IV.

Definition 2. The $A \to B$ merging cost $C^{AV}_{m,\to}(\mathcal{X})$ *of the compound source* $\mathcal{X}$ *is defined by*

$$C_{m,\to}(\mathcal{X}) := \inf\left\{ R_q \in \mathbb{R} : \begin{array}{l} R_q \textit{ is an achievable entanglement cost for } A \to B \\ \textit{merging of the compound source } \mathcal{X} \text{ with} \\ \textit{some classical communication rate } R_c \end{array} \right\}. \tag{18}$$

We recall the following theorem proven in Ref. 7

Theorem 3 (cf. Ref. 7).

$$C_{m,\to}(\mathcal{X}) = \sup_{\rho\in\mathcal{X}} S(A|B,\rho). \tag{19}$$

Definition 4. A number $R_q \in \mathbb{R}$ *is called an* achievable entanglement cost for $A \to B$ merging of the AVQS $\mathcal{X}$ with classical communication rate R_c *if there exists a sequence* $\{\mathcal{M}_l\}_{l\in\mathbb{N}}$ *of* (l, k_l, D_l) $A \to B$ *mergings satisfying*

1. $\lim_{l\to\infty} \inf_{s^l\in\mathbf{S}^l} F_m(\rho_{s^l}, \mathcal{M}_l) = 1$,

2. $\limsup_{l\to\infty} \frac{1}{l}\log k_l \leq R_q$,

3. $\limsup_{l\to\infty} \frac{1}{l}\log D_l \leq R_c$.

Definition 5. The $A \to B$ *merging cost* $C^{AV}_{m,\to}(\mathcal{X})$ *of the AVQS* $\mathcal{X}$ *is defined by*

$$C^{AV}_{m,\to}(\mathcal{X}) := \inf\left\{ R_q \in \mathbb{R} : \begin{array}{l} R_q \textit{ is an achievable entanglement cost for } A \to B \textit{ merging} \\ \textit{of the AVQS } \mathcal{X} \textit{ with some classical communication rate } R_c \end{array} \right\}. \tag{20}$$

B. Entanglement distillation

Concerning entanglement distillation, we are interested in the asymptotically entanglement gain of one-way LOCC distillation procedures. We use the following definitions.

Definition 6. A non-negative number R is an achievable $A \to B$ entanglement distillation rate for the AVQS generated by a set $\mathcal{X}$ with classical rate R_c, *if there exists a sequence* $\{\mathcal{D}_l\}_{l\in\mathbb{N}}$ *of* $A \to B$ *LOCC channels*,

$$\mathcal{D}_l = \sum_{m=1}^{M_l} \mathcal{A}_{m,l} \otimes \mathcal{B}_{m,l} \qquad (l \in \mathbb{N}) \tag{21}$$

such that the conditions

1. $\lim_{l\to\infty} \inf_{s^l\in\mathbf{S}^l} F(\mathcal{D}_l(\rho_{s^l}), \phi_l) = 1$,

2. $\liminf_{l\to\infty} \frac{1}{l}\log \mathrm{sr}(\phi_l) \geq R$,
3. $\limsup_{l\to\infty} \frac{1}{l}\log M_l \leq R_c$

are fulfilled, where ϕ_l is a maximally entangled state shared by A and B for each $l \in \mathbb{N}$.

In this paper, we will be primarily interested in the entanglement gain of one-way entanglement distillation. Regarding the classical communication cost of entanglement distillation, no general cost results are known even in case that the source is memoryless with perfectly known source state.[12]

Definition 7. The $A \to B$ entanglement distillation capacity for the AVQS generated by $\mathcal{X}$ is defined

$$D_{\to}^{AV}(\mathcal{X}) := \sup\left\{R : \begin{array}{l} R \text{ is an achievable } A \to B \text{ entanglement distillation rate for} \\ \text{the AVQS } \mathcal{X} \text{with some classical communication rate } R_c \end{array}\right\}. \tag{22}$$

The corresponding definitions for achievable rates and entanglement distillation capacity of compound sources can be easily guessed (see Ref. 7). To introduce some notation we use in this paper, we state a theorem from Ref. 12, where the $A \to B$ entanglement distillation capacity $D_{\to}(\rho)$ of a memoryless bipartite quantum source with perfectly known density matrix ρ was considered.

Theorem 8 (Ref. 12, Theorem 3.4). *Let ρ be a state on $\mathcal{H}_{AB}$. It holds*

$$D_{\to}(\rho) = \lim_{k\to\infty}\frac{1}{k}\sup_{\mathcal{T}\in\Theta_k} D^{(1)}(\rho^{\otimes k}, \mathcal{T}) \tag{23}$$

with

$$D_{\to}^{(1)}(\sigma, \mathcal{T}) := \sum_{\substack{j\in[J]:\\ \lambda_j(\sigma)\neq 0}} \lambda_j(\sigma)\, I_c(A\rangle B, \sigma_j), \tag{24}$$

where Θ_k is the set of finite-valued quantum instruments on A's site, i.e.,

$$\Theta_k := \left\{\{\mathcal{T}_j\}_{j=1}^J \subset \mathcal{C}^{\downarrow}(\mathcal{H}_A^{\otimes k}, \mathcal{K}_A) : \sum_{j=1}^J \mathcal{T}_j \in \mathcal{C}(\mathcal{H}_A^{\otimes k}, \mathcal{K}_A),\ J < \infty,\ \dim \mathcal{K}_A < \infty\right\}. \tag{25}$$

For each state σ and quantum instrument $\mathcal{T} := \{\mathcal{T}_j\}_{j=1}^J$ on A's site and definitions

$$\lambda_j(\sigma) := \mathrm{tr}(\mathcal{T}_j(\sigma_A)), \text{ and } \quad \sigma_j := \frac{1}{\lambda_j}(\sigma)(\mathcal{T}_j \otimes \mathrm{id}_{\mathcal{H}_B})(\sigma) \tag{26}$$

for each j with $\lambda_j(\sigma) \neq 0$.

Remark 9. It is known[12] that the limit in (23) exists for each state, and maximization over instruments in this formula is always realized by an instrument $\mathcal{T} = \{\mathcal{T}_j\}_{j=1}^J$ with $J \leq \dim \mathcal{H}_A^{2k}$ and the operation $\mathcal{T}_j$ described by only one Kraus operator for $1 \leq j \leq J$.

In order to obtain a compact notation for the capacity functions arising in the entanglement distillation scenarios we consider in this paper, we introduce a one-way LOCC $\hat{\mathcal{T}} := \sum_{j=1}^J \mathcal{T}_j \otimes |e_j\rangle\langle e_j|$ for each instrument $\{\mathcal{T}_j\}_{j=1}^J$ with domain $\mathcal{H}_A$ and an orthonormal system $\{e_j\}_{j=1}^J$ in a suitable space $\mathcal{H}'_B \simeq \mathbb{C}^J$ assigned to B, it holds

$$D^{(1)}(\sigma, \mathcal{T}) = I_c(A\rangle BB', \hat{\mathcal{T}}(\sigma)) \tag{27}$$

in (24) for each given state σ.

IV. QUANTUM STATE MERGING FOR COMPOUND QUANTUM SOURCES

In this section, we derive, for any given bipartite compound source $\mathcal{X}$, asymptotically faithful state merging protocols, which are approximately optimal regarding their entanglement as well as classical $A \to B$ communication cost given the corresponding converse statement.[7] While the merging cost was determined in Ref. 7 before (see Theorem 3 above also), the protocols used there are suboptimal, in general, regarding their classical $A \to B$ communication requirements. However, it was shown there (see Sec. V in Ref. 7) that

$$R_c = \sup_{\rho \in \mathcal{X}} I(A; E, \rho) \tag{28}$$

(supremum of the quantum mutual information between A and a purifying environment E) is a lower bound on the $A \to B$ classical communication cost for merging a compound source $\mathcal{X}$ by protocols which have fidelity one in the limit of large blocklengths.

Proposition 13 below states that this bound actually is achievable, and thus together with results from Ref. 7 provides a full solution of the quantum state merging problem for compound quantum sources. The assertions proved in this section will be utilized in Sec. VI, where we compare the $A \to B$ merging as well as the classical communication cost of a certain AVQS merging protocol for a set $\mathcal{X}$ with the optimal costs of state merging protocols for the compound source generated by $\mathrm{conv}(\mathcal{X})$.

The preliminary Proposition 10 below is a slight generalization of Theorem 6 in Ref. 7. It states existence of protocols achieving the optimal entanglement cost, but with generally suboptimal classical communication rates. However, these protocols will be utilized to derive protocols suitable for the proof of Proposition 13.

Proposition 10 (cf. Ref. 7, Theorem 6). Let $\mathcal{X} \subset \mathcal{S}(\mathcal{H}_{AB})$. For each $\delta > 0$, there is a number $l_0 \in \mathbb{N}$, such that for each blocklength $l > l_0$ there is an (l, k_l, D_l) $A \to B$ merging $\mathcal{M}_l$, such that

$$\inf_{\rho \in \mathcal{X}} F(\rho^{\otimes l}, \mathcal{M}_l) \geq 1 - 2^{-lc_1} \tag{29}$$

with a constant $c_1 = c_1(\mathcal{X}, \delta) > 0$,

$$\frac{1}{l} \log k_l \leq \sup_{\rho \in \mathcal{X}} S(A|B, \rho) + \delta \tag{30}$$

and

$$\frac{1}{l} \log D_l \leq \sup_{\rho \in \mathcal{X}} S(\rho_A) + \sup_{\rho \in \mathcal{X}} S(A|B, \rho) + \delta. \tag{31}$$

Proof. The assertion to prove includes both a strengthening of the fidelity convergence rates in Ref. 7, Theorem 4 to exponentially decreasing trade-offs, and a generalization of Theorem 6 in Ref. 7 to arbitrary (not necessary finite or countable) sets of states. Approximating $\mathcal{X}$ by a τ_l-net $\mathcal{X}_{\tau_l} := \{\rho_i\}_{i=1}^{N_{\tau_l}} \subset \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ for each blocklength l (see Ref. 7 for details) and using the result for finite sets, we infer by careful observation of the merging fidelities in Ref. 7 (see Eqs. (36), (37), and (58) therein) that for given $\delta > 0$ and large enough blocklength l, there exists a (l, k_l, D_l) $A \to B$ merging $\mathcal{M}_l$, where

$$\inf_{\rho \in \mathcal{X}} F_m(\rho^{\otimes l}, \mathcal{M}_l) \geq 1 - N_{\tau_l}^2 \cdot 2^{-l\theta} - 4\sqrt{l \cdot \tau_l} \tag{32}$$

is valid for the merging fidelities with a constant $\theta = \theta(\delta) > 0$, and

$$\frac{1}{l} \log k_l \leq \sup_{\rho \in \mathcal{X}} S(A|B, \rho) + \frac{\delta}{2} \tag{33}$$

(see (57) in Ref. 7). Moreover, we can bound the number of messages for the classical $A \to B$-communication (see (101) in Ref. 7) by

$$\frac{1}{l}\log D_l \leq \max_{1\leq i\leq N_{\tau_l}} S(\rho_{A,i}) + \max_{1\leq i\leq N_{\tau_l}} S(A|B,\rho_i) + \frac{\delta}{2} + \frac{1}{l}\log N_{\tau_l} \tag{34}$$

$$\leq \sup_{\rho\in\mathcal{X}} S(\rho_A) + \sup_{\rho\in\mathcal{X}} S(A|B,\rho) + \overline{\nu}(\tau_l) + \frac{\delta}{2} + \frac{1}{l}\log N_{\tau_l}, \tag{35}$$

where the summand $\overline{\nu}(\tau_l) := 3\tau_l \log\frac{\dim\mathcal{H}_{AB}}{\tau_l}$ follows from threefold application of Fannes' inequality,[13] i.e.,

$$\left|\max_{1\leq i\leq N_{\tau_l}} S(\rho_{A,i}) + \max_{1\leq i\leq N_{\tau_l}} S(A|B,\rho_i) - \sup_{\rho\in\mathcal{X}} S(\rho_A) + \sup_{\rho\in\mathcal{X}} S(A|B,\rho)\right| \leq \overline{\nu}(\tau_l). \tag{36}$$

