CrossMark

Quantum Conditional Mutual Information and Approximate Markov Chains

Omar Fawzi[1,2]**, Renato Renner**[1]

[1] Institute for Theoretical Physics, ETH Zurich, Zurich, Switzerland.
   E-mail: renner@ethz.ch
[2] LIP, UMR 5668 ENS Lyon-CNRS-UCBL-INRIA, Universite de Lyon, ENS de Lyon, Lyon, France.
   E-mail: omar.fawzi@ens-lyon.fr

Abstract: A state on a tripartite quantum system $A \otimes B \otimes C$ forms a Markov chain if it can be reconstructed from its marginal on $A \otimes B$ by a quantum operation from B to $B \otimes C$. We show that the quantum conditional mutual information $I(A : C|B)$ of an arbitrary state is an upper bound on its distance to the closest reconstructed state. It thus quantifies how well the Markov chain property is approximated.

1. Introduction

The *conditional mutual information* $I(A : C|B)_\rho = H(\rho_{AB}) + H(\rho_{BC}) - H(\rho_B) - H(\rho_{ABC})$ of a state ρ_{ABC} on a tripartite system $A \otimes B \otimes C$ is meant to quantify the correlations between A and C from the point of view of B. Here $H(\rho) = -\text{tr}(\rho \log_2 \rho)$ is the von Neumann entropy. Apart from its central role in traditional information theory, the conditional mutual information has recently found applications in new areas of computer science and physics. Examples include communication and information complexity (see [10] and references therein), de Finetti type theorems [8,9] and also the study of quantum many-body systems [36]. The importance of the conditional mutual information for such applications is due to its various useful properties. In particular, it has an additivity property called the *chain rule*: $I(A_1 \ldots A_n : C|B) = I(A_1 : C|B) + I(A_2 : C|BA_1) + \cdots + I(A_n : C|BA_1 \ldots A_{n-1})$.

When the B system is classical, the conditional mutual information $I(A : C|B)$ has a simple interpretation: it is the average over the values b taken by B of the (unconditional) mutual information evaluated for the conditional state on the system $A \otimes C$. This is crucial for applications because the (unconditional) mutual information can be related to operational quantities, such as the distance to product states using Pinsker's inequality for instance. However, when B is quantum, the conditional mutual information is significantly more complicated and much less is known about it. In fact, even the fact that $I(A : C|B) \geq 0$, also known as *strong subadditivity* of the von Neumann entropy, is a highly non-trivial theorem [39]. The structure of states that satisfy $I(A : C|B)_\rho = 0$

was also studied [28,43]. It has been found that a zero conditional mutual information characterises states ρ_{ABC} whose C system can be reconstructed just by acting on B, i.e., there exists a quantum operation $\mathcal{T}_{B \to BC}$ from the B to the $B \otimes C$ system such that

$$\rho_{ABC} = \mathcal{T}_{B \to BC}(\rho_{AB}). \tag{1}$$

States ρ_{ABC} that satisfy this condition are called *(quantum) Markov chains*. When B is classical the condition (1) simply means that, for all values b taken by B, the conditional state on $A \otimes C$ is a product state. We say that *A and C are independent given B*.

A natural question that is very relevant for applications is to characterise states for which the conditional mutual information is approximately zero, i.e., for which it is guaranteed that $I(A : C|B) \leq \epsilon$ for some $\epsilon > 0$. In applications involving n systems $A_1, \ldots, A_n$, such a guarantee is often obtained from an upper bound on the total conditional mutual information $I(A_1 \ldots A_n : C|B) \leq c$ (which can even be the trivial bound $2 \log_2 \dim C$). The chain rule mentioned above then implies that, on average over i, we have $I(A_i : C|BA_1 \ldots A_{i-1}) \leq c/n$. The authors of [30] gave evidence for the difficulty of characterising such states in the quantum setting by finding states for which the conditional mutual information is small whereas their distance to any Markov chain is large (see also [16] for more extreme examples). In unpublished works, Winter and Li[1] and Kim made the important observation that instead of considering the distance to a (perfect) Markov chain, another possibly more appropriate measure would be the accuracy with which Eq. 1 is satisfied (see also [56]). In fact, Kim[2] conjectured that the conditional mutual information is lower bounded by the trace distance between the two sides of Eq. 1 for a specific form for the map $\mathcal{T}_{B \to BC}$, known sometimes as the Petz map (cf. Eq. 15 below). Later, in the context of studying Rényi generalisations of the conditional mutual information, the authors of [4] refined this conjecture by replacing the trace distance with the negative logarithm of the fidelity (see also [48]). Here, we prove a variant of this last conjecture where the map $\mathcal{T}_{B \to BC}$ does not necessarily have the form of a Petz map.

Main result. We prove that for any state ρ_{ABC} on $A \otimes B \otimes C$ there exists a quantum operation $\mathcal{T}_{B \to BC}$ from the B system to the $B \otimes C$ system such that the fidelity of the reconstructed state

$$\sigma_{ABC} = \mathcal{T}_{B \to BC}(\rho_{AB}) \tag{2}$$

is at least[3]

$$F(\rho_{ABC}, \sigma_{ABC}) \geq 2^{-\frac{1}{2} I(A:C|B)_\rho}. \tag{3}$$

We refer to Theorem 5.1 for a more precise statement.

Reformulations and implications. A first immediate implication of our inequality is the *strong subadditivity* of the von Neumann entropy, $I(A : C|B)_\rho \geq 0$ [39]. The latter may be rewritten in terms of the conditional von Neumann entropy, $H(A|B)_\rho = H(\rho_{AB}) - H(\rho_B)$, as

$$H(A|B)_\rho \geq H(A|BC)_\rho \tag{4}$$

---

[1] This manuscript was recently updated and published online [38].

[2] See http://www.physics.usyd.edu.au/quantum/Coogee2013/Presentations/Kim.

[3] The *fidelity* of ρ and σ is defined as $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$.

and is also known as the *data processing inequality*. Furthermore, (3) implies that if (4) holds with equality for some state ρ_{ABC} then it satisfies the Markov chain condition (1), reproducing the result from [28,43]. The work presented here may thus be viewed as a robust extension of this result—if (4) holds with *approximate* equality then the Markov chain condition is fulfilled *approximately*.

Our result may also be rewritten as

$$\inf_{\sigma_{ABC}} D_{\frac{1}{2}}(\rho_{ABC}, \sigma_{ABC}) \leq I(A : C|B)_\rho, \tag{5}$$

where the infimum ranges over all *recovered* states, i.e., states of the form (2), and where $D_{1/2}(\rho\|\sigma) = -2\log_2 F(\rho, \sigma)$ is the Rényi divergence of order $\alpha = 1/2$ [40,55]. We remark that the quantity on the left hand side is equal to the *surprisal of the fidelity of recovery*, which has been introduced and studied in detail in [48].

Finally, we note that (3) also implies an upper bound on the trace distance, which we denote by $\Delta(\cdot, \cdot)$, between ρ_{ABC} and the recovered state σ_{ABC},

$$\frac{1}{\ln 2}\Delta(\rho_{ABC}, \sigma_{ABC})^2 \leq I(A : C|B)_\rho. \tag{6}$$

The bound is readily verified using $\Delta(\cdot, \cdot)^2 \leq 1 - F(\cdot, \cdot)^2$ (cf. Lemma B.1) and $1 - 2^{-x} \leq \ln(2)x$.

Tightness. One may ask whether, conversely to our main result, the conditional mutual information of a state ρ_{ABC} also gives a lower bound on its distance to any reconstructed state σ_{ABC} of the form (2). To answer this question, we note that, as a consequence of the data processing inequality, we have

$$I(A : C|B)_\rho = H(A|B)_\rho - H(A|BC)_\rho \leq H(A|BC)_\sigma - H(A|BC)_\rho. \tag{7}$$

The entropy difference on the right hand side can be bounded by the Alicki–Fannes inequality [1] in terms of the trace distance between the two states, yielding[4]

$$I(A : C|B)_\rho \leq 8\Delta \log_2(\dim A) - 4\Delta \log_2(2\Delta) - 2(1 - 2\Delta)\log_2(1 - 2\Delta)$$
$$\text{for } \Delta \leq \frac{1}{2}. \tag{8}$$

This can be seen as a converse to (6). To simplify the comparison, we may use

$$8\Delta - 4\Delta \log_2(2\Delta) - 2(1 - 2\Delta)\log_2(1 - 2\Delta) \leq 7\sqrt{\Delta} \quad \text{for } \Delta \leq \frac{1}{11}, \tag{9}$$

which gives

$$I(A : C|B)_\rho \leq 7\log_2(\dim A)\sqrt{\Delta(\rho_{ABC}, \sigma_{ABC})}. \tag{10}$$

Note that a term proportional to the logarithm of the dimension of A is necessary in general as the trace distance is always upper bounded by 1, whereas the conditional mutual information may be as large as $2\log_2 \dim A$.

---

[4] We refer to [4] for a more detailed discussion, including a proof that the same bound holds also when the conditional mutual information is evaluated for σ instead of ρ.

The classical case. Inequality (3) is easily obtained in the case where B is classical, i.e., when ρ_{ABC} is a qcq-state,

$$\rho_{ABC} = \sum_b P_B(b) |b\rangle\langle b|_B \otimes \rho_{AC,b}, \tag{11}$$

for some probability distribution P_B, an orthonormal basis $\{|b\rangle\}_b$ of B, and a family of states $\{\rho_{AC,b}\}_b$ on $A \otimes C$. Let $\mathcal{T}_{B \to BC}$ be any map such that

$$\mathcal{T}_{B \to BC}(|b\rangle\langle b|) = |b\rangle\langle b| \otimes \rho_{C,b} \qquad (\forall b), \tag{12}$$

where $\rho_{C,b} = \mathrm{tr}_A(\rho_{AC,b})$. Then the reconstructed state $\sigma_{ABC} = \mathcal{T}_{B \to BC}(\rho_{AB})$ is the qcq-state

$$\sigma_{ABC} = \sum_b P_B(b)\rho_{A,b} \otimes |b\rangle\langle b| \otimes \rho_{C,b}, \tag{13}$$

where $\rho_{A,b} = \mathrm{tr}_C(\rho_{AC,b})$. We remark that σ_{ABC} is a Markov chain. Furthermore, a straightforward calculation shows that the relative entropy[5] $D(\rho_{ABC}\|\sigma_{ABC})$ between ρ_{ABC} and σ_{ABC} is given by

$$D(\rho_{ABC}\|\sigma_{ABC}) = I(A : C|B)_\rho. \tag{14}$$

Inequality (3) then follows from Lemma B.2.

Related results. While the conditional mutual information is well understood in the classical case and has various interesting properties (see, e.g., [47]), these properties do not necessarily hold for quantum states. For example, identity (14) cannot be generalised directly to the case where B is non-classical (see [38] for a discussion). Furthermore, it has been discovered that there exist states ρ_{ABC} that have a large distance to the closest Markov chain, while the conditional mutual information is small [16,23,30]. We remark that this is not in contradiction to (3) as the reconstructed state σ_{ABC}, defined by (2), is not necessarily a Markov chain. (Note that this is a major difference to the classical case sketched above.)

As mentioned above, the special case of (3) where $I(A : C|B) = 0$ has been studied in earlier work [28,43]. There, it has also been shown that the relevant reconstruction map $\mathcal{T}_{B \to BC}$ is of the form

$$X_B \mapsto \rho_{BC}^{\frac{1}{2}}(\rho_B^{-\frac{1}{2}} X_B \rho_B^{-\frac{1}{2}} \otimes \mathrm{id}_C)\rho_{BC}^{\frac{1}{2}}. \tag{15}$$

However, it remained unclear whether this particular map also works in the case where $I(A : C|B)$ is strictly larger than zero, even though several conjectures in this direction were proposed and studied [4,56]. We refer to [38] for a detailed account of the evolution of these conjectures. We note that our result provides some information about the structure of the map for which (3) holds (cf. Theorem 5.1), but leaves open the question of whether it is of this particular form.

There is a large body of literature underlying the fundamental role that the conditional mutual information plays in quantum information theory. Notably, it has been shown to characterise the communication rate for the task of *quantum state redistribution* in the asymptotic limit of many independent copies of a resource state [20]. Furthermore, the

---

[5] See Sect. 2 for a definition.

quantum conditional mutual information is the basis for an important measure of entanglement, known as *squashed entanglement* [17]. The properties of this entanglement measure thus hinge on the properties of $I(A : C|B)$. In this context, lower bounds on $I(A : C|B)$ in terms of the distance between the marginal ρ_{AC} from the set of separable states have been proved in [7] and later improved in [37]. We also note that another lower bound on the conditional mutual information in terms of a distance between certain operators derived from ρ_{ABC} has recently been stated in [57]. This bound is based on a novel monotonicity bound for the relative entropy [11]. Our work may be used to obtain strengthened versions of some of these results. We are going to illustrate this for the case of squashed entanglement.

Applications. For us, one motivation to study how well the conditional mutual information characterises approximate Markov chains is in the context of device-independent quantum key distribution [22]. Another implication, proposed in [38], is a novel lower bound on the squashed entanglement of any bipartite state. The bound depends only on the trace distance to the closest k-extendible[6] state, and also implies a strong lower bound in terms of the trace distance to the closest separable state (cf. Appendix D for details).

It would be interesting to investigate whether inequality (3) can lead to better quantum de Finetti theorems. In fact, the authors of [8,9] recently gave beautiful proofs of various de Finetti theorems using the conditional mutual information. For the quantum version, they apply an informationally complete measurement to reduce the problem to the classical case, but this comes at the cost of a factor that is exponential in the number of systems. We also believe that inequality (3) will be helpful in proving communication complexity lower bounds via the *quantum information complexity* [31,32,35,51].

Structure of the proof. The proof of inequality (3) is based on two main ideas, which we discuss in separate sections. The first is the use of *one-shot entropy measures* [44] to bound the von Neumann relative entropy (Sect. 2). The second is an extension of the method of *de Finetti reductions* [14,15,45,46] (Sect. 3). We use the latter to derive a general tool for evaluating the fidelity of permutation-invariant states (Sect. 4). The proof of (3) then proceeds in two main steps in which these techniques are applied successively (Sect. 5).

2. Typicality Bounds on the Relative Entropy

In this section we are going to derive bounds on the relative entropy that will be used in the proof of Theorem 5.1. The method we use to obtain these bounds is inspired by a recent approach [2] to prove strong subadditivity of the von Neumann entropy (see Eq. 4). The idea there was to first prove strong subadditivity for one-shot entropies [44] and then use *typicality* or, more precisely, the *Asymptotic Equipartition Property* [50] to obtain the desired statement for the von Neumann entropy. Here we proceed analogously: we use one-shot versions of the relative entropy (defined in Appendix A) to obtain bounds on the von Neumann relative entropy.

The *(von Neumann) relative entropy* $D(\rho\|\sigma)$ for two non-negative operators ρ and σ is defined as

$$D(\rho\|\sigma) = \frac{1}{\operatorname{tr}(\rho)}\operatorname{tr}\big(\rho(\log_2 \rho - \log_2 \sigma)\big), \tag{16}$$

---

[6] A non-negative operator ω_{AC} is called *k-extendible* if there exists a non-negative operator $\bar{\omega}_{AC_1...C_n}$ such that $\bar{\omega}_{AC_i} = \omega_{AC}$ for all $i = 1, \ldots, k$.

where we set $D(\rho \| \sigma) = \infty$ if the support of ρ is not contained in the support of σ. Our statements also refer to the *trace distance*. While this distance is often defined for density operators only, we define it here more generally for any non-negative operators ρ and σ by

$$\Delta(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1 + \frac{1}{2} |{\rm tr}(\rho - \sigma)| \tag{17}$$

(see Sect. 3.2 of [49]). Note that the second term is zero if ρ and σ are both density operators. We also remark that the trace distance may be rewritten as

$$\Delta(\rho, \sigma) = \max\left[{\rm tr}(Y^+), {\rm tr}(Y^-)\right], \tag{18}$$

where $Y+$ and Y^- are the positive and negative parts of $\rho - \sigma$, i.e., $\rho - \sigma = Y^+ - Y^-$ with $Y^+ \geq 0$, $Y^- \geq 0$, and ${\rm tr}(Y^+ Y^-) = 0$. It follows that we can write Δ as

$$\Delta(\rho, \sigma) = \sup_{0 \leq Q \leq {\rm id}} |{\rm tr}(Q(\rho - \sigma))|. \tag{19}$$

One can easily see from this expression that for any trace non-increasing completely positive map $\mathcal{W}$ we have

$$\Delta(\mathcal{W}(\rho), \mathcal{W}(\sigma)) \leq \Delta(\rho, \sigma). \tag{20}$$

Our first lemma provides an upper bound on the relative entropy in terms of sequences of operators that satisfy an operator inequality.