Due to the bound given in Ref. 7, Lemma 9, it is known that the nets can be chosen with cardinality bounded by

$$N_{\tau_l} \leq \left(\frac{3}{\tau_l}\right)^{2(\dim\mathcal{H}_{AB})^2} \tag{37}$$

for each $l \in \mathbb{N}$. Choosing net parameter $\tau_l = 2^{-l\theta'}$ with $\theta' := \min\{\theta/8(\dim\mathcal{H}_{AB})^2, \delta/4\}$ for each l, we infer

$$\inf_{\rho\in\mathcal{X}} F_m(\rho^{\otimes l},\mathcal{M}_l) \geq 1 - 2^{-l\frac{\theta}{2}} - 2^{-l\frac{\theta'}{4}} \geq 1 - 2^{-lc_1} \tag{38}$$

with a constant $c_1 = c_1(\delta) > 0$, and

$$\frac{1}{l}\log D_l \leq \sup_{\rho\in\mathcal{X}} S(\rho_A) + \sup_{\rho\in\mathcal{X}} S(A|B,\rho) + \delta \tag{39}$$

from (35) if l is large enough to satisfy $\overline{\nu}(\tau_l) \leq \frac{\delta}{4}$. Collecting the bounds in (33), (38), and (39), we are done. □

Before we state and prove Proposition 13, we collect some results from representation theory of the symmetric groups, which we utilize in the proof.

We denote by $YF_{d,l}$ the set of young frames with at most d rows and l boxes for $d, l \in \mathbb{N}$. A young frame $\lambda \in YF_{d,l}$ is determined by a tuple $(\lambda_1, \ldots, \lambda_d)$ of nonnegative integers summing to l. The box-lengths $\lambda_1, \ldots, \lambda_d$ of λ define a probability distribution $\overline{\lambda}$ on $[d]$ in a natural way via the definition $\overline{\lambda}(i) := \frac{1}{l}\lambda_i$ for each $1 \leq i \leq d$. To each Young frame $\lambda \in YF_{d,l}$, there is an invariant subspace of $(\mathbb{C}^d)^{\otimes l}$, and we denote by $P_{\lambda,l}$ the projector onto the subspace belonging to λ.

Theorem 11 below allows to asymptotically estimate the spectrum of a density operator ρ by projection valued measurements on i.i.d. (independent and identically distributed) sequences of the form $\rho^{\otimes l}$, and is an important ingredient of our proof of Proposition 13. A variant of the first statement of the theorem was first proven in by Keyl and Werner.[15] The actual bounds stated below are from Ref. 10, while the remaining statements of the theorem are well-known facts in group representation theory (Ref. 10 and references therein are recommended for further information).

Theorem 11 (cf. Refs. 15 and 10). *The following assertions are valid for each $d, l \in \mathbb{N}$.*

1. *For $\lambda \in YF_{d,l}$ and $\rho \in \mathcal{S}(\mathbb{C}^d)$, it holds*
$$\mathrm{tr}(P_{\lambda,l}\rho^{\otimes l}) \leq (l+1)^{d(d-1)/2}\exp(-lD(\overline{\lambda}||r)), \tag{40}$$
where $\overline{\lambda} \in \mathfrak{P}([d])$ is the probability distribution given by the normalized box-lengths of λ, and r is the probability distribution on $[d]$ induced by the decreasingly ordered spectrum of ρ (with multiplicities of eigenvalues counted).
2. $|YF_{d,l}| \leq (l+1)^d$.
3. *For $\lambda, \lambda' \in YF_{d,l}$, it holds $P_{\lambda,l}P_{\lambda',l} = 0$ if $\lambda \neq \lambda'$.*

Lemma 12 (Refs. 19 and 17). Let τ, X be matrices with $\tau \geq 0$, $\mathrm{tr}(\tau) \leq 1$, and $0 \leq X \leq 1$, $\epsilon \in (0, 1)$. If $\mathrm{tr}(\rho X) \geq 1 - \epsilon$, it holds

$$\|\sqrt{X}\rho\sqrt{X} - \rho\|_1 \leq 2\sqrt{\epsilon}. \tag{41}$$

The following proposition is the main result of this section.

Proposition 13. Let $\mathcal{X} \subset \mathcal{S}(\mathcal{H}_{AB})$ be a set of states on $\mathcal{H}_{AB}$. For each $\delta > 0$, there exists a number $l_0 = l_0(\delta)$, such that for each $l > l_0$ there is an (l, k_l, D_l) $A \to B$ merging $\mathcal{M}_l$ with

$$\inf_{\rho \in \mathcal{X}} F_m(\rho^{\otimes l}, \mathcal{M}_l) \geq 1 - 2^{-lc_2} \tag{42}$$

with a constant $c_2 = c_2(\mathcal{X}, \delta) > 0$,

$$\frac{1}{l} \log k_l \leq \sup_{\rho \in \mathcal{X}} S(A|B, \rho) + \delta \tag{43}$$

and

$$\frac{1}{l} \log D_l \leq \sup_{\rho \in \mathcal{X}} I(A; E, \rho) + \delta, \tag{44}$$

where the quantum mutual information in (44) is evaluated on the AE marginal state of any purification ψ of ρ (notice that the above abuse of notation does not lead to ambiguities, since $I(A; E, \mathrm{tr}_{AE}(\psi)) = S(\rho_A) + S(A|B, \rho)$ holds for any purification ψ of ρ).

Remark 14. Regarding the classical communication cost a quite restrictive converse statement was shown to be valid.[7] Asymptotically faithful one-way state merging schemes demand classical communication at rate

$$R_c \geq \sup_{\rho \in \mathcal{X}} I(A; E, \rho) \tag{45}$$

regardless of the entanglement rate achieved, i.e., even investing more entanglement resources (choosing protocols with suboptimal merging rates) does not lead to a reduction of the classical communication cost in a significant way.

Proof of Proposition 13. One half of the above assertion was already proven (see Ref. 7 and Proposition 10 at the beginning of this section). Explicitly, it was shown there that $A \to B$ LOCC channels exist for each set of bipartite states, which for sufficiently large blocklengths fulfill the conditions formulated in (42) and (43). We complete the proof by demonstrating that also the constraint (44) on the classical $A \to B$ communication rate can be met simultaneously with (42) and (43) by certain protocols. The strategy of our proof will be as follows. We decompose $\mathcal{X}$ into disjoint subsets $\mathcal{X}_1, \ldots, \mathcal{X}_N$, each containing only states with approximately equal entropy on the A-marginal system and combine an entropy estimating instrument on the A-system with a suitable merging scheme for each set $\mathcal{X}_i$ according to Proposition 10. We fix $\delta > 0$, and assume, to simplify the argument, that

$$\tilde{s} := \sup_{\rho \in \mathcal{X}} S(A|B, \rho) < 0 \tag{46}$$

holds (i.e., merging is possible without input entanglement resources for large enough blocklengths). Otherwise the argument below can be carried out using further input entanglement and wasting it before action of the protocol. We define $d := \dim \mathcal{H}_A$ and fix $\eta \in (0, 1]$ to be determined later. Consider the sequence

$$s_0 := 0 < s_1 < \ldots < s_N := \log d, \quad s_i := s_{i-1} + \eta \text{ for each } 1 \leq i < N. \tag{47}$$

Define intervals $I_1 := [s_0, s_1]$ and $I_i := (s_{i-1}, s_i]$ for $i = 2, \ldots, N$, which generate a decomposition of $\mathcal{X}$ into disjoint sets $\mathcal{X}_1, \ldots, \mathcal{X}_N$ by definitions

$$\mathcal{X}_i := \{\rho \in \mathcal{X} : S(\rho_A) \in I_i\} \qquad (i \in [N]), \tag{48}$$

and set

$$\tilde{\mathcal{X}}_i := \bigcup_{j \in n(i)} \mathcal{X}_i (i \in [N]), \tag{49}$$

where $n(i)$ is defined $n(i) := \{j \in [N] : |j - i| \leq 1\}$ for all i. In order to construct an entropy estimating instrument in the A marginal systems, we define an operation $\mathcal{P}_l^{(i)} \in \mathcal{C}^{\downarrow}(\mathcal{H}_{AB}^{\otimes l}, \mathcal{H}_{AB}^{\otimes l})$ by

$$\mathcal{P}_l^{(i)}(\cdot) := p_{i,l} \otimes \mathbf{1}_{\mathcal{H}_B^{\otimes l}}(\cdot) p_{i,l}^* \otimes \mathbf{1}_{\mathcal{H}_B^{\otimes l}} \text{ with } \quad p_{i,l} := \sum_{\substack{\lambda \in YF_{d,l}: \\ H(\bar{\lambda}) \in I_i}} P_{\lambda,l} \tag{50}$$

for each $i \in [N]$ using the notation from Theorem 11. Notice that $p_1, \ldots, p_N$ form a projection valued measure on $\mathcal{H}_A^{\otimes l}$ due to Theorem 11.3. By construction, we have for each state $i \in [N]$, $\rho \in \mathcal{X}_i$,

$$\sum_{j \in [N] \setminus n(i)} \operatorname{tr}(\mathcal{P}_l^{(j)}(\rho^{\otimes l})) = \sum_{j \in [N] \setminus n(i)} \operatorname{tr}(p_{i,l} \rho_A^{\otimes l}) \tag{51}$$

$$= \sum_{\substack{\lambda \in YF_{d,l}: \\ |H(\lambda) - S(\rho_A)| \geq \eta}} \operatorname{tr}(P_{\lambda,l} \rho_A^{\otimes l}) \tag{52}$$

$$\leq |YF_{d,l}| \cdot (l+1)^{d(d-1)/2} \tag{53}$$

$$\times \exp\left(-l \left(\min_{r : H(r) \in I_i} \min_{\substack{\lambda \in YF_{d,l}: \\ |H(\bar{\lambda}) - H(r)| \geq \eta}} D(\bar{\lambda} || r)\right)\right), \tag{54}$$

where (51) and (52) are valid due to construction and (54) follows from Theorem 11.1. Since the relative entropy term in the exponent on the r.h.s. of (54) is bounded away from zero for each fixed number $\eta > 0$ (consult the Appendix of this paper for a proof of this fact), i.e.,

$$\min_{r : H(r) \in I_i} \min_{\substack{\lambda \in YF_{d,l}: \\ |H(\bar{\lambda}) - H(r)| \geq \eta}} D(\bar{\lambda} || r) \geq 2c_3 \qquad (i \in [N]) \tag{55}$$

with a constant $c_3 = c_3(\eta) > 0$, and the functions outside the exponential term are growing polynomially for $l \to \infty$ (see Theorem 11.2), we deduce

$$\sum_{j \in [N] \setminus n(i)} \operatorname{tr}(\mathcal{P}_l^{(j)}(\rho^{\otimes l})) \geq 2^{-lc_3} \qquad (i \in [N]) \tag{56}$$

provided that l is large enough. Define index sets $J := \{i : \mathcal{X}_i \neq \emptyset\}$ and $\tilde{J} := \{i : \tilde{\mathcal{X}}_i \neq \emptyset\}$. We know from Proposition 10 that for each sufficiently large l, we find an $(l, k_l, D_l^{(i)})$ $A \to B$ merging $\widetilde{\mathcal{M}}_l^{(i)}$ for each $i \in \tilde{J}$ such that