Lemma 2.1. *Let ρ be a density operator, let σ be a non-negative operator, and let $\{\bar{\rho}_n\}_{n\in\mathbb{N}}$ be a sequence of non-negative operators such that for some $s \in \mathbb{R}$*

$$\bar{\rho}_n \leq 2^{sn} \sigma^{\otimes n} \quad (\forall n \in \mathbb{N}) \quad and \quad \lim_{n\to\infty} \Delta(\bar{\rho}_n, \rho^{\otimes n}) < 1. \tag{21}$$

Then $D(\rho \| \sigma) \leq s$.

Proof. By assumption, there exist $c < 1$ and $n_0 \in \mathbb{N}$ such that

$$\Delta(\bar{\rho}_n, \rho^{\otimes n}) \leq c \tag{22}$$

holds for all $n \geq n_0$. Let $\epsilon \in (c, 1)$. By Lemma A.5 we have

$$D_H^\epsilon(\rho^{\otimes n} \| \sigma^{\otimes n}) \leq sn - \log_2(1 - c/\epsilon), \tag{23}$$

where $D_H^\epsilon(\cdot \| \cdot)$ is the generalised relative entropy defined in Appendix A. Setting $C = 1 - c/\epsilon > 0$ we conclude that

$$\lim_{n\to\infty} \frac{1}{n} D_H^\epsilon(\rho^{\otimes n} \| \sigma^{\otimes n}) \leq \lim_{n\to\infty} \left(s + \frac{1}{n} \log_2 \frac{1}{C}\right) = s. \tag{24}$$

The claim then follows from the Asymptotic Equipartition Property of $D_H^\epsilon(\cdot \| \cdot)$ (Lemma A.8).   □

The following lemma is in some sense a converse of Lemma 2.1.

Lemma 2.2. *Let ρ be a density operator, let σ be non-negative operator, and let $s >$ $D(\rho\|\sigma)$. Then there exists $\kappa > 0$ and a sequence of non-negative operators $\{\bar{\rho}_n\}_{n\in\mathbb{N}}$ with $\mathrm{tr}(\bar{\rho}_n) \leq 1$ such that*

$$\bar{\rho}_n \leq 2^{sn}\sigma^{\otimes n} \quad (\forall n \in \mathbb{N}) \quad and \quad \lim_{n\to\infty} 2^{n\kappa}\Delta(\bar{\rho}_n, \rho^{\otimes n}) = 0. \tag{25}$$

Proof. The proof uses the smooth relative max-entropy $D_{\max}^{\epsilon}(\cdot\|\cdot)$ defined in Appendix A. The Asymptotic Equipartition Property for this entropy measure (Lemma A.7) asserts that there exists $n_0 \in \mathbb{N}$ such that for any $n \geq n_0$

$$D_{\max}^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) < ns \tag{26}$$

for $\epsilon_n > 0$ chosen such that

$$D(\rho\|\sigma) + c\sqrt{\frac{\log_2(2/\varepsilon_n^2)}{n}} = s, \tag{27}$$

where c is independent of n. Inserting this into the definition of $D_{\max}^{\epsilon}(\cdot\|\cdot)$ we find that there exists a non-negative operator $\bar{\rho}_n$ with $\mathrm{tr}(\bar{\rho}_n) \leq 1$ such that

$$\bar{\rho}_n \leq 2^{sn}\sigma^{\otimes n} \tag{28}$$

and

$$\sqrt{1 - F(\bar{\rho}_n, \rho^{\otimes n})^2} \leq \epsilon_n. \tag{29}$$

Equation (27) may be rewritten as

$$\epsilon_n = \sqrt{2}\,2^{-\kappa'n/2} \quad \text{with} \quad \kappa' = \left(\frac{s - D(\rho\|\sigma)}{c}\right)^2. \tag{30}$$

Inserting this in (29) and using Lemma B.1, we conclude that

$$\Delta(\bar{\rho}_n, \rho^{\otimes n}) \leq \sqrt{2}\,2^{-\kappa'n/2}. \tag{31}$$

This proves (25) for any $\kappa < \kappa'/2$. (Note that for $n < n_0$ we may simply set $\bar{\rho}_n = 0$ so that the left hand side of (25) holds for all $n \in \mathbb{N}$.) $\quad\square$

The next lemma asserts that the relative entropy, evaluated for n-fold product states, has the following stability property: if one acts with the same trace non-increasing map on the two arguments then the relative entropy cannot substantially increase. This property is used in the proof of Theorem 5.1 (but see also Remark 2.4).

Lemma 2.3. *Let ρ be a density operator, let σ be a non-negative operator on the same space, and let $\{\mathcal{W}_n\}_{n\in\mathbb{N}}$ be a sequence of trace non-increasing completely positive maps on the n-fold tensor product of this space. If $\mathrm{tr}(\mathcal{W}_n(\rho^{\otimes n}))$ decreases less than exponentially in n, i.e.,*

$$\liminf_{n\to\infty} e^{\xi n}\mathrm{tr}\big(\mathcal{W}_n(\rho^{\otimes n})\big) > 0 \tag{32}$$

for any $\xi > 0$, then

$$\limsup_{n\to\infty} \frac{1}{n}D\big(\mathcal{W}_n(\rho^{\otimes n})\|\mathcal{W}_n(\sigma^{\otimes n})\big) \leq D(\rho\|\sigma). \tag{33}$$

Proof. Let $\delta > 0$. Lemma 2.2 tells us that there exists $\kappa > 0$ and a sequence of non-negative operators $\{\bar{\rho}_m\}_{m \in \mathbb{N}}$ such that

$$\bar{\rho}_m \leq 2^{m(D(\rho\|\sigma)+\delta)}\sigma^{\otimes m} \tag{34}$$

and

$$\lim_{m \to \infty} e^{\kappa m}\Delta(\bar{\rho}_m, \rho^{\otimes m}) = 0. \tag{35}$$

To abbreviate notation, we define $r_n = 1/\mathrm{tr}(\mathcal{W}_n(\rho^{\otimes n}))$. Note that, by assumption, r_n grows less than exponentially in n, so that

$$r_n < e^{\kappa n} \tag{36}$$

holds for n sufficiently large.

Let now $k, n \in \mathbb{N}$ and set $m = kn$. Applying $\mathcal{W}_n$ and multiplying with the factor r_n on the two sides of (34) yields

$$r_n^k \mathcal{W}_n^{\otimes k}(\bar{\rho}_{nk}) \leq 2^{kn(D(\rho\|\sigma)+\delta)}\left(r_n \mathcal{W}_n(\sigma^{\otimes n})\right)^{\otimes k}. \tag{37}$$

As $\mathcal{W}_n$ is trace non-increasing,

$$\lim_{k \to \infty} \Delta\left(r_n^k \mathcal{W}_n^{\otimes k}(\bar{\rho}_{nk}), \left(r_n \mathcal{W}_n(\rho^{\otimes n})\right)^{\otimes k}\right) = \lim_{k \to \infty} r_n^k \Delta\left(\mathcal{W}_n^{\otimes k}(\bar{\rho}_{nk}), \mathcal{W}_n^{\otimes k}(\rho^{\otimes nk})\right)$$

$$\leq \lim_{k \to \infty} r_n^k \Delta(\bar{\rho}_{nk}, \rho^{\otimes nk})$$

$$\leq \lim_{k \to \infty} e^{\kappa nk}\Delta(\bar{\rho}_{nk}, \rho^{\otimes nk}) = 0, \tag{38}$$

where the first inequality follows from the monotonicity property of the trace distance (20), the second inequality follows from (36), and where the final equality follows from (35). We can now apply Lemma 2.1 to the density operator $r_n \mathcal{W}_n(\rho^{\otimes n})$ and the non-negative operator $r_n \mathcal{W}_n(\sigma^{\otimes n})$, which gives

$$D\left(r_n \mathcal{W}_n(\rho^{\otimes n})\|r_n \mathcal{W}_n(\sigma^{\otimes n})\right) \leq n\left(D(\rho\|\sigma) + \delta\right). \tag{39}$$

Noting that multiplying both arguments of the relative entropy with the same factor leaves it unchanged we conclude

$$\frac{1}{n}D\left(\mathcal{W}_n(\rho^{\otimes n})\|\mathcal{W}_n(\sigma^{\otimes n})\right) \leq D(\rho\|\sigma) + \delta. \tag{40}$$

Taking the limit $n \to \infty$ and noting that $\delta > 0$ was arbitrary, the claim of the lemma follows. □

Remark 2.4. Lemma 2.3 will be used in one of the steps of the proof of Theorem 5.1. We note that, alternatively, this step may also be based on the inequality (cf. Lemma 25 of [3])

$$(1 - \epsilon)D\left(\mathcal{W}(\rho)\|\mathcal{W}(\sigma)\right) \leq D(\rho\|\sigma) + \epsilon \log_2(\mathrm{tr}(\sigma)/\epsilon), \tag{41}$$

which holds for any density operator ρ, any non-negative operator σ, any trace non-increasing completely positive map $\mathcal{W}$, and $\epsilon = 1 - \mathrm{tr}\left(\mathcal{W}(\rho)\right)$ (see also Footnote 10). However, Lemma 2.3 provides a stronger stability condition for the relative entropy of product states (notably when $\epsilon \gg 0$) and may therefore be useful for generalisations of our results.

As a corollary of Lemma 2.3 we also obtain the well known *Asymptotic Equipartition Property* (see, e.g., Chapter 3 of [18]). We state it here explicitly as Lemma 2.5 because we are going to use it within the proof of Theorem 5.1 and because it illustrates the use of Lemma 2.3.

Lemma 2.5. *Let ρ be a density operator. For any $n \in \mathbb{N}$ let $\rho^{\otimes n} = \sum_{s \in S_n} s\, \Pi_s$ where S_n is the set of eigenvalues of $\rho^{\otimes n}$ and where Π_s, for $s \in S_n$, is the projector onto the corresponding eigenspace. Furthermore, for any $\delta > 0$, let S_n^{δ} be the subset of S_n defined by*

$$S_n^{\delta} = \left\{ s \in S_n : s \in [2^{-n(H(\rho)+\delta)}, 2^{-n(H(\rho)-\delta)}] \right\}. \tag{42}$$

Then $\lim_{n \to \infty} \sum_{s \in S_n^{\delta}} \operatorname{tr}(\Pi_s \rho^{\otimes n}) = 1$ and the convergence is exponentially fast in n, i.e., there exists $\xi > 0$ such that $\sum_{s \in S_n^{\delta}} \operatorname{tr}(\Pi_s \rho^{\otimes n}) \geq 1 - e^{-\xi n}$.

Proof. Let $n \in \mathbb{N}$ and consider the projector $\bar{\Pi}_n^+ = \sum_{s \in S_n^+} \Pi_s$, where $S_n^+ = \{s \in S_n : s > 2^{-n(H(\rho)-\delta)}\}$. An explicit evaluation of the relative entropy shows that

$$
\begin{aligned}
D(\bar{\Pi}_n^+ \rho^{\otimes n} \bar{\Pi}_n^+ \| \bar{\Pi}_n^+) &= \frac{\operatorname{tr}\left(\rho^{\otimes n} \bar{\Pi}_n^+ \log_2(\rho^{\otimes n} \bar{\Pi}_n^+)\right)}{\operatorname{tr}(\rho^{\otimes n} \bar{\Pi}_n^+)} \\
&\geq \frac{\operatorname{tr}\left(\rho^{\otimes n} \bar{\Pi}_n^+ \log_2(2^{-n(H(\rho)-\delta)} \bar{\Pi}_n^+)\right)}{\operatorname{tr}(\rho^{\otimes n} \bar{\Pi}_n^+)} = -n\left(H(\rho) - \delta\right).
\end{aligned}
\tag{43}
$$

Using

$$-H(\rho) = D(\rho \| \mathrm{id}) \tag{44}$$

and defining the map $\mathcal{W}_n^+ : X \mapsto \bar{\Pi}_n^+ X \bar{\Pi}_n^+$ we can rewrite this bound as

$$\frac{1}{n} D(\mathcal{W}_n^+(\rho^{\otimes n}) \| \mathcal{W}_n^+(\mathrm{id}^{\otimes n})) \geq D(\rho \| \mathrm{id}) + \delta. \tag{45}$$

If we now assume, by contradiction, that $\operatorname{tr}(\rho^{\otimes n} \bar{\Pi}_n^+) = \operatorname{tr}(\mathcal{W}_n^+(\rho^{\otimes n}))$ decreases less than exponentially fast in n, Lemma 2.3 tells us that

$$\limsup_{n \to \infty} \frac{1}{n} D(\mathcal{W}_n^+(\rho^{\otimes n}) \| \mathcal{W}_n^+(\mathrm{id}^{\otimes n})) \leq D(\rho \| \mathrm{id}). \tag{46}$$

This is obviously in contradiction to (45) and thus proves that $\operatorname{tr}(\rho^{\otimes n} \bar{\Pi}_n^+)$ decreases exponentially fast in n.

Similarly, we may consider the projector $\bar{\Pi}_n^- = \sum_{s \in S_n^-} \Pi_s$ where $S_n^- = \{s \in S_n : s < 2^{-n(H(\rho)+\delta)}\}$. Here, instead of (44), we use that for any purification ρ_{DR} of $\rho_D = \rho$

$$H(\rho) = -H(D|R)_\rho = D(\rho_{DR} \| \mathrm{id}_D \otimes \rho_R). \tag{47}$$

We may choose the purification such that $(\bar{\Pi}_n^- \otimes \mathrm{id}_{R^n})\rho_{DR}^{\otimes n}(\bar{\Pi}_n^- \otimes \mathrm{id}_{R^n}) = (\bar{\Pi}_n^- \otimes \bar{\Pi}_n^-)\rho_{DR}^{\otimes n}(\bar{\Pi}_n^- \otimes \bar{\Pi}_n^-)$. Then

$$
\begin{aligned}
D\big((\bar{\Pi}_n^- &\otimes \mathrm{id}_{R^n})\rho_{DR}^{\otimes n}(\bar{\Pi}_n^- \otimes \mathrm{id}_{R^n}) \big\| \bar{\Pi}_n^- \otimes \rho_R^{\otimes n}\big) \\
&= \log_2 \mathrm{tr}(\rho_D^{\otimes n}\bar{\Pi}_n^-) - \frac{\mathrm{tr}\big(\rho_R^{\otimes n}\bar{\Pi}_n^- \log_2(\rho_R^{\otimes n}\bar{\Pi}_n^-)\big)}{\mathrm{tr}(\rho_D^{\otimes n}\bar{\Pi}_n^-)} \\
&\geq \log_2 \mathrm{tr}(\rho_D^{\otimes n}\bar{\Pi}_n^-) - \frac{\mathrm{tr}\big(\rho_R^{\otimes n}\bar{\Pi}_n^- \log_2(2^{-n(H(\rho)+\delta)}\bar{\Pi}_n^-)\big)}{\mathrm{tr}(\rho_D^{\otimes n}\bar{\Pi}_n^-)} \\
&= \log_2 \mathrm{tr}(\rho_D^{\otimes n}\bar{\Pi}_n^-) + n\big(H(\rho) + \delta\big). \quad\quad\quad\quad\quad (48)
\end{aligned}
$$

Defining $\mathcal{W}_n^- : X_{DR} \mapsto (\bar{\Pi}_n^- \otimes \mathrm{id}_R)X_{DR}(\bar{\Pi}_n^- \otimes \mathrm{id}_R)$ and inserting (47) we obtain the bound

$$
\frac{1}{n}D\big(\mathcal{W}_n^-(\rho_{DR}^{\otimes n})\big\|\mathcal{W}_n^-(\mathrm{id}_D^{\otimes n}\otimes\rho_R^{\otimes n})\big) \geq D(\rho_{DR}\|\mathrm{id}_D\otimes\rho_R) + \delta + \frac{\log_2 \mathrm{tr}\big(\mathcal{W}_n^-(\rho_{DR}^{\otimes n})\big)}{n}.
$$
$$(49)$$

Assume now, by contradiction, that $\mathrm{tr}(\rho_D^{\otimes n}\bar{\Pi}_n^-) = \mathrm{tr}(\mathcal{W}_n^-(\rho_{DR}^{\otimes n}))$ decreases less then exponentially fast in n. Then the last term of (49) approaches 0 in the limit of large n. In particular, we have

$$
\limsup_{n\to\infty} \frac{1}{n}D\big(\mathcal{W}_n^-(\rho_{DR}^{\otimes n})\big\|\mathcal{W}_n^-(\mathrm{id}_D^{\otimes n}\otimes\rho_R^{\otimes n})\big) \geq D(\rho_{DR}\|\mathrm{id}_D\otimes\rho_R) + \delta, \quad (50)
$$

which contradicts the statement of Lemma 2.3. We have thus shown that both $\mathrm{tr}(\rho^{\otimes n}\bar{\Pi}_n^-)$ and $\mathrm{tr}(\rho^{\otimes n}\bar{\Pi}_n^+)$ decrease exponentially fast in n. The claim of the lemma follows because $\sum_{s\in S_n^\delta} \Pi_s = \mathrm{id} - \bar{\Pi}_n^+ - \bar{\Pi}_n^-$. $\quad\square$

We conclude this section with a remark that is going to be useful for our proof of Theorem 5.1.