$$\inf_{\rho \in \tilde{\mathcal{X}}_i} F_m(\rho^{\otimes l}, \widetilde{\mathcal{M}}_l^{(i)}) \geq 1 - 2^{-l\tilde{c}_i} \tag{57}$$

holds with a constant $\tilde{c}_i > 0$,

$$-\frac{1}{l} \log k_l \leq \sup_{\rho \in \tilde{\mathcal{X}}_i} S(A|B, \rho) + \frac{\delta}{2} \tag{58}$$

and

$$\frac{1}{l} \log D_l^{(i)} \leq \sup_{\rho \in \tilde{\mathcal{X}}_i} S(\rho_A) + \sup_{\rho \in \tilde{\mathcal{X}}_i} S(A|B, \rho) + \frac{\delta}{2} \tag{59}$$

for the classical $A \to B$ communication rate. By construction of the sets $\tilde{\mathcal{X}}_i$, $i \in \tilde{J}$, it also holds

$$I(A;E,\rho) \;=\; S(\rho_A) + S(A|B,\rho) \;\geq\; \sup_{\rho\in\tilde{\mathcal{X}}_i} S(\rho_A) - 3\eta + S(A|B,\rho) \tag{60}$$

for each $\rho \in \tilde{\mathcal{X}}_i$. Taking suprema over the set $\tilde{\mathcal{X}}_i$ on both sides of the above inequality in combination with (59) leads us to the estimate

$$\frac{1}{l}\log D_l^{(i)} \;\leq\; \sup_{\rho\in\tilde{\mathcal{X}}_i} I(A;E,\rho) + \frac{\delta}{2} + 3\eta \;\leq\; \sup_{\rho\in\mathcal{X}} I(A;E,\rho) + \frac{\delta}{2} + 3\eta \tag{61}$$

for each $i \in [\tilde{J}]$. Combining the entropy estimating instrument $\{\mathcal{P}_l^{(j)}\}_{j=1}^N$ with the corresponding merging protocols, we define

$$\mathcal{M}_l(\cdot) := \sum_{i=1}^N \widetilde{\mathcal{M}}_l^{(i)} \circ \mathcal{P}_l^{(i)}(\cdot). \tag{62}$$

The maps $\widetilde{\mathcal{M}}_l^{(i)}$ are yet undefined for all numbers $i \in [N] \setminus \widetilde{J}$. Since they will not be relevant for the fidelity, they may be defined by any trivial local operations, with $D_l^{(i)} = 1$ for $i \in [N] \setminus \widetilde{J}$. Moreover, we assume that the merging rate of $\mathcal{M}_l^{(i)}$ for each i is stuck to the worst and each $\mathcal{M}_l^{(i)}$ outputs approximately the same maximally entangled resource output state ϕ_l. We can always achieve this by partial tracing and local unitaries, which do not further affect the classical communication rates. By inspection of the definition in (62) one readily verifies that $\mathcal{M}_l$ is, in fact, an (l, k_l, D_l) $A \to B$ merging, with

$$D_l = \sum_{i\in\tilde{J}} D_l^{(i)} + |N - \widetilde{J}|, \tag{63}$$

and therefore, classical communication rate bounded by

$$\frac{1}{l}\log D_l = \frac{1}{l}\log\left(\sum_{i\in\tilde{J}} D_l^{(i)} + |N - \widetilde{J}|\right) \tag{64}$$

$$\leq \frac{1}{l}\log\left(N \cdot \max_{i\in[N]} D_l^{(i)}\right) \tag{65}$$

$$\leq \sup_{\rho\in\mathcal{X}} I(A;E,\rho) + \frac{\delta}{2} + 3\eta + \frac{\log N}{l}. \tag{66}$$

It remains to show that we achieve merging fidelity one with $\{\mathcal{M}_l\}_{l\in\mathbb{N}}$ for each $\rho \in \mathcal{X}$ with exponentially decreasing trade-offs for large enough blocklengths. Assume ρ is a member of $\mathcal{X}_i$ for any index $i \in J$. Then, it holds

$$F_m(\rho^{\otimes l}, \mathcal{M}_l) \geq \sum_{j\in n(i)} F_m(\rho^{\otimes l}, \widetilde{\mathcal{M}}_l^{(j)} \circ \mathcal{P}_l^{(j)}) \tag{67}$$

$$= \sum_{j\in n(i)} F(\rho^{\otimes l}, \widetilde{\mathcal{M}}_l^{(j)} \circ \widetilde{\mathcal{P}}_l^{(i)}) - \sum_{j\in n(i)} \sum_{\substack{k\in n(i)\\ k\neq j}} F(\rho^{\otimes l}, \widetilde{\mathcal{M}}_l^{(j)} \circ \mathcal{P}_l^{(k)}). \tag{68}$$

The inequality above holds because the merging fidelity is linear in the operation and all summands are nonnegative together with the definition of $\mathcal{M}_l$. The equality is by some zero-adding of terms and using the definition $\widetilde{\mathcal{P}}_l^{(i)} := \sum_{j\in n(i)} \mathcal{P}_l^{(j)}$ together with linearity of the merging fidelity in the operation again. We bound the terms in (68) separately. Beginning with the second term, we notice that the fidelity is homogeneous in its inputs and bounded by one for states, it holds

$$F(\widetilde{\mathcal{M}}_l^{(j)} \circ \mathcal{P}_l^{(k)} \otimes \mathrm{id}_{\mathcal{H}_E^{\otimes n}}(\psi_l), \phi_l \otimes \psi_l') \leq \mathrm{tr}(\mathcal{P}_l^{(k)}(\rho_A^{\otimes n})). \tag{69}$$

Summing up the bounds in (69), rearranging the summands and using the definition of $\widetilde{\mathcal{P}}_l^{(i)}$, we obtain the bound

$$\sum_{j\in n(i)}\sum_{\substack{k\in n(i)\\ k\neq j}} F_m(\rho^{\otimes l}, \widetilde{\mathcal{M}}_l^{(j)}\circ\mathcal{P}_l^{(k)}) \leq \sum_{j\in n(i)}\sum_{\substack{k\in n(i)\\ k\neq j}} \operatorname{tr}(\mathcal{P}_l^{(k)}(\rho^{\otimes n})) \tag{70}$$

$$= (|n(i)|-1)\operatorname{tr}(\widetilde{\mathcal{P}}_l^{(i)}(\rho^{\otimes l})) \tag{71}$$

$$\leq |n(i)|-1. \tag{72}$$

To bound the first terms in (68), we use the relation between fidelity and trace norm in (3). It then holds for each $j \in n(i)$,

$$F_m(\rho^{\otimes l}, \widetilde{\mathcal{M}}_l^{(j)}\circ\widetilde{\mathcal{P}}_l^{(i)}) \;\geq\; \operatorname{tr}(\widetilde{\mathcal{P}}_l^{(i)}(\rho^{\otimes l})) - \|\widetilde{\mathcal{M}}_l^{(j)}\circ\widetilde{\mathcal{P}}_l^{(i)}\otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_l) - \phi_l\otimes\psi_l'\|_1. \tag{73}$$

For the second term in (73) it holds by zero adding, triangle inequality, and monotonicity of the trace norm under action of partial traces

$$\|\widetilde{\mathcal{M}}_l^{(j)}\circ\widetilde{\mathcal{P}}_l^{(i)}\otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_l) - \phi_l\otimes\psi_l'\|_1 \leq \|\widetilde{\mathcal{M}}_l^{(j)}\otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_l), \phi_l\otimes(\psi')^{\otimes l}\|_1 + \|\widetilde{\mathcal{P}}_l^{(i)}(\rho^{\otimes l}) - \rho^{\otimes l}\|_1. \tag{74}$$

We further yield the bound

$$\|\widetilde{\mathcal{M}}_l^{(j)}\otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_l) - \phi_l\otimes\psi_l'\|_1 \tag{75}$$

$$\leq 2\left(1 - F(\widetilde{\mathcal{M}}_l^{(j)}\otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_l), \phi_l\otimes\psi_l')\right)^{\frac{1}{2}} \tag{76}$$

$$\leq 2\cdot 2^{-l\frac{\tilde{c}_i}{2}} \tag{77}$$

by (4) together with (57), and

$$\|\widetilde{\mathcal{P}}_l^{(i)}(\rho^{\otimes l}) - \rho^{\otimes l}\|_1 \;\leq\; 2\sqrt{1-\operatorname{tr}(\widetilde{\mathcal{P}}_l^{(i)}(\rho^{\otimes l}))} \;\leq\; 2\cdot 2^{-l\frac{c_3}{2}}, \tag{78}$$

where the first inequality is by Lemma 12, and the second inequality is valid due to the bound in (56) along with the fact that (because $p_{1,l},\ldots, p_{N,l}$ is a resolution of the identity into pairwise orthogonal projections)

$$1-\operatorname{tr}(\widetilde{P}^{(i)}(\rho^{\otimes l})) \;=\; \operatorname{tr}\left(\left(\mathrm{id}_{\mathcal{H}_{AB}^{\otimes l}} - \widetilde{\mathcal{P}}_l^{(i)}\right)(\rho^{\otimes l})\right) \;=\; \sum_{j\in[N]\setminus n(i)} \operatorname{tr}(\mathcal{P}_l^{(j)}(\rho^{\otimes l})) \tag{79}$$

holds. We define the constant c_4 by $c_4 := \min\{\tilde{c}_1,\ldots,\tilde{c}_N, c_3\}$. Combining (73) with (74)–(78) leads us to the estimate

$$F_m(\rho^{\otimes l}\widetilde{\mathcal{M}}_l^{(j)}\circ\widetilde{\mathcal{P}}_l^{(i)}) \geq \operatorname{tr}(\widetilde{\mathcal{P}}_l^{(i)}(\rho^{\otimes l})) - 4\cdot 2^{-l\frac{c_4}{2}} \tag{80}$$

$$\geq 1 - 5\cdot 2^{-l\frac{c_4}{2}} \tag{81}$$

for each $j \in n(i)$, where the last of the above inequalities, again is by the bound in (56). By inserting the bounds given in (72) and (81) into (68), we yield

$$F_m(\rho^{\otimes l}, \mathcal{M}_l) \geq |n(i)|(1-5\cdot 2^{-l\frac{c_4}{2}}) - (|n(i)|-1) \tag{82}$$

$$\geq 1 - 5|n(i)|\cdot 2^{-l\frac{c_4}{2}} \tag{83}$$

$$\geq 1 - 15\cdot 2^{-l\frac{c_4}{2}}. \tag{84}$$

If we now choose η small enough and assume l_0 large enough, to suffice

$$3\eta + \frac{\log N}{l_0} \leq \delta, \tag{85}$$

(58), (66), and (84) show that $\mathcal{M}_l$ has the desired properties for each $l > l_0$.

The assertion can be proven for the remaining case $\tilde{s} \geq 0$ by considering a compound set $\{\rho \otimes \phi_0 : \rho \in \mathcal{X}\}$ with a maximally entangled state ϕ_0 having Schmidt rank large enough to ensure $\sup_{\rho\in\mathcal{X}} S(A|B, \rho \otimes \phi_0) < 0$ and repeat the argument given above for the first case (note that $I(A; E, \rho \otimes \phi_0) = I(A; E, \rho)$ holds for each state $\rho \in \mathcal{X}$). □

Corollary 15. Asymptotically faithful $A \to B$-one-way quantum state merging of a compound source $\mathcal{X}$ is possible with (quantum) merging cost

$$C_{m,\to}(\mathcal{X}) = \sup_{\rho\in\mathcal{X}} S(A|B, \rho) \tag{86}$$

and classical cost

$$R_c(\mathcal{X}) = \sup_{\rho\in\mathcal{X}} I(A; E, \rho) \tag{87}$$

(again with the quantum mutual information evaluated on the AE marginal system of a purification $\psi \in \mathcal{S}(\mathcal{H}_{ABE})$ of ρ for each $\rho \in \mathcal{X}$).