Remark 2.6. Considering the decomposition $\rho = \sum_{r\in R} r\pi_r$, it is easy to see that all eigenvalues of $\rho^{\otimes n}$ have the form $\prod_{r\in R} r^{n_r}$, where $(n_r)_{r\in R}$ are partitions of n, i.e., elements from the set

$$
Q_n = \Big\{(n_r)_{r\in R}, \ n_r \in \mathbb{N}_0, \ \sum_{r\in R} n_r = n\Big\}. \quad\quad (51)
$$

Hence, the set S_n of eigenvalues of $\rho^{\otimes n}$ used within Lemma 2.5 has size at most $|S_n| \leq |Q_n|$. Since $|Q_n| = \binom{n+|R|-1}{n} \leq (n+1)^{|R|}$, where $|R| \leq \mathrm{rank}(\rho)$ is the number of different eigenvalues of ρ, we can upper bound the size of S_n by

$$
|S_n| \leq (n+1)^{\mathrm{rank}(\rho)}. \quad\quad (52)
$$

3. Generalised de Finetti Reduction

The main result of this section, stated as Lemma 3.1, is motivated by a variant of the method of *de Finetti reductions* proposed in [15]. (This variant is also known as *post-selection technique*; we refer to [46] for a not too technical presentation.) De Finetti reductions are generally used to study states on n-fold product systems $S^{\otimes n}$ that are invariant under permutations of the subsystems [14,45]. More precisely, the idea is to reduce the analysis of any density operator ρ_{S^n} in the symmetric subspace $\mathrm{Sym}^n(S)$ of $S^{\otimes n}$ to the—generally simpler—analysis of states of the form $\sigma_S^{\otimes n}$, where σ_S is pure. We extend this method to the case where $S = D \otimes E$ is a bipartite space and where the marginal of $\rho_{S^n} = \rho_{D^n E^n}$ on $D^{\otimes n}$ is known to have the form

$$\mathrm{tr}_{E^n}(\rho_{D^n E^n}) = \rho_{D^n} = \sigma_D^{\otimes n} \tag{53}$$

for some given state σ_D on D. Lemma 3.1 implies that, in this case, the analysis can be reduced to states of the form $\sigma_{DE}^{\otimes n}$, where σ_{DE} is a purification of σ_D. (We note that a similar extension has been proposed earlier for another variant of the de Finetti reduction method; see Remark 4.3.3 of [44].) Lemma 3.1 will play a central role for the derivation of the claims of Sect. 4 below. Its proof uses concepts from representation theory, which are presented in Appendix C.

Lemma 3.1. *Let D and E be Hilbert spaces and let σ_D be a non-negative operator on D. Then there exists a probability measure $\mathrm{d}\phi$ on the set of purifications $|\phi\rangle\langle\phi|_{DE}$ of σ_D such that*

$$\rho_{D^n E^n} \leq (n+1)^{d^2-1} \int |\phi\rangle\langle\phi|_{DE}^{\otimes n} \mathrm{d}\phi \tag{54}$$

holds for any $n \in \mathbb{N}$, any permutation-invariant purification $\rho_{D^n E^n}$ of $\sigma_D^{\otimes n}$, and $d = \max[\dim(D), \dim(E)]$.

Proof. For the following argument, we assume without loss of generality that $d = \dim(D) = \dim(E)$, and that σ_D has full rank and is therefore invertible on D. (If this is not the case one may embed the smaller space in one of dimension d and replace σ_D by $\sigma_D + \epsilon \, \mathrm{id}_D$ for $\epsilon > 0$. The claim is then obtained in the limit $\epsilon \to 0$.) We define

$$|\theta\rangle_{DE} = \sum_i |d_i\rangle_D \otimes |e_i\rangle_E, \tag{55}$$

where $\{|d_i\rangle_D\}_i$ and $\{|e_i\rangle_E\}_i$ are orthonormal bases of D and E, respectively. Let now

$$T_{D^n E^n} = \int (\mathrm{id}_{D^n} \otimes U_E^{\otimes n}) |\theta\rangle\langle\theta|_{DE}^{\otimes n} (\mathrm{id}_{D^n} \otimes U_E^{\otimes n})^\dagger \mathrm{d}U, \tag{56}$$

where $\mathrm{d}U$ is the Haar probability measure on the group of unitaries on E. Because $(\sigma_D^{\frac{1}{2}} \otimes U_E)|\theta\rangle\langle\theta|(\sigma_D^{\frac{1}{2}} \otimes U_E^\dagger)$ is a purification of σ_D for any U_E, the operator

$$\begin{aligned} \tau_{D^n E^n} &= (\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{\frac{1}{2}} T_{D^n E^n} (\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{\frac{1}{2}} \\ &= \int \left((\sigma_D^{\frac{1}{2}} \otimes U_E)|\theta\rangle\langle\theta|_{DE}(\sigma_D^{\frac{1}{2}} \otimes U_E^\dagger)\right)^{\otimes n} \mathrm{d}U \end{aligned} \tag{57}$$

is obviously of the form

$$\tau_{D^n E^n} = \int |\phi\rangle\langle\phi|_{DE}^{\otimes n} \mathrm{d}\phi, \tag{58}$$

for some suitably chosen measure $\mathrm{d}\phi$ on the set of purifications $|\phi\rangle\langle\phi|_{DE}$ of σ_D. It therefore suffices to show that

$$\rho_{D^n E^n} \le (n+1)^{d^2-1}\tau_{D^n E^n}. \tag{59}$$

We do this by analysing the structure of $T_{D^n E^n}$. For this we employ the Schur–Weyl duality, which equips the product space $(D \otimes E)^{\otimes n}$ with a convenient structure (see Appendix C). Specifically, according to Lemma C.1, the vector $|\theta\rangle_{DE}^{\otimes n}$ decomposes as

$$|\theta\rangle_{DE}^{\otimes n} = \sum_\lambda |\phi_\lambda\rangle_{U_{D,\lambda}U_{E,\lambda}} \otimes |\Psi_\lambda\rangle_{V_{D,\lambda}V_{E,\lambda}}. \tag{60}$$

where, for each Young diagram λ,

$$|\Psi_\lambda\rangle_{V_{D,\lambda}V_{E,\lambda}} = \sqrt{\dim(V_\lambda)}|\psi_\lambda\rangle_{V_{D,\lambda}V_{E,\lambda}} = \sum_k |v_k\rangle_{V_{D,\lambda}} \otimes |\bar{v}_k\rangle_{V_{E,\lambda}} \tag{61}$$

for orthonormal bases $\{|v_k\rangle_{V_{D,\lambda}}\}_k$ and $\{|\bar{v}_k\rangle_{V_{E,\lambda}}\}_k$ of $V_{D,\lambda}$ and $V_{E,\lambda}$, respectively, and $|\phi_\lambda\rangle_{U_{D,\lambda}U_{E,\lambda}}$ is a vector in $U_{D,\lambda} \otimes U_{E,\lambda}$. The latter may always be written in the Schmidt decomposition as

$$|\phi_\lambda\rangle_{U_{D,\lambda}U_{E,\lambda}} = \sum_j \alpha_{\lambda,j}|u_j\rangle_{U_{D,\lambda}} \otimes |\bar{u}_j\rangle_{U_{E,\lambda}}, \tag{62}$$

where $\{|u_j\rangle_{U_{D,\lambda}}\}_j$ and $\{|\bar{u}_j\rangle_{U_{E,\lambda}}\}_j$ are orthonormal bases of of $U_{D,\lambda}$ and $U_{E,\lambda}$, respectively, and $\alpha_{\lambda,j}$ are appropriately chosen coefficients, which we assume without loss of generality to be real. The marginal of $|\theta\rangle\langle\theta|_{DE}^{\otimes n}$ on $D^{\otimes n} \cong \bigoplus_{D,\lambda} U_{D,\lambda} \otimes V_{D,\lambda}$ is equal to the identity and can thus be written as

$$\mathrm{tr}_{E^n}(|\theta\rangle\langle\theta|_{DE}^{\otimes n}) = \mathrm{id}_{D^n} = \sum_\lambda \mathrm{id}_{U_{D,\lambda}} \otimes \mathrm{id}_{V_{D,\lambda}}. \tag{63}$$