Especially, the above lines show that the merging cost as well as the classical $A \to B$ communication cost exhibit regular behavior: If two nonempty sets $\mathcal{X}, \mathcal{X}'$ are near in the Hausdorff distance (see Sec. II for a definition), the costs will be nearly equal as well.

V. ENTANGLEMENT DISTILLATION FOR ARBITRARILY VARYING QUANTUM SOURCES

In this section, we prove a regularized formula for the one-way entanglement distillation capacity where the source is an AVQS generated by a set $\mathcal{X} \subset \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. We first prove the achievability part in case that $\mathcal{X}$ is finite, where we derive suitable one-way entanglement distillation protocols for the AVQS $\mathcal{X}$ from entanglement distillation protocols which are universal for the compound source $\mathrm{conv}(\mathcal{X})$ with fidelity approaching one exponentially fast. In a second step, we generalize this result allowing $\mathcal{X}$ to be any (not necessarily finite or countable) set on $\mathcal{H}_A \otimes \mathcal{H}_B$. To this end, we approximate $\mathcal{X}$ by a polytope (which is known to be the convex hull of a finite set of states), where we utilize methods we borrow from Ref. 4. First we state some facts concerning the continuity of the one-way entanglement distillation capacity functions.

A. Continuity of entanglement distillation capacities

Continuity was shown for the capacity functions appearing in coding theorems of several quantum channel coding scenarios,[16] here we state and prove uniform continuity for the entanglement distillation capacity functions.

Lemma 16. Let $\mathcal{Y}, \mathcal{Y}' \subset \mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_Y)$ be two nonempty sets of bipartite states with Hausdorff distance $0 \leq d_H(\mathcal{Y}, \mathcal{Y}') < \epsilon \leq \frac{1}{2}$. It holds for each $k \in \mathbb{N}$ and c.p.t.p. map $\mathcal{N}$ with domain $\mathcal{L}(\mathcal{H}_{XY}^{\otimes k})$

$$\left| \inf_{\tau\in\mathcal{Y}} I_c(X\rangle Y, \mathcal{N}(\tau^{\otimes k})) - \inf_{\sigma\in\mathcal{Y}'} I_c(X\rangle Y, \mathcal{N}(\sigma^{\otimes k})) \right| \leq k\nu(\epsilon), \tag{88}$$

where the function ν is defined by $\nu(x) := 4x \log \dim \mathcal{H}_X + 2h(x)$ for $x \in (0\frac{1}{2})$ and h being the binary entropy $h(x) := -x \log x - (1-x)\log(1-x)$.

Proof. We show this assertion with sets containing only one state defined $\mathcal{Y} := \{\tau\}, \mathcal{Y}' := \{\sigma\}$. The general assertion in (88) follows directly by definition of the Hausdorff distance. The argument parallels the one given in Ref. 16, Theorem 6 for continuity of the entropy exchange for channels.

Introduce a state $\gamma_{k,n} := \tau^{\otimes n} \otimes \sigma^{\otimes(k-n)}$ for each $0 \le n \le k$. By assumption, it holds

$$\|\gamma_{k,n-1} - \gamma_{k,n}\|_1 \le \epsilon \tag{89}$$

for each $0 < n \le k$, which implies via the Alicki-Fannes inequality[6] for the conditional von Neumann entropy

$$\left|I_c(X\rangle Y, \mathcal{N}(\gamma_{k,n-1})) - I_c(X\rangle Y, \mathcal{N}(\gamma_{k,n}))\right| \le \nu(\epsilon) \tag{90}$$

for each $0 < n \le k$ by (89) and monotonicity of the trace distance under action of $\mathcal{N}$. Further, it holds

$$\left|I_c(X\rangle Y, \mathcal{N}(\tau^{\otimes k})) - I_c(X\rangle Y, \mathcal{N}(\sigma^{\otimes k}))\right| \tag{91}$$

$$= \left|I_c(X\rangle Y, \mathcal{N}(\gamma_{k,k})) - I_c(X\rangle Y, \mathcal{N}(\gamma_{k,0}))\right| \tag{92}$$

$$= \left|\sum_{n=1}^{k} \left(I_c(X\rangle Y, \mathcal{N}(\gamma_{k,n-1})) - I_c(X\rangle Y, \mathcal{N}(\gamma_{k,n}))\right)\right| \tag{93}$$

$$\le \sum_{n=1}^{k} \left|I_c(X\rangle Y, \mathcal{N}(\gamma_{k,n-1})) - I_c(X\rangle Y, \mathcal{N}(\gamma_{k,n}))\right|, \tag{94}$$

where the first equality above is by definition, and the second by adding some zeros. Estimating each summand in (94) by (90) concludes the proof. □

Corollary 17. The one-way entanglement distillation capacity $D_\rightarrow$ for memoryless sources with perfectly known source state in (23) is a uniformly continuous function (considering the trace distance). Explicitly, it holds for $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\|\rho - \sigma\|_1 < \epsilon \le \frac{1}{2}$, it holds

$$|D_\rightarrow(\rho) - D_\rightarrow(\sigma)| \le \nu(\epsilon). \tag{95}$$

B. AVQS generated by finite sets

In this section, we assume $\mathcal{X}$ to be a finite set of bipartite states. We show that sequences of one-way entanglement distillation protocols for the compound source $\mathrm{conv}(\mathcal{X})$ with fidelity going to one exponentially fast can be modified to faithful entanglement distillation schemes for the AVQS $\mathcal{X}$. We apply Ahlswede's robustification[3] and elimination[1] techniques. This method of proof is well-known in classical information theory, and found application also in the quantum setting where it was shown to be a useful approach to determine the entanglement transmission capacity of arbitrarily varying quantum channels (AVQC).[4] Proposition 18 below is a generalization and sharpening of Lemma 12 in Ref. 7 required for our considerations. It asserts achievability of each rate below the one-way entanglement capacity for a compound source generated by a set $\mathcal{Y}$, where we drop the condition of finiteness imposed on $\mathcal{Y}$ in Ref. 7, Lemma 12. Moreover, we show, that each of these rates is achievable by protocols with fidelity approaching one exponentially fast.

Proposition 18. Let $\mathcal{Y} \subset \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a set of bipartite states. For each $k \in \mathbb{N}$, $\delta > 0$, there exist a number $l_0 = l_0(k, \delta)$ and a constant $c_5 = c_5(k, \delta, \mathcal{X}) > 0$, such that for each $l > l_0$, there exists an $A \rightarrow B$ LOCC $\mathcal{D}_l$ fulfilling

$$\inf_{\rho \in \mathcal{X}} F(\mathcal{D}_l(\rho^{\otimes l}), \phi_l) \ge 1 - 2^{-lc_5}, \tag{96}$$

where ϕ_l is a maximally entangled state shared by A and B with

$$\frac{1}{l} \log \mathrm{sr}(\phi_l) \ge \lim_{k\to\infty} \frac{1}{k} \sup_{\mathcal{T} \in \Theta_k} \inf_{\rho \in \mathcal{X}} D_\rightarrow^{(1)}(\rho^{\otimes k}, \mathcal{T}) - \delta. \tag{97}$$

The function $D_\rightarrow^{(1)}$ is defined in (24), and Θ_k is defined as in (25) for each $k \in \mathbb{N}$.

Proof. The line of proof is similar to that of the proofs given for Lemma 12 and Theorem 8 in Ref. 7 where we replace usage of Theorem 4 therein by the sharper and more general result Proposition 10 proven in Sec. IV above. We only briefly indicate the line of proof and restrict ourselves to the case $k = 1$. For other k, the argument is nearly the same. For given instrument $\mathcal{T}$ on $A's$ systems, and $\delta > 0$, we apply Proposition 10 to the set $\{\hat{\mathcal{T}}(\rho)\}_{\rho\in\mathcal{X}}$ (remember the notation introduced in (27)). In this way, we find for each large enough l an $A \to B$ LOCC $\mathcal{D}_l$ (incorporating $\hat{\mathcal{T}}$), such that

$$\inf_{\rho\in\mathcal{X}} F(\mathcal{D}_l(\rho^{\otimes l}), \phi_l) \geq 1 - 2^{-lc_5} \tag{98}$$

holds with a maximally entangled state ϕ_l with

$$\frac{1}{l}\log \mathrm{sr}(\phi_l) \geq -\sup_{\rho\in\mathcal{X}} S(A|BB', \hat{\mathcal{T}}(\rho)) - \frac{\delta}{2} \tag{99}$$

$$= \inf_{\rho\in\mathcal{X}} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho)) - \frac{\delta}{2}. \tag{100}$$

Since this can be done for each instrument $\mathcal{T}$ on A's site. Maximization over instruments on A's site shows the assertion. □

Proposition 18 allows to drop the finiteness condition on the compound generating set in Ref. 7, Theorem 8. We obtain the following corollary.

Corollary 19 (cf. Refs. 7 and 8). Let $\mathcal{Y} \subset \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

1. *It holds*

$$D_\to(\mathcal{Y}) = \lim_{k\to\infty} \frac{1}{k} \sup_{\mathcal{T}\in\Theta_k} \inf_{\rho\in\mathcal{Y}} D_\to^{(1)}(\rho^{\otimes k}, \mathcal{T}), \tag{101}$$

   where the set Θ_k is defined as in (25) for each $k \in \mathbb{N}$.
2. *The function in (101) behaves regular for compound sources in the following sense. If $\mathcal{Y}, \mathcal{Y}' \subset \mathcal{S}(\mathcal{H}_{AB})$ are two nonempty sets of bipartite states with $d_H(\mathcal{Y}, \mathcal{Y}') < \delta \leq \frac{1}{2}$, it holds*

$$|D_\to(\mathcal{Y}) - D_\to(\mathcal{Y}')| \leq \nu(\delta). \tag{102}$$

Proof. Achievability of the r.h.s. in (101) directly follows from Proposition 18. For the converse statement, we refer to the proof of Theorem 8 in Ref. 7 for finite sets of states. The argument given there directly carries over to the general case. It remains to show validity of the inequality in (102). Assume $d_H(\mathcal{Y}, \mathcal{Y}') < \delta \leq \frac{1}{2}$. Let $\tau > 0$ be an arbitrary but fixed number, and $\mathcal{Q}$ be an instrument with domain $\mathcal{L}(\mathcal{H}_A^{\otimes l})$, such that

$$\inf_{\rho\in\mathcal{Y}} I_c(A\rangle BB', \hat{\mathcal{Q}}(\rho^{\otimes l})) \geq \sup_{\mathcal{T}\in\Theta_k} \inf_{\rho\in\mathcal{Y}} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho^{\otimes l})) - \tau \tag{103}$$

holds, where we used our notation from (27). Lemma 16 implies

$$\inf_{\rho\in\mathcal{Y}} I_c(A\rangle BB', \hat{\mathcal{Q}}(\rho^{\otimes l})) \geq \inf_{\rho\in\mathcal{Y}'} I_c(A\rangle BB', \hat{\mathcal{Q}}(\rho^{\otimes l})) - k\nu(\delta), \tag{104}$$

which, together with (103), implies

$$\sup_{\mathcal{T}\in\Theta_k} \inf_{\rho\in\mathcal{Y}'} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho^{\otimes l})) \geq \sup_{\mathcal{T}} \inf_{\rho\in\mathcal{Y}} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho^{\otimes l})) - \tau - k\nu(\delta). \tag{105}$$

Since the above line of reasoning also holds with $\mathcal{Y}, \mathcal{Y}'$ interchanged and τ can be chosen arbitrarily small, we obtain

$$\left| \sup_{\mathcal{T}\in\Theta_k} \inf_{\rho\in\mathcal{Y}} D^{(1)}(\rho^{\otimes k}, \mathcal{T}) - \sup_{\mathcal{T}\in\theta_k} \inf_{\rho\in\mathcal{Y}'} D^{(1)}(\rho^{\otimes k}, \mathcal{T}) \right| \leq k\nu(\delta). \tag{106}$$

The above inequality together with the first assertion of the corollary proves the second one. □

The following theorem is the core of the robustification technique. It was first proven in Ref. 2. The version below (with a better constant) is from Ref. 3.

Theorem 20 (Robustification technique, cf. Theorem 6 in Ref. 3).

Let $\mathbf{S}$ *be a set with* $|\mathbf{S}| < \infty$ *and* $l \in \mathbb{N}$. *If a function* $f : \mathbf{S}^l \to [0, 1]$ *satisfies*

$$\sum_{s^l\in\mathbf{S}^l} f(s^l)q(s_1)\cdot\ldots\cdot q(s_l) \geq 1 - \gamma \tag{107}$$

for each type q *of sequences in* $\mathbf{S}^l$ *for some* $\gamma \in [0, 1]$, *then*

$$\frac{1}{l!}\sum_{\sigma\in\mathfrak{S}_l} f(\sigma(s^l)) \geq 1 - (l+1)^{|\mathbf{S}|}\cdot\gamma \qquad \forall s^l \in \mathbf{S}^l. \tag{108}$$

The following theorem is the main result of this section.

Theorem 21. *Let* $\mathcal{X} := \{\rho_s\}_{s\in\mathbf{S}} \subset \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $|\mathbf{S}| \leq \infty$. *For the AVQS generated by* $\mathcal{X}$, *it holds*

$$D_{\rightarrow}^{AV}(\mathcal{X}) \geq D_{\rightarrow}(\mathrm{conv}(\mathcal{X})) = \lim_{k\to\infty} \sup_{\mathcal{T}\in\Theta_k} \inf_{p\in\mathfrak{P}(\mathbf{S})} D^{(1)}(\mathcal{T}, \rho_p^{\otimes k}), \tag{109}$$

where we use the definition

$$\rho_p := \sum_{s\in\mathbf{S}} p^l(s^l)\,\rho_{s^l} \tag{110}$$

for each $p \in \mathfrak{P}(\mathbf{S})$.

Remark 22. The above statement actually holds with equality in (109) which we show in the proof of Corollary 25 below.

Proof. We show that each rate R, which is achievable for $A \to B$ entanglement distillation for the compound source generated by $\mathrm{conv}(\mathcal{X})$, is also an achievable rate for $A \to B$ entanglement distillation for the AVQS generated by $\mathcal{X}$. We indicate the elements of $\mathrm{conv}(\mathcal{X})$ by probability distributions on $\mathbf{S}$, since

$$\mathrm{conv}(\mathcal{X}) = \left\{ \rho_p : \ \rho_p = \sum_{s\in\mathbf{S}} p(s)\rho_s, \ p \in \mathfrak{P}(\mathbf{S}) \right\} \tag{111}$$

holds. We know from Proposition 18 that for an achievable $A \to B$ entanglement distillation rate R for the compound source generated by $\mathrm{conv}(\mathcal{X})$, $\delta > 0$ and each sufficiently large blocklength l, there exists a one-way LOCC channel $\tilde{\mathcal{D}}_l$, such that the condition

$$\min_{p\in\mathfrak{P}(\mathbf{S})} F(\tilde{\mathcal{D}}_l(\rho_p^{\otimes l}), \phi_l) \geq 1 - 2^{-lc_5} \tag{112}$$

is fulfilled with a maximally entangled state ϕ_l shared by A and B, such that

$$\frac{1}{l}\log \mathrm{sr}(\phi_l) \geq R - \delta \tag{113}$$

holds. Note that the minimization in (112) is because of (111). We define a function $f : \mathbf{S}^l \to [0, 1]$ by $f(s^l) := F(\tilde{\mathcal{D}}_l(\rho_{s^l}), \phi_l)$ for each $s^l \in \mathbf{S}^l$, and infer from (112) that

$$\sum_{s^l \in \mathbf{S}^l} p(s_1) \cdot \ldots \cdot p(s_l)\ f(s^l)\ \geq 1 - 2^{-lc_5} \tag{114}$$

holds for each $p \in \mathfrak{P}(\mathbf{S})$ with a constant $c_5 > 0$. Let

$$\mathcal{U}_\sigma(\cdot) := U_{A,\sigma} \otimes U_{B,\sigma}(\cdot) U_{A,\sigma}^* \otimes U_{B,\sigma}^*, \tag{115}$$

for each permutation $\sigma \in \mathfrak{S}_l$, be the unitary channel, which permutes the tensor factors in $\mathcal{H}_{AB}^{\otimes l}$ according to σ (with unitary matrices $U_{A,\sigma}, U_{B,\sigma}$ permuting the tensor bases on $\mathcal{H}_A^{\otimes l}$ resp. $\mathcal{H}_B^{\otimes l}$). It holds

$$\rho_{\sigma(s^l)} = \mathcal{U}_\sigma(\rho_{s^l}), \tag{116}$$

and consequently

$$f(\sigma(s^l)) = F(\tilde{\mathcal{D}}_l \circ \mathcal{U}_\sigma(\rho_{s^l}), \phi_l) \tag{117}$$

for each $s^l \in \mathbf{S}^l$, $\sigma \in \mathfrak{S}_l$. The functions in (117) fulfill the conditions of Theorem 20, which in turn implies that

$$(1 - (l+1)^{|\mathbf{S}|}) \cdot 2^{-lc_5} \leq \frac{1}{l!} \sum_{\sigma \in \mathfrak{S}_l} F(\tilde{\mathcal{D}}_l \circ \mathcal{U}_\sigma(\rho_{s^l}), \phi_l) \tag{118}$$

$$= F(\hat{\mathcal{D}}_l(\rho_{s^l}), \phi_l) \tag{119}$$

is valid with the definition $\hat{\mathcal{D}}_l := \frac{1}{l!} \sum_{\sigma \in \mathfrak{S}_l} \tilde{\mathcal{D}}_l \circ \mathcal{U}_\sigma$. Notice that $\hat{\mathcal{D}}_l$ is an $A \to B$ LOCC channel either. However, $\hat{\mathcal{D}}_l$ is not a reasonable protocol for entanglement distillation regarding the classical communication cost. Implementation of $\hat{\mathcal{D}}_l$ demands $A \to B$ communication of a number of classical messages increased by a factor $l!$ compared to the requirements of $\tilde{\mathcal{D}}_l$, which leads to super-exponential growth of required classical messages and consequently unbounded classical communication rates. We remark here that for a coordination of the permutations in $\hat{\mathcal{D}}_l$, common randomness accessible to A and B, which is known to be a weaker resource than $A \to B$ communication, would suffice. Nonetheless, the asymptotic common randomness consumption of the protocol would be above any rate either. We will apply the well-known derandomization technique which first appeared in Ref. 1 to construct $A \to B$ LOCC channel with reasonable classical communication requirements (actually, we will show that we can approximate the classical cost of $A \to B$ distillation of the compound source conv$(\mathcal{X})$.

Let $X_1, \ldots, X_{K_l}$ be a sequence of i.i.d. random variables, each distributed uniformly on $\mathfrak{S}_l$. We define a function $g : \mathfrak{S}_l \times \mathbf{S}^l \to [0, 1]$ by

$$g(\sigma, s^l) = 1 - F(\tilde{\mathcal{D}}_l \circ \mathcal{U}_\sigma(\rho_{s^l}), \phi_l) \qquad (\sigma \in \mathfrak{S}_l, s^l \in \mathbf{S}^l). \tag{120}$$

One readily verifies that

$$\mathbb{E}\left[g(X_1, s^l)\right] = 1 - F(\hat{\mathcal{D}}_l(\rho_{s^l}), \phi_l) \leq (l+1)^{|\mathbf{S}|}\, 2^{-lc_5} := \epsilon_l \tag{121}$$

holds for each $s^l \in \mathbf{S}^l$. Thus, for each $s^l \in \mathbf{S}^l$, and $\nu_l \in (0, 1)$, we yield

$$\Pr\left(\sum_{k=1}^{K_l} g(X_k, s^l) > K_l \nu_l\right) = \Pr\left(\prod_{k=1}^{K_l} \exp(g(X_k, s^l)) > 2^{K_l \nu_l}\right) \tag{122}$$

$$\leq 2^{-K_l \nu_l} \cdot \mathbb{E}\left[\exp(g(X_k, s^l)\right]^{K_l} \tag{123}$$

$$\leq 2^{-K_l \nu_l} \cdot (1 + \mathbb{E}\left[\exp(g(X_k, s^l)\right])^{K_l} \tag{124}$$

$$\leq 2^{-K_l \nu_l} \cdot 2^{K_l \log(1+\epsilon_l)} \tag{125}$$

$$\leq 2^{-K_l(\nu_l - 2\epsilon_l)}. \tag{126}$$

Equation (123) above is by Markov's inequality, (124) follows from the fact that $\exp(x) \leq 1 + x$ holds for $x \in [0, 1]$, (125) is by (121), and (126) follows from the inequality $\log(1 + x) \leq 2x$ being valid for $x \in (0, 1)$. From (122)–(126) and application of de Morgan's laws, it follows

$$\Pr\left(\forall s^l \in \mathbf{S}^l : \frac{1}{K_l}\sum_{k=1}^{K_l} g(X_k, s^l) \leq \nu_l\right) \geq 1 - |\mathbf{S}|^l \cdot 2^{-K_l(\nu_l - 2\epsilon_l)} \tag{127}$$

$$\geq 1 - 2^{-l\frac{(\theta-\kappa)}{2}}, \tag{128}$$

for large enough l, where the last line results from the choosing $\nu_l = 2^{-l\kappa}$ and $K_l = 2^{l\theta}$ with $\theta, \kappa > 0$. If we choose κ and θ in a way that they fulfill $0 < \kappa < \theta < c$, the r.h.s. of (128) is strictly positive and we find a realization $\sigma_1, \ldots, \sigma_{K_l}$ of $X_1, \ldots, X_{K_l}$, such that for each $s^l \in \mathbf{S}^l$

$$2^{-l\kappa} \geq \frac{1}{K_l}\sum_{k=1}^{K_l} g(\sigma_k, s^l) \tag{129}$$

$$= 1 - \frac{1}{K_l}\sum_{k=1}^{K_l} F(\tilde{\mathcal{D}}_l \circ \mathcal{U}_{\sigma_k}(\rho_{s^l}), \phi_l) \tag{130}$$

$$= 1 - F(\mathcal{D}_l(\rho_{s^l}), \phi_l), \tag{131}$$

where we defined $\mathcal{D}_l := \frac{1}{K_l}\sum_{k=1}^{K_l} \tilde{\mathcal{D}}_l \circ \mathcal{U}_{\sigma_k}$. With (113) and (131), it is shown that for each sufficiently large blocklength l, we find a one-way entanglement distillation protocol with

$$\min_{s^l \in \mathbf{S}^l} F(\mathcal{D}_l(\rho_{s^l}), \phi_l) \geq 1 - 2^{-l\kappa}, \quad \text{and} \quad \frac{1}{l}\log \mathrm{sr}(\phi_l) \geq R - \delta. \tag{132}$$

Notice that the number of different classical messages to be communicated by A within application of $\mathcal{D}_l$ is increased by a factor $2^{l\theta}$ compared to the message transmission demanded by $\tilde{\mathcal{D}}_l$, i.e., the communication rate is increased by θ (which we can choose to be an arbitrarily small fixed number). □

C. General AVQS

In this section, we generalize the results of Sec. V B, admitting the AVQS to be generated by any not necessarily finite or countable set $\mathcal{X}$ of states on $\mathcal{H}_A \otimes \mathcal{H}_B$. We approximate the closed convex hull of $\mathcal{X}$ by a polytope, which is known as the convex hull of a finite set of points and apply Theorem 21, together with continuity properties of the capacity function. The proof strategy has some similarities with the argument given in Ref. 4 for entanglement transmission over general arbitrarily varying quantum channels. To prepare ourselves for the approximation, we need some notation and results from convex geometry which we state first. For a subset A of a normed space $(V, \|\cdot\|)$, $\overline{A}$ is the closure and $\mathrm{aff}A$ is the affine hull of A. If A is a convex set, the relative interior $\mathrm{ri}A$ is the interior and the relative boundary $\mathrm{rebd}A$ of A are the interior and boundary of A regarding the topology on $\mathrm{aff}A$ induced by $\|\cdot\|$.