Comparing this to (60) shows that all coefficients $\alpha_{\lambda,j}$ in (62) must be equal to 1, i.e.,

$$|\phi_\lambda\rangle_{U_{D,\lambda}U_{E,\lambda}} = \sum_j |u_j\rangle_{U_{D,\lambda}} \otimes |\bar{u}_j\rangle_{U_{E,\lambda}}. \tag{64}$$

Note also that, according to the Schur–Weyl duality (see, e.g., Theorem 1.10 of [13]), $U_E^{\otimes n}$ acts on $E^{\otimes n} \cong \bigoplus_{E,\lambda} U_{E,\lambda} \otimes V_{E,\lambda}$ as $\sum_\lambda U_{E,\lambda}(U) \otimes \mathrm{id}_{V_{E,\lambda}}$. We thus have

$$(\mathrm{id}_{D^n} \otimes U_E^{\otimes n})|\theta\rangle^{\otimes n} = \sum_{\lambda,j} |u_j\rangle_{U_{D,\lambda}} \otimes U_{E,\lambda}(U)|\bar{u}_j\rangle_{U_{E,\lambda}} \otimes |\Psi_\lambda\rangle_{V_{D,\lambda}V_{E,\lambda}}. \tag{65}$$

We may therefore write

$$\begin{aligned}
T_{D^n E^n} = \sum_{\lambda,\lambda',j,j'} |u_j\rangle\langle u_{j'}|_{U_{D,\lambda}\leftarrow U_{D,\lambda'}} \otimes (T_{\lambda,\lambda',j,j'})_{U_{E,\lambda}\leftarrow U_{E,\lambda'}} \\
\otimes |\Psi_\lambda\rangle\langle\Psi_{\lambda'}|_{(V_{D,\lambda}V_{E,\lambda})\leftarrow(V_{D,\lambda'}V_{E,\lambda'})},
\end{aligned} \tag{66}$$

where $T_{\lambda,\lambda',j,j'}$ is the homomorphism between $U_{E,\lambda'}$ and $U_{E,\lambda}$ defined by

$$(T_{\lambda,\lambda',j,j'})_{U_{E,\lambda}\leftarrow U_{E,\lambda'}} = \int U_{E,\lambda}(U)|\bar{u}_j\rangle\langle\bar{u}_{j'}|_{U_{E,\lambda}\leftarrow U_{E,\lambda'}}U_{E,\lambda'}(U)^\dagger dU. \quad (67)$$

Since this operator manifestly commutes with the action of the unitary, Schur's lemma (see, e.g., Lemma 0.8 of [13]), together with the fact that $U_{E,\lambda}$ and $U_{E,\lambda'}$ are inequivalent for $\lambda \neq \lambda'$, implies that it has the form

$$(T_{\lambda,\lambda',j,j'})_{U_{E,\lambda}\leftarrow U_{E,\lambda'}} = \mu_{\lambda,j,j'}\delta_{\lambda,\lambda'}\mathrm{id}_{U_{E,\lambda}} \quad (68)$$

for appropriately chosen coefficients $\mu_{\lambda,j,j'}$. Inserting this in (66) gives

$$T_{D^n E^n} = \sum_{\lambda,j,j'} \mu_{\lambda,j,j'}|u_j\rangle\langle u_{j'}|_{U_{D,\lambda}} \otimes \mathrm{id}_{U_{E,\lambda}} \otimes |\Psi_\lambda\rangle\langle\Psi_\lambda|_{V_{D,\lambda}V_{E,\lambda}}. \quad (69)$$

Because the marginal of $T_{D^n E^n}$ on $D^{\otimes n}$,

$$T_{D^n} = \sum_{\lambda,j,j'} \mu_{\lambda,j,j'}\dim(U_\lambda)|u_j\rangle\langle u_{j'}|_{U_{D,\lambda}} \otimes \mathrm{id}_{V_{D,\lambda}}, \quad (70)$$

must be equal to the marginal of $|\theta\rangle\langle\theta|_{DE}^{\otimes n}$, we conclude from (63) that $\mu_{\lambda,j,j'} = \frac{1}{\dim(U_\lambda)}\delta_{j,j'}$. Hence,

$$T_{D^n E^n} = \sum_\lambda \frac{\dim(V_\lambda)}{\dim(U_\lambda)} \mathrm{id}_{U_{D,\lambda}} \otimes \mathrm{id}_{U_{E,\lambda}} \otimes |\psi_\lambda\rangle\langle\psi_\lambda|_{V_{D,\lambda}V_{E,\lambda}}, \quad (71)$$

where $|\psi_\lambda\rangle_{V_{E,\lambda}}$ is normalised.

Defining the invertible operator

$$\kappa_{D^n} = \sum_\lambda \frac{\dim(V_\lambda)}{\dim(U_\lambda)} \mathrm{id}_{U_{D,\lambda}} \otimes \mathrm{id}_{V_{D,\lambda}} \quad (72)$$

we have

$$S_{D^n E^n} = (\kappa_{D^n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}} T_{D^n E^n}(\kappa_{D^n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}$$
$$= \sum_\lambda \mathrm{id}_{U_{D,\lambda}} \otimes \mathrm{id}_{U_{E,\lambda}} \otimes |\psi_\lambda\rangle\langle\psi_\lambda|_{V_{D,\lambda}V_{E,\lambda}}. \quad (73)$$

Note that κ_{D^n} commutes with any permutation, because, according to the Schur–Weyl duality, permutations act like $\sum_\lambda \mathrm{id}_{U_\lambda} \otimes V_\lambda(\pi)$ on the decomposition of $D^{\otimes n}$. Consequently, because the support of $T_{D^n E^n}$ is contained in the symmetric subspace $\mathrm{Sym}^n(D \otimes E)$, the same must hold for $S_{D^n E^n}$. Furthermore, for any vector $|\Omega\rangle \in \mathrm{Sym}^n(D \otimes E)$, it follows from its representation according to Lemma C.1 that $S_{D^n E^n}|\Omega\rangle = |\Omega\rangle$. This proves that

$$S_{D^n E^n} = \mathrm{id}_{\mathrm{Sym}^n(D\otimes E)}. \quad (74)$$

Consider now the operator

$$Q_{D^n E^n} = (\kappa_{D^n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}\rho_{D^n E^n}(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}(\kappa_{D^n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}. \quad (75)$$

Since the support of $\rho_{D^n E^n}$ is contained in $\mathrm{Sym}^n(D \otimes E)$, the same must hold for $Q_{D^n E^n}$ and we find

$$Q_{D^n E^n} \leq \|Q_{D^n E^n}\|_\infty \mathrm{id}_{\mathrm{Sym}^n(D \otimes E)} \leq \mathrm{tr}(Q_{D^n E^n})\mathrm{id}_{\mathrm{Sym}^n(D \otimes E)} = \mathrm{tr}(Q_{D^n E^n})S_{D^n E^n}. \tag{76}$$

This, in turn, implies that

$$\rho_{D^n E^n} \leq \mathrm{tr}(Q_{D^n E^n})(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{\frac{1}{2}} T_{D^n E^n}(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{\frac{1}{2}} = \mathrm{tr}(Q_{D^n E^n})\tau_{D^n E^n}. \tag{77}$$

To conclude the proof of (59), we note that $\rho_{D^n} = \sigma_D^{\otimes n} = \tau_{D^n}$, which implies

$$
\begin{aligned}
\mathrm{tr}(Q_{D^n E^n}) &= \mathrm{tr}\big((\kappa_{D^n} \otimes \mathrm{id}_{E^n})^{-1}(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}\rho_{D^n E^n}(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}\big) \\
&= \mathrm{tr}\big(\kappa_{D^n}^{-1}(\sigma_D^{\otimes n})^{-\frac{1}{2}}\rho_{D^n}(\sigma_D^{\otimes n})^{-\frac{1}{2}}\big) \\
&= \mathrm{tr}\big(\kappa_{D^n}^{-1}(\sigma_D^{\otimes n})^{-\frac{1}{2}}\tau_{D^n}(\sigma_D^{\otimes n})^{-\frac{1}{2}}\big) \\
&= \mathrm{tr}\big((\kappa_{D^n} \otimes \mathrm{id}_{E^n})^{-1}(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}\tau_{D^n E^n}(\sigma_D^{\otimes n} \otimes \mathrm{id}_{E^n})^{-\frac{1}{2}}\big) \\
&= \mathrm{tr}\big((\kappa_{D^n} \otimes \mathrm{id}_{E^n})^{-1} T_{D^n E^n}\big) \\
&= \mathrm{tr}(S_{D^n E^n}) = \mathrm{tr}(\mathrm{id}_{\mathrm{Sym}^n(D \otimes E)}) \\
&= \dim(\mathrm{Sym}^n(D \otimes E)) \leq (n+1)^{d^2-1}. \tag{78}
\end{aligned}
$$

$\square$

Because any permutation-invariant density operator has a permutation-invariant purification, Lemma 3.1 can be easily extended so that $\rho_{D^n E^n}$ does not need to be pure.

Corollary 3.2. *Let D and E be Hilbert spaces and let σ_D be a non-negative operator on D. Then there exists a probability measure $\mathrm{d}\sigma_{DE}$ on the set of non-negative extensions σ_{DE} of σ_D such that*

$$\rho_{D^n E^n} \leq (n+1)^{d^2-1} \int \sigma_{DE}^{\otimes n}\mathrm{d}\sigma_{DE} \tag{79}$$

holds for any $n \in \mathbb{N}$, any permutation-invariant non-negative extension $\rho_{D^n E^n}$ of $\sigma_D^{\otimes n}$, and $d = \dim(D)\dim(E)^2$.

Proof. According to Lemma B.10, $\rho_{D^n E^n}$ has a permutation-invariant purification $\rho_{D^n E^n R^n}$ with purifying system $R^{\otimes n}$, where $\dim R \leq \dim(D \otimes E)$. Lemma 3.1 with E replaced by $E \otimes R$, applied to $\rho_{D^n E^n R^n}$, yields

$$\rho_{D^n E^n R^n} \leq (n+1)^{d^2-1} \int |\phi\rangle\langle\phi|_{DER}^{\otimes n}\mathrm{d}\phi, \tag{80}$$

where $\mathrm{d}\phi$ is a probability measure on the purifications $|\phi\rangle\langle\phi|_{DER}$ of σ_D and $d = \max[\dim D, \dim E \otimes R] \leq \dim D(\dim E)^2$. Taking the partial trace over $R^{\otimes n}$ on both sides gives

$$\rho_{D^n E^n} \leq (n+1)^{d^2-1} \int \mathrm{tr}_R(|\phi\rangle\langle\phi|_{DER})^{\otimes n}\mathrm{d}\phi. \tag{81}$$

The claim follows because the probability measure $\mathrm{d}\phi$ on the pure states $|\phi\rangle\langle\phi|_{DER}$ can be replaced by the induced measure $\mathrm{d}\sigma_{DE}$ on the marginal states $\sigma_{DE} = \mathrm{tr}_R(|\phi\rangle\langle\phi|_{DER})$. $\square$

Even though we do not need it here, it is worth pointing out that, by virtue of the Choi–Jamiołkowski isomorphism [12,34], the claim above can be rephrased in terms of completely positive trace-preserving maps. As shown in [6], this is useful to derive an improved variant of inequality (3).

Corollary 3.3. *Let D and E be Hilbert spaces. Then there exists a probability measure* $d\tau$ *on the set of completely positive trace-preserving maps $\tau_{D \to E}$ such that[7]*

$$\mathcal{W}_{D^n \to E^n} \leq (n+1)^{d^2-1} \int \tau_{D \to E}^{\otimes n} d\tau \tag{82}$$

holds for any $n \in \mathbb{N}$, any completely positive trace-preserving map $\mathcal{W}_{D^n \to E^n}$ that is permutation-invariant (i.e., $\mathcal{W} \circ \pi = \pi \circ \mathcal{W}$ for all permutations π), and $d = \dim(D) \dim(E)^2$.

Proof. Let $\rho_{D^n E^n} = J^{\otimes n}(\mathcal{W}_{D^n \to E^n})$ where J denotes the Choi–Jamiołkowski isomorphism on the mappings from D to E. That is, $\rho_{D^n E^n} = (\mathcal{W}_{\bar{D}^n \to E^n} \otimes \mathcal{I}_{D^n})(\Psi_{\bar{D}D}^{\otimes n})$, where $\Psi_{\bar{D}D}$ is a maximally entangled state. The marginal of $\rho_{D^n E^n}$ on D^n equals $\sigma_D^{\otimes n}$ with $\sigma_D = \frac{\mathrm{id}_D}{\dim D}$. Furthermore, as the map $\mathcal{W}_{\bar{D}^n \to E^n} \otimes \mathcal{I}_{D^n}$ is permutation invariant, so is the state $\rho_{D^n E^n}$. Hence (79) holds. Since the corresponding probability measure $d\sigma_{DE}$ is restricted to the set of density operators σ_{DE} with marginal σ_D, each σ_{DE} is the image of a trace-preserving completely positive map $\tau_{D \to E}$ under the isomorphism J, i.e., $\sigma_{DE} = J(\tau_{D \to E})$. The claim thus follows by applying the inverse of $J^{\otimes n}$ to both sides of (79). $\square$

4. Fidelity Between Permutation-Invariant Operators

The purpose of this section is to provide techniques to approximate the fidelity of permutation-invariant states. They play a key role in the proof of Theorem 5.1. The derivation of the statements below is based on the generalised de Finetti reduction method introduced in Sect. 3. Furthermore, we will use several established facts about the fidelity, which are summarised in Appendix B.

Lemma 4.1. *Let $\rho_{D^n E^n}$ be a permutation-invariant non-negative operator on $(D \otimes E)^{\otimes n}$ and let σ_D be a non-negative operator on D. Then there exists a non-negative extension σ_{DE} of σ_D on $D \otimes E$ such that*

$$F(\rho_{D^n E^n}, \sigma_{DE}^{\otimes n}) \geq (n+1)^{-d^2/2} F(\rho_{D^n}, \sigma_D^{\otimes n}), \tag{83}$$

where $d = \dim(D) \dim(E)^2$. Furthermore, if $\rho_{D^n E^n}$ is pure then σ_{DE} is pure and $d \leq \max[\dim(D), \dim(E)]$.

Proof. Let $\rho_{D^n E^n R^n} = |\Psi\rangle\langle\Psi|_{D^n E^n R^n}$ be a permutation-invariant purification of $\rho_{D^n E^n}$, i.e., $|\Psi\rangle \in \mathrm{Sym}^n(D \otimes E \otimes R)$, where $\dim(R) \leq \dim(D \otimes E)$. (That such a purification exists is the statement of Lemma B.10. We also note that, if $\rho_{D^n E^n}$ is already pure, then R can be chosen to be the trivial space $\mathbb{C}$, i.e., $\dim(R) = 1$.) According to Lemma B.11, there exists a permutation-invariant purification $\bar{\sigma}_{D^n E^n R^n}$ of $\sigma_D^{\otimes n}$ such that

$$F(\rho_{D^n}, \sigma_D^{\otimes n}) = F(\rho_{D^n E^n R^n}, \bar{\sigma}_{D^n E^n R^n}) = \sqrt{\langle\Psi|\bar{\sigma}_{D^n E^n R^n}|\Psi\rangle}. \tag{84}$$

---

[7] The inequality means that the difference between the right hand side and the left hand side is a completely positive map.