Lemma 23 (Ref. 4, Lemma 34). Let A, B be compact sets in $\mathbb{C}^n$ with $A \subset B$ and

$$d_H(\mathrm{rebd}B, A) = t > 0, \tag{133}$$

where $\|\cdot\|$ denotes any norm on $\mathbb{C}^n$. Let P a polytope with $A \subset P$ and $d_H(A, P) \leq \delta$, where $\delta \in (0, t]$ and d_H is the Hausdorff distance induced by $\|\cdot\|$. Then $P' := P \cap \operatorname{aff} A$ is also a polytope and $P \subset B$.

With the above statement and the assertions Sec. V B, we are prepared to prove the following theorem which is the main result of this section.

Theorem 24. *Let $\mathcal{X} := \{\rho_s\}_{s\in\mathbf{S}}$ be a set of states on $\mathcal{H}_A \otimes \mathcal{H}_B$. For each $\delta > 0$ and $k \in \mathbb{N}$, there exists a number $l_0 \in \mathbb{N}$, such that for each $l > l_0$, there is an $A \to B$ LOCC channel $\mathcal{D}_l$ fulfilling*

$$\inf_{s\in\mathbf{S}^l} F(\mathcal{D}_l(\rho_{s^l}), \phi_l) \geq 1 - 2^{-lc_6} \tag{134}$$

with a maximally entangled state ϕ_l shared by A and B and a constant $c_6 > 0$, such that

$$\frac{1}{l}\log \operatorname{sr}(\phi_l) \geq \frac{1}{k}\sup_{\mathcal{T}\in\Theta_k}\inf_{\tau\in\operatorname{conv}(\mathcal{X})} D_{\rightarrow}^{(1)}(\tau^{\otimes k}, \mathcal{T}) - \delta \tag{135}$$

holds, where the function $D_{\rightarrow}^{(1)}$ is defined in (24).

Proof. Let $\mathcal{T} := \{\mathcal{T}_j\}_{j=1}^J$ be any instrument with domain $\mathcal{L}(\mathcal{H}_A^{\otimes k})$, $\delta > 0$. Dealing only with the nontrivial case, we show that

$$\inf_{\rho\in\operatorname{conv}(\mathcal{X})} \frac{1}{k} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho^{\otimes k})) - \delta > 0 \tag{136}$$

is an achievable rate (remember our notation from (27)). Since the Hausdorff distance between $\operatorname{conv}(\mathcal{X})$ and $\overline{\operatorname{conv}(\mathcal{X})}$ is zero, it makes no difference if we consider the set $\overline{\operatorname{conv}(\mathcal{X})}$ instead. We briefly describe the strategy of our proof. We approximate the set $\overline{\operatorname{conv}(\mathcal{X})}$ from the outside by a polytope P_η. Since P_η, as a polytope, is the convex hull of a finite set of points, Theorem 21 can be applied. A technical issue (cf. Ref. 4) is to ensure that the approximating polytope completely consists of density matrices, i.e., $P_\eta \subset \mathcal{S}(\mathcal{H}_{AB})$. We achieve this by a slight depolarization of the states in $\operatorname{conv}(\mathcal{X})$, such that the resulting set does not touch the boundary of $\mathcal{S}(\mathcal{H}_{AB})$. Define, for $\gamma \in [0, 1]$, the channel $\mathcal{N}_\gamma \in \mathcal{C}(\mathcal{H}_A \otimes \mathcal{H}_B)$ by $\mathcal{N}_\gamma := \mathcal{N}_{A,\gamma} \otimes \mathcal{N}_{B,\gamma}$, where $\mathcal{N}_{X,\gamma}$ is the γ-depolarizing channel on the subsystem X, $X = A, B$, i.e.,

$$\mathcal{N}_\gamma(\tau) = (1-\gamma)^2\tau + \gamma(1-\gamma)(\tau_A \otimes \pi_B + \pi_A \otimes \tau_B) + \gamma^2(\pi_A \otimes \pi_B) \tag{137}$$

for each $\tau \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, where π_A, π_B are maximally mixed states and τ_A, τ_B are the marginals of τ on $\mathcal{H}_A, \mathcal{H}_B$. Notice that $\mathcal{N}_\gamma$ is defined in terms of local depolarizing channels on the subsystems. This is required, since we are restricted to one-way LOCC channels. It holds

$$\|\mathcal{N}_\eta(\tau) - \tau\|_1 \leq \|(1-\eta)^2\tau - \tau\|_1 + \eta(1-\eta)\|\tau_A \otimes \pi_B + \pi_A \otimes \tau_B\|_1 \tag{138}$$

$$+ \eta\|\pi_A \otimes \pi_B\|_1 \tag{139}$$

$$\leq 6\eta \tag{140}$$

for each state τ on $\mathcal{H}_A \otimes \mathcal{H}_B$. Moreover, it holds $\overline{\mathcal{N}_\eta(\operatorname{conv}(\mathcal{X}))} = \mathcal{N}_\eta(\overline{\operatorname{conv}(\mathcal{X})}) \subset \operatorname{ri}\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, which implies

$$\inf\left\{\|\rho - \rho'\|_1 : \ \rho \in \overline{\mathcal{N}_\eta(\operatorname{conv}(\mathcal{X}))}, \rho' \in \operatorname{rebd}(\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B))\right\} > 0. \tag{141}$$

Therefore, due to of Lemma 23 and Theorem 3.1.6 in Ref. 18, there exists, for each small enough number $\eta > 0$, a polytope $P_\eta := \operatorname{conv}(\{\tau_e\}_{e\in E_\eta}) \subset \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ such that $\mathcal{N}_\eta(\operatorname{conv}(\mathcal{X})) \subset P_\eta$ and

$$d_H(\mathcal{N}_\eta(\operatorname{conv}(\mathcal{X})), P_\eta) \leq \eta. \tag{142}$$

Applying Theorem 21 to the finite AVQS generated by the extremal set $\{\tau_e\}_{e\in E}$ of the polytope P_η, we know that for each sufficiently large blocklength l, there exists an $A \to B$ LOCC channel $\hat{\mathcal{D}}_l$

such that

$$F(\hat{\mathcal{D}}_l(\tau_{e^l}), \phi_l) \geq 1 - 2^{-lc_6} \tag{143}$$

holds with a maximally entangled state ϕ_l shared by A and B for each $e^l \in E^l$ with Schmidt rank fulfilling

$$\frac{1}{l} \log \mathrm{sr}(\phi_l) \geq \frac{1}{k} \inf_{\tau \in P_\eta} I_c(A\rangle BB', \hat{\mathcal{T}}(\tau^{\otimes k})) - \frac{\delta}{2}. \tag{144}$$

Since $\mathcal{N}_\eta(\mathrm{conv}(\mathcal{X})) \subset P_\eta$ holds, the depolarized version $\mathcal{N}_\eta(\rho_s)$ of each state ρ_s, $s \in \mathbf{S}$ can be written as a convex combination of elements from $\{\tau_e\}_{e \in E_\eta}$, i.e.,

$$\mathcal{N}_\eta(\rho_s) = \sum_{e \in E_\eta} q(e|s)\, \tau_e \tag{145}$$

with a probability distribution $q(\cdot|s)$ on E_η for each $s \in \mathbf{S}$. We define a one-way LOCC channel $\mathcal{D}_l$ by $\mathcal{D}_l := \hat{\mathcal{D}}_l \circ \mathcal{N}_\eta^{\otimes l}$ and deduce

$$F(\mathcal{D}_l(\rho_{s^l}), \phi_l) = F(\hat{\mathcal{D}}_l(\mathcal{N}_\eta^{\otimes l}(\rho_{s^l})), \phi_l) \tag{146}$$

$$= F\left(\hat{\mathcal{D}}_l\left(\bigotimes_{i=1}^{l} \mathcal{N}_\eta(\rho_{s_i})\right), \phi_l\right) \tag{147}$$

$$= F\left(\hat{\mathcal{D}}_l\left(\bigotimes_{i=1}^{l} \sum_{e_i \in E} q(e_i|s_i)\tau_{e_i}\right), \phi_l\right) \tag{148}$$

$$= \sum_{e_1 \in E} \cdots \sum_{e_l \in E} \prod_{i=1}^{l} p(e_i|s_i) F\left(\hat{\mathcal{D}}_l(\tau_{e_i}), \phi_l\right) \tag{149}$$

$$= \sum_{e^l \in E_\eta^l} q^l(e^l|s^l)\, F(\hat{\mathcal{D}}_l(\tau_{e^l}), \phi_l) \tag{150}$$

$$\geq 1 - 2^{-lc_6} \tag{151}$$

for each $s^l = (s_1, \ldots, s_l) \in \mathbf{S}^l$ where we used (145) in (148) and (151) is by (143). To complete the proof, we show that for small enough η,

$$\inf_{\rho \in \mathrm{conv}(\mathcal{X})} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho^{\otimes k})) \geq \inf_{\tau \in P_\eta} I_c(A\rangle BB', \hat{\mathcal{T}}(\tau^{\otimes k})) - \frac{k\delta}{2} \tag{152}$$

holds. For each $\rho \in \mathrm{conv}(\mathcal{X})$, $\tau \in P_\eta$, we have

$$\|\rho - \tau\|_1 \leq \|\rho - \mathcal{N}_\eta(\rho)\|_1 + \|\mathcal{N}_\eta(\rho) - \tau\|_1 \tag{153}$$

$$\leq 6\eta + \|\mathcal{N}_\eta(\rho) - \tau\|_1, \tag{154}$$

where the last estimation is by (140). From (154), we can conclude that

$$d_H(\mathrm{conv}(\mathcal{X}), P_\eta) \leq d_H(\mathcal{N}_\eta(\mathrm{conv}(\mathcal{X}), P_\eta) + 6\eta \leq 7\eta \tag{155}$$

holds, which implies, via Lemma 16,

$$\left| \inf_{\rho \in \mathrm{conv}(\mathcal{X})} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho^{\otimes k})) - \inf_{\tau \in P_\eta} I_c(A\rangle BB', \hat{\mathcal{T}}(\tau^{\otimes k})) \right| \leq k\nu(7\eta). \tag{156}$$

If now η is chosen small enough to ensure $\nu(7\eta) < \frac{\delta}{2}$, (152), and we conclude, collecting inequalities, that the entanglement rate of $\mathcal{D}_l$ is

$$\frac{1}{l}\log \mathrm{sr}(\phi_l) \geq \frac{1}{k}\inf_{\tau\in P_\eta} I_c(A\rangle BB', \hat{\mathcal{T}}(\tau^{\otimes k})) - \frac{\delta}{2} \tag{157}$$

$$\geq \frac{1}{k}\inf_{\rho\in\mathrm{conv}(\mathcal{X})} I_c(A\rangle BB', \hat{\mathcal{T}}(\rho^{\otimes k})) - \delta, \tag{158}$$

where (157) is (144), (158) is by (152). □

Corollary 25. Let $\mathcal{X}$ be a set of states on $\mathcal{H}_A \otimes \mathcal{H}_B$. For the AVQS generated by $\mathcal{X}$, it holds

$$D_{\rightarrow}^{AV}(\mathcal{X}) = D_{\rightarrow}(\mathrm{conv}(\mathcal{X})) = \lim_{l\to\infty}\frac{1}{k}\sup_{\mathcal{T}\in\Theta_k}\inf_{\tau\in\mathrm{conv}(\mathcal{X})} D_{\rightarrow}^{(1)}(\tau^{\otimes k}, \mathcal{T}) \tag{159}$$

with $D_{\rightarrow}^{(1)}$ being the function defined in (24), and maximization over instruments on A's systems.

Proof. The rightmost equality in (159) is Corollary 19.1. We prove the first equality. Achievability directly follows from Theorem 24. For the converse statement, let $\mathcal{X} := \{\rho_s\}_{s\in\mathbf{S}}$ and $\sigma \in \mathrm{conv}(\mathcal{X})$. By Carathéordory's theorem (see, e.g., Ref. 18, Theorem 2.2.4.), σ can be written as a finite convex combination of elements of $\mathcal{X}$, say

$$\sigma = \sum_{s\in\mathbf{S}'} p(s)\rho_s, \tag{160}$$

with $|\mathbf{S}'| < \infty$. Thus, for an $A \to B$ LOCC channel $\mathcal{D}_l$ for blocklength l with suitable maximally entangled state ϕ_l, it holds

$$\inf_{s^l\in\mathbf{S}^l} F(\mathcal{D}_l(\rho_{s^l}), \phi_l) \leq \min_{s^l\in\mathbf{S}'^l} F(\mathcal{D}_l(\rho_{s^l}), \phi_l) \tag{161}$$

$$\leq \sum_{s^l\in\mathbf{S}'^l} p^l(s^l) F(\mathcal{D}_l(\rho_{s^l}), \phi_l) \tag{162}$$

$$= F(\mathcal{D}_l(\sigma^{\otimes l}), \phi_l). \tag{163}$$

Since (163) holds for each element of $\mathrm{conv}(\mathcal{X})$, each rate R which is an $A \to B$ achievable entanglement distillation rate for the AVQS generated by $\mathcal{X}$ is also achievable for the compound quantum source generated by $\mathrm{conv}(\mathcal{X})$, thus the converse statement in Corollary 19.1 applies. □

Having determined the one-way entanglement distillation capacity $D_{\rightarrow}^{AV}$, the continuity properties of the capacity function on the r.h.s. of (156) imply the following corollary.

Corollary 26. Identifying each set of states with its closure, $D_{\rightarrow}^{AV}$ is uniformly continuous in the metric defined by the Hausdorff distance on compact sets of density matrices. If $\mathcal{X}, \mathcal{X}' \subset \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ are two compact sets with $d_H(\mathcal{X}', \mathcal{X}) < \epsilon \leq \frac{1}{2}$ it holds

$$|D_{\rightarrow}^{AV}(\mathcal{X}') - D_{\rightarrow}^{AV}(\mathcal{X})| \leq \nu(\epsilon). \tag{164}$$

Remark 27. Corollary 26 classifies the AVQS one-way entanglement distillation task as well-behaved in the following sense. Two different AVQS with generating sets being near in the Hausdorff sense will have approximately equal capacities.

An example for a situation where "capacity" is a more fragile quantity is transmission of classical messages over an arbitrarily varying quantum channel. The capacity C_{random} for classical message transmission using correlated random codes is continuous, while it can be shown that in some cases, the capacity using deterministic codes, C_{det}, is discontinuous on certain points.[8]

VI. ON QUANTUM STATE MERGING FOR AVQS

In this section, we consider quantum state merging in case that the bipartite sources A and B have to merge is an AVQS. In Ref. 7 and Sec. IV of this paper, we have determined the optimal entanglement as well as classical communication cost in case of a compound quantum source, and achieved these rates by protocols with merging fidelity going to one exponentially fast (see Sec. IV). Therefore, one would expect that we can proceed as we did for one-way entanglement distillation for AVQS in Sec. V and build state merging protocols for the AVQS generated by a set $\mathcal{X}$ from suitable protocols for the compound source generated by $\overline{\text{conv}(\mathcal{X})}$. One would expect to be able to prove the equalities,

$$C_{m,\to}^{AV}(\mathcal{X}) = C_{m,\to}(\text{conv}(\mathcal{X})) = \sup_{\rho\in\text{conv}(\mathcal{X})} S(A|B,\rho), \tag{165}$$

to hold, where again, no difference has to be made between $\text{conv}(\mathcal{X})$ and its closure, because of continuity of the conditional von Neumann entropy. Actually, it seems possible to prove the relation

$$C_{m,\to}^{AV}(\mathcal{X}) \leq C_{m,\to}(\text{conv}(\mathcal{X})) \tag{166}$$

using Ahlswede's elimination and derandomization techniques (at least if the AVQS is generated by a finite set of states). We do not carry out the argument here. Instead, we give a simple counterexample to the relation in (165).

Consider a finite set $\hat{\mathcal{X}} := \{\rho_s\}_{s=1}^N$ of bipartite states on a Hilbert space $\mathcal{H}_A\otimes\mathcal{H}_B$, which is generated by unitaries in the following sense. Let $\rho_1\in\mathcal{S}(\mathcal{H}_A\otimes\mathcal{H}_B)$, where we assume $S(A|B,\rho_1)<0$ and $\dim\mathcal{H}_A \geq N\cdot\dim\text{supp}(\rho_{A,1})$, $U_1 = \mathbf{1}_{\mathcal{H}_A}$ and $U_2,\ldots,U_N$ unitaries on $\mathcal{H}_A$ such that with the definitions

$$\rho_s := U_s\otimes\mathbf{1}_{\mathcal{H}_B}(\rho_1)U_s^*\otimes\mathbf{1}_{\mathcal{H}_B} \qquad (s\in[N]), \tag{167}$$

the supports of the A-marginals are pairwise orthogonal, i.e.,

$$\text{supp}(\rho_{A,s}) \perp \text{supp}(\rho_{A,s'}) \qquad (s,s'\in[N], s'\neq s). \tag{168}$$

Note, that our definitions also imply the relations $\rho_{B,s}=\rho_{B,1}$ $(s\in[N])$ and

$$\text{supp}(\rho_s)\perp\text{supp}(\rho_{s'}) \qquad (s,s'\in[N], s\neq s'). \tag{169}$$

In the following we show that sets constructed in the above described manner are counterexamples to (165) if $N>1$.

Example 28. For the AVQS generated by $\hat{\mathcal{X}} := \{\rho_s\}_{s=1}^N$, it holds

$$C_{m,\to}^{AV}(\hat{\mathcal{X}}) = C_{m\to}\left(\text{conv}(\hat{\mathcal{X}})\right) - \log N. \tag{170}$$

The classical $A\to B$ communication cost for merging of the AVQS $\hat{\mathcal{X}}$ is upper bounded by

$$\sup_{\sigma\in\text{conv}(\hat{\mathcal{X}})} I(A;E,\sigma) - \log N, \tag{171}$$

where $\rho_p := \sum_{s=1}^N p(s)\rho_s$ for each $p\in\mathfrak{P}([N])$.

Proof of Example 28. Before we prove the claims made in the example, we briefly sketch the argument. Since the A marginals are supported on pairwise orthogonal subspaces, A can perfectly detect, given a block of l outputs of the AVQS, which of the $s^l\in\mathbf{S}^l$ is actually realized. In this way, A obtains state knowledge which helps to achieve the desired rates. We introduce unitary channels $\mathcal{V}_{A,1},\ldots,\mathcal{V}_{A,N}$ and $\mathcal{V}_{B',1},\ldots,\mathcal{V}_{B',N}$ where we define $\mathcal{V}_{A,s}(\cdot) := U_s(\cdot)U_s^*$ with the unitaries from (167) and consider $\mathcal{V}_{B',s}$ to be the corresponding unitary channel on the space $\mathcal{H}_{B'}$ for each $s\in[N]$. For given blocklength l, we define unitary channels

$$\mathcal{V}_{A,s^l}(\cdot) := \mathcal{V}_{A,s_1}\otimes\ldots\otimes\mathcal{V}_{A,s_l} \quad\text{and}\quad \mathcal{V}_{B',s^l}(\cdot) := \mathcal{V}_{B',s_1}\otimes\ldots\otimes\mathcal{V}_{B',s_l} \tag{172}$$

for each $s^l = (s_1, \dots, s_l) \in \mathbf{S}^l$ accordingly. Thus, the definitions in (167) imply

$$\rho_{s^l} = \mathcal{V}_{A,s^l} \otimes \mathrm{id}_{\mathcal{H}_B^{\otimes l}}(\rho_1^{\otimes l}). \qquad (s^l \in [N]^l). \tag{173}$$

Using the projection P_s onto the support of $\rho_{A,s}$ for each $s \in [N]$, we define a quantum instrument

$$\hat{\mathcal{A}} := \{\hat{\mathcal{A}}_s\}_{s=1}^N \tag{174}$$

with $\hat{\mathcal{A}}_s(\cdot) := \mathcal{U}_{A,s} \circ P_s(\cdot)P_s^*$ for each $s \in [N]$, which implies

$$\hat{\mathcal{A}}_{s'} \otimes \mathrm{id}_{\mathcal{H}_B}(\rho_s) = \delta_{ss'}\rho_1 \qquad (s \in [N]). \tag{175}$$

It is known from Ref. 14 and Sec. IV that for each $\delta > 0$ and sufficiently large blocklength l, there exists an $(l, k_l, \tilde{D}_l)$ $A \to B$ merging $\tilde{\mathcal{M}}_l$ such that

$$F(\tilde{\mathcal{M}}_l \otimes \mathrm{id}_{\mathcal{H}_E^{\otimes n}}(\psi_1^{\otimes l}), \phi_l \otimes \psi_1'^{\otimes l}) \geq 1 - 2^{-lc} \tag{176}$$

holds with a constant $c > 0$, where ψ_1 is a purification of ρ_1 and ϕ_l a maximally entangled state shared by A and B with

$$-\frac{1}{l}\log \mathrm{sr}(\phi_l) \leq S(A|B, \rho_1) + \delta \tag{177}$$

and where for the classical communication rate

$$\frac{1}{l}\log \tilde{D}_l \leq I(A; E, \rho_1) + \delta \tag{178}$$

holds. We combine the instrument $\hat{\mathcal{A}}$ and the unitary channels from (172) with $\tilde{\mathcal{M}}_l$ to build a merging LOCC $\mathcal{M}_l$ suitable for merging the AVQS generated by $\hat{\mathcal{X}}$ and define

$$\mathcal{M}_l := \sum_{s^l \in [N]^l} (\mathcal{V}_{B',s^l} \otimes \mathrm{id}_{\mathcal{H}_B^{\otimes l}}) \circ \tilde{\mathcal{M}}_l \circ (\hat{\mathcal{A}}_{s^l} \otimes \mathrm{id}_{\mathcal{H}_B^{\otimes l}}). \tag{179}$$

Clearly, $\mathcal{M}_l$ is an $A \to B$ LOCC channel. Explicitly, inspection of the above definition shows that $\mathcal{M}_l$ is an (l, k_l, D_l) $A \to B$ merging where one of

$$D_l = \tilde{D}_l \cdot N^l \tag{180}$$

different classical messages has to be communicated within action of $\mathcal{M}_l$. Moreover, for each $s^l \in [N]^l$, it holds

$$F(\mathcal{M}_l \otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_{s^l}), \phi_l \otimes \psi'_{s^l}) \tag{181}$$

$$\overset{(a)}{=} \sum_{m^l \in [N]^l} F((\mathcal{V}_{B',m^l} \otimes \mathrm{id}_{\mathcal{H}_B^{\otimes l}}) \circ \tilde{\mathcal{M}}_l \circ (\hat{\mathcal{A}}_{m^l} \otimes \mathrm{id}_{\mathcal{H}_B^{\otimes l}}) \otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_{s^l}), \phi_l \otimes \psi'_{s^l}) \tag{182}$$

$$\overset{(b)}{=} F(\tilde{\mathcal{M}}_l \otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_1^{\otimes l}), \phi_l \otimes (\mathcal{V}^*_{B',s^l} \otimes \mathrm{id}_{\mathcal{H}_{BE}^{\otimes l}})(\psi'_{s^l})) \tag{183}$$

$$\overset{(c)}{=} F(\tilde{\mathcal{M}}_l \otimes \mathrm{id}_{\mathcal{H}_E^{\otimes l}}(\psi_1^{\otimes l}), \phi_l \otimes (\psi'_1)^{\otimes l}) \tag{184}$$

$$\overset{(d)}{\geq} 1 - 2^{-lc}, \tag{185}$$

where (a) is the definition of $\mathcal{M}_l$ plus linearity of the fidelity in the first argument in the present situation, (b) is because

$$\hat{\mathcal{A}}_{m^l} \otimes \mathrm{id}_{\mathcal{H}_{BE}^{\otimes l}}(\psi_{s^l}) = \delta_{m^l s^l}\psi_1^{\otimes l} \tag{186}$$

holds implied by (175) together with the fact that the fidelity is invariant under action of unitary channels applied simultaneously on both arguments. Equality (c) follows from (167), and (d) is by

(176). It remains to evaluate the rates. It is well known that for each ensemble $\{q(x), \rho_x\}_{x\in\mathbf{X}}$ of quantum states having pairwise orthogonal supports, it holds

$$S\left(\sum_{x\in\mathbf{X}} q(x)\rho_x\right) = \sum_{x\in\mathbf{X}} q(x)S(\rho_x) + H(q), \tag{187}$$

where $H(q)$ is the Shannon entropy of q. Thus, for each $p \in \mathfrak{P}([N])$, $\rho_p := \sum_{s\in[N]} p(s)\rho_s$ we yield

$$S(A|B, \rho_p) = S(\rho_p) - S(\rho_{B,p}) \tag{188}$$

$$= \sum_{s\in[N]} p(s)S(\rho_s) + H(p) - S(\rho_{B,1}) \tag{189}$$

$$= S(A|B, \rho_1) + H(p) \tag{190}$$

and

$$I(A; E, \rho_p) = S(\rho_{A,p}) + S(A|B, \rho_p) \tag{191}$$

$$= \sum_{s\in[N]} p(s)S(\rho_{A,s}) + S(A|B, \rho_1) + 2H(p) \tag{192}$$

$$= I(A; E, \rho_1) + 2H(p). \tag{193}$$

Taking maxima over all $p \in \mathfrak{P}([N])$ and rearranging equations, we arrive at

$$S(A|B, \rho_1) = \max_{p\in\mathfrak{P}([N])} S(A|B, \rho_p) - \log N \tag{194}$$

and

$$I(A; E, \rho_1) = \max_{p\in\mathfrak{P}([N])} I(A; E, \rho_p) - 2\log N. \tag{195}$$

Note that

$$C_{m,\rightarrow}(\mathrm{conv}(\hat{\mathcal{X}})) = \max_{p\in\mathfrak{P}([N])} S(A|B, \rho_p) \tag{196}$$

by Proposition 13. Combining (194) with (177) and (196) together with (185) shows that

$$C_m^{AV}(\mathcal{X}) \leq C_m(\mathrm{conv}(\hat{\mathcal{X}})) - \log N + \delta \tag{197}$$

holds. The converse is valid by the merging cost converse for single states.[14] Moreover, by (195), our protocols have classical $A \rightarrow B$ classical communication rates with

$$\limsup_{l\rightarrow\infty} \frac{1}{l} \log D_l = \limsup_{l\rightarrow\infty} \frac{1}{l} \log(\tilde{D}_l \cdot N^l) \tag{198}$$

$$\leq I(A; E, \rho_1) + \delta + \log N \tag{199}$$

$$= \max_{p\in\mathfrak{P}([N])} I(A; E, \rho_p) - \log N + \delta, \tag{200}$$

where (198) follows from (180), (199) is by (178), and (200) is by (195). Since $\delta > 0$ was an arbitrary positive number, we are done. □

VII. CONCLUSION

In this work, we have shown simultaneous achievability of the optimal entanglement as well as classical communication cost of one-way quantum state merging in case, that the source to merge is a compound quantum source. In this way, we completed our work on quantum state merging for compound sources begun in Ref. 7.

We also determined the optimal entanglement rates for one-way entanglement distillation in case that the source from which the entanglement is distilled is an AVQS. In this case, Ahlswede's robustification and elimination technique turned out to be appropriate tools, and we can in fact, by the elimination technique, achieve each rate below the entanglement capacity with fidelity going to one exponentially fast and simultaneously approximate the classical communication rate of the utilized protocols for the compound source generated by the convex hull of the AVQS generating set.

Imposing a simple example of a class of AVQS, we demonstrated that applying the robustification and elimination technique to suitable protocols for the corresponding compound source (generated by the convex hull of the AVQS generating set) is insufficient in general. Another situation, where the above standard approach is not suitable, is the problem of proving achievability of the strong subspace capacity of an AVQC. In this case, the problem is not immediately accessible for the robustification technique, and this deficiency was overcome in Ref. 4 by first determining the capacity of the AVQC for entanglement transmission, and then showing equality of the capacities utilizing fairly nontrivial results from convex high-dimensional convex geometry.

The quantum state merging problem for AVQS, in contrast, seems accessible to robustification and elimination. However, application leads to suboptimal rates in some cases, as Example 28 shows.

In fact, a closer look to Example 28 reveals that the achievability result asserted by the inequality in (166) is not only suboptimal, but also meaningless in a qualitative sense for some AVQS.

Imagine a situation, in which the communication parties do not have any access to shared pure entanglement resources and they want to merge the AVQS generated by a set $\hat{\mathcal{X}}$ as in Example 28 with $C_m\left(\text{conv}(\hat{\mathcal{X}})\right) > 0$, where the number N of states in $\hat{\mathcal{X}}$ is assumed to be bounded

$$N > \exp(C_m(\text{conv}(\hat{\mathcal{X}})). \tag{201}$$

Having only protocols according to the achievability result (166) at hand, they infer that merging is impossible in their situation, while Example 28 shows that merging of the AVQS is, in fact, possible without external entanglement resources.

Summarizing our considerations, we notice with some regret that in case of quantum state merging, the merging cost of an AVQS generated by a set $\mathcal{X}$ seems, at least not immediately, related to the merging cost of the corresponding compound source generated by $\text{conv}(\mathcal{X})$. A merging cost function presumably will involve LOCC pre- and post-processing maximization. Probably, the merging cost for AVQS will require a multi-letter characterization.

ACKNOWLEDGMENTS

We express our deep gratitude to Igor Bjelaković, our former colleague and constant source of inspiration for years. Having prepared Ref. 7 with us, he also stimulated the initial discussions leading to the present paper. Partial results of this paper were presented in a talk at the 2014 IEEE International Symposium on Information Theory (ISIT 2014) without detailed proofs.[9]

The work of H.B. is supported by the DFG via Grant No. BO 1734/20-1 and by the BMBF via Grant No. 01BQ1050. G.J. gratefully acknowledges the support of the TUM Graduate School/Faculty Graduate Center FGC-EI at Technische Universität München, Germany.

APPENDIX: PROOF OF THE BOUND IN EQ. (55)

1. Proof of (55)

Let $\eta > 0$ be fixed and p, q probability distributions on $[d]$, such that

$$|H(p) - H(q)| \geq \eta \tag{A1}$$

holds. It is well known that the Shannon entropy is uniformly continuous in the variation distance (see, e.g., Ref. 11), it holds

$$|H(p) - H(q)| \leq f(\|p - q\|_1) \tag{A2}$$

with a strictly monotonically increasing function f. Therefore, (A1) and (A2) lead to

$$0 < 2c_3 := \frac{1}{2\ln 2} f^{-1}(\eta)^2 \leq \frac{1}{2\ln 2}\|p - q\|_1^2 \leq D(p||q), \tag{A3}$$

where the rightmost inequality is Pinsker's inequality $D(p||q) \geq \frac{1}{2\ln 2}\|p - q\|_1^2$. Since p and q were arbitrary probability distributions on $[d]$ with entropy distance bounded below by η, for each $i \in [N]$, the bound in (55) is valid for each $i \in [N]$.

[1] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," Z. Wahr. Verw. Geb. **44**, 159–175 (1978).
[2] R. Ahlswede, "Coloring hypergraphs: A new approach to multi-user source coding II," J. Comb., Inf. Syst. Sci. **5**(3), 220–268 (1980).
[3] R. Ahlswede, "Arbitrarily varying channels with state sequence known to the sender," IEEE Trans. Inf. Theory **32**, 621–629 (1986).
[4] R. Ahlswede, I. Bjelaković, H. Boche, and J. Nötzel, "Quantum capacity under adversarial quantum noise: Arbitrarily varying quantum channels," Commun. Math. Phys. **317**, 103–156 (2013).
[5] R. Ahlswede and V. Blinovsky, "Classical capacity of classical-quantum arbitrarily varying channels," IEEE Trans. Inf. Theory **53**, 526–533 (2007).
[6] R. Alicki and M. Fannes, "Continuity of quantum conditional information," J. Phys. A **37**, L55–L57 (2004).
[7] I. Bjelaković, H. Boche, and G. Janßen, "Universal quantum state merging," J. Math. Phys. **54**, 032204 (2013).
[8] H. Boche and J. Nötzel, "Positivity, discontinuity, finite resources and nonzero error for arbitrarily varying quantum channels," pre-print arXiv:1401.5360 (2014).
[9] H. Boche and G. Janßen, "Resource cost results for entanglement distillation and state merging under source uncertainties," in *2014 IEEE International Symposium on Information Theory Proceedings (ISIT)* (IEEE, 2014), pp. 721–725.
[10] M. Christandl and G. Mitchinson, "The spectra of quantum states and the Kronecker coefficients of the symmetric group," Commun. Math. Phys. **261**, 789–797 (2006).
[11] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. (Cambridge University Press, 2011).
[12] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," Proc. R. Soc. London, Ser. A **461**, 207–235 (2005).
[13] M. Fannes, "A continuity property of the entropy density for spin lattice systems," Commun. Math. Phys. **31**, 291–294 (1973).
[14] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum state merging and negative information," Commun. Math. Phys. **269**, 107–136 (2007).
[15] M. Keyl and R. F. Werner, "Estimating the spectrum of a density operator," Phys. Rev. A **64**, 052311 (2001).
[16] D. Leung and G. Smith, "Continuity of quantum channel capacities," Commun. Math. Phys. **292**, 201–215 (2009).
[17] T. Ogawa and H. Nagaoka, "Making good codes for classical-quantum channel coding via quantum hypothesis testing," IEEE Trans. Inf. Theory **53**, 2261–2266 (2007).
[18] R. Webster, *Convexity* (Oxford University Press, 1994).
[19] A. Winter, "Coding theorem and strong converse for quantum channels," IEEE Trans. Inf. Theory **45**, 2481–2485 (1999).