Let Γ be the set of vectors $|\phi\rangle_{DER}$ on $D \otimes E \otimes R$ such that $\mathrm{tr}_{ER}(|\phi\rangle\langle\phi|_{DER}) = \sigma_D$. According to Lemma 3.1 there exists a probability measure $\mathrm{d}\phi$ on Γ such that

$$\bar{\sigma}_{D^n E^n R^n} \le (n+1)^{d^2-1} \int |\phi\rangle\langle\phi|_{DER}^{\otimes n} \mathrm{d}\phi, \tag{85}$$

where $d = \max[\dim(D), \dim(E)\dim(R)] \le \dim(D)\dim(E)^2$. Using this we find

$$(n+1)^{-d^2} \langle\Psi|\bar{\sigma}_{D^n E^n R^n}|\Psi\rangle \le \int \langle\Psi|(|\phi\rangle\langle\phi|_{DER}^{\otimes n})|\Psi\rangle \mathrm{d}\phi \le \max_{|\phi\rangle \in \Gamma} \langle\Psi|(|\phi\rangle\langle\phi|_{DER}^{\otimes n})|\Psi\rangle. \tag{86}$$

We now set $\sigma_{DER} = |\phi\rangle\langle\phi|_{DER}$, where $|\phi\rangle_{DER} \in \Gamma$ is a vector that maximises the above expression. Note that, by the definition of the set Γ, σ_{DE} is then a valid extension of the given operator σ_D. (Furthermore, if R is the trivial space $\mathbb{C}$ then σ_{DE} is pure.) Combining (84) with (86), and using the monotonicity of the fidelity under the partial trace (Lemma B.4), we conclude that

$$(n+1)^{-d^2/2} F(\rho_{D^n}, \sigma_D^{\otimes n}) \le \sqrt{\langle\Psi|\sigma_{DER}^{\otimes n}|\Psi\rangle} = F(|\Psi\rangle\langle\Psi|_{D^n E^n R^n}, \sigma_{DER}^{\otimes n})$$
$$\le F(\rho_{D^n E^n}, \sigma_{DE}^{\otimes n}). \tag{87}$$

$\square$

Lemma 4.2. *Let $\rho_{R^n S^n}$ be a permutation-invariant non-negative operator on $(R \otimes S)^{\otimes n}$ and let σ_{RS} be a non-negative operator on $R \otimes S$. Furthermore, let W_{R^n} be a permutation-invariant operator on $R^{\otimes n}$ with $\|W_{R^n}\|_\infty \le 1$. Then there exists a unitary U_R on R such that[8]*

$$F\big(\rho_{R^n S^n}, U_R^{\otimes n} \sigma_{RS}^{\otimes n} (U_R^{\otimes n})^\dagger\big) \ge (n+1)^{-d^2} F\big(W_{R^n}\rho_{R^n S^n} W_{R^n}^\dagger, \sigma_{RS}^{\otimes n}\big), \tag{88}$$

where $d = \dim(R)\dim(S)^2$.

Proof. Let $\rho_{R^n S^n E^n}$ be a permutation-invariant purification of $\rho_{R^n S^n}$ on $(R \otimes S \otimes E)^{\otimes n}$, where $\dim(E) = \dim(R \otimes S)$ (cf. Lemma B.10). Then $W_{R^n}\rho_{R^n S^n E^n} W_{R^n}^\dagger$ is a permutation-invariant purification of $W_{R^n}\rho_{R^n S^n} W_{R^n}^\dagger$. Hence, according to Lemma 4.1, there exists a purification σ_{RSE} of σ_{RS} such that

$$(n+1)^{d_1^2/2} F(W_{R^n}\rho_{R^n S^n E^n} W_{R^n}^\dagger, \sigma_{RSE}^{\otimes n}) \ge F(W_{R^n}\rho_{R^n S^n} W_{R^n}^\dagger, \sigma_{RS}^{\otimes n}), \tag{89}$$

where $d_1 = \max[\dim(R \otimes S), \dim(E)] = \dim(R)\dim(S)$. We then use Lemma B.8 which asserts that

$$F(\rho_{S^n E^n}, \sigma_{SE}^{\otimes n}) \ge F(W_{R^n}\rho_{R^n S^n E^n} W_{R^n}^\dagger, \sigma_{RSE}^{\otimes n}). \tag{90}$$

Furthermore, again by Lemma 4.1, there exists a purification $\tilde{\sigma}_{RSE}$ of σ_{SE} such that

$$(n+1)^{d_2^2/2} F(\rho_{R^n S^n E^n}, \tilde{\sigma}_{RSE}^{\otimes n}) \ge F(\rho_{S^n E^n}, \sigma_{SE}^{\otimes n}), \tag{91}$$

---

[8] Here and in the following we simplify our notation by omitting identity operators that are clear from the context, e.g., we write U_R instead of $U_R \otimes \mathrm{id}_S$.

where $d_2 = \max[\dim(S \otimes E), \dim(R)] = \dim(R)\dim(S)^2$. Because all purifications are unitarily equivalent, there exists a unitary U_R on R such that $U_R \sigma_{RSE} U_R^\dagger = \tilde{\sigma}_{RSE}$, that is,

$$F\big(\rho_{R^n S^n E^n}, U_R^{\otimes n} \sigma_{RSE}^{\otimes n} (U_R^{\otimes n})^\dagger\big) = F(\rho_{R^n S^n E^n}, \tilde{\sigma}_{RSE}^{\otimes n}). \tag{92}$$

Because the fidelity is non-decreasing under the partial trace (cf. Lemma B.4), we also have

$$F\big(\rho_{R^n S^n}, U_R^{\otimes n} \sigma_{RS}^{\otimes n} (U_R^{\otimes n})^\dagger\big) \geq F\big(\rho_{R^n S^n E^n}, U_R^{\otimes n} \sigma_{RSE}^{\otimes n} (U_R^{\otimes n})^\dagger\big). \tag{93}$$

Combining all these equations, we obtain the desired claim. □

Remark 4.3. If for some orthonormal basis $\{|r\rangle\}_r$ of R the operator W_{R^n} is diagonal in the corresponding product basis $\{|r_1\rangle \otimes \cdots \otimes |r_n\rangle\}_{r_1,\ldots r_n}$ then inequality (88) also holds for an operator U_R which is diagonal in the basis $\{|r\rangle\}_r$ and satisfies $\|U_R\|_\infty \leq 1$, and for $d = \dim(R)^2 \dim(S)^2$.

To see this, let $\bar{R}$ be a system that is isomorphic to R and let C be the isometry from R to $R \otimes \bar{R}$ defined by

$$C = \sum_r \big(|r\rangle_R \otimes |r\rangle_{\bar{R}}\big)\langle r|_R. \tag{94}$$

It is straightforward to verify that, for W_{R^n} diagonal in the product basis $\{|r_1\rangle \otimes \cdots \otimes |r_n\rangle\}_{r_1,\ldots r_n}$, we have

$$W_{R^n} = (C^\dagger)^{\otimes n}(W_{R^n} \otimes \mathrm{id}_{\bar{R}^n})C^{\otimes n}. \tag{95}$$

Let now $\rho_{R^n S^n E^n}$ and σ_{RSE} be pure operators such that (89) holds. Furthermore, define $\bar{\rho}_{R^n \bar{R}^n S^n E^n} = C^{\otimes n} \rho_{R^n S^n E^n} (C^\dagger)^{\otimes n}$ and $\bar{\sigma}_{R\bar{R}SE} = C\sigma_{RSE}C^\dagger$. Using Lemma B.6 we find

$$F(W_{R^n} \rho_{R^n S^n E^n} W_{R^n}^\dagger, \sigma_{RSE}^{\otimes n}) = F\big((C^\dagger)^{\otimes n} W_{R^n} C^{\otimes n} \rho_{R^n S^n E^n} (C^\dagger)^{\otimes n} W_{R^n}^\dagger C^{\otimes n}, \sigma_{RSE}^{\otimes n}\big)$$
$$= F(W_{R^n} \bar{\rho}_{R^n \bar{R}^n S^n E^n} W_{R^n}^\dagger, \bar{\sigma}_{R\bar{R}SE}^{\otimes n}). \tag{96}$$

Furthermore, we can carry out the proof steps as in (90) and (91), while keeping the system $\bar{R}^{\otimes n}$, to obtain

$$(n+1)^{d^2/2} F(\bar{\rho}_{R^n \bar{R}^n S^n E^n}, \tilde{\sigma}_{R\bar{R}SE}^{\otimes n}) \geq F(\bar{\rho}_{\bar{R}^n S^n E^n}, \bar{\sigma}_{\bar{R}SE}^{\otimes n})$$
$$\geq F(W_{R^n} \bar{\rho}_{R^n \bar{R}^n S^n E^n} W_{R^n}^\dagger, \bar{\sigma}_{R\bar{R}SE}^{\otimes n}), \tag{97}$$

for some purification $\tilde{\sigma}_{R\bar{R}SE}$ of $\bar{\sigma}_{\bar{R}SE}$. Because $\bar{\sigma}_{R\bar{R}SE}$ is pure there must exist a unitary $\bar{U}_R$ on R such that $\bar{U}_R \bar{\sigma}_{R\bar{R}SE} \bar{U}_R^\dagger = \tilde{\sigma}_{R\bar{R}SE}$. Using this and the fact that the fidelity is non-decreasing under the partial trace we find

$$F\big(\bar{\rho}_{R^n \bar{R}^n S^n}, \bar{U}_R^{\otimes n} \bar{\sigma}_{R\bar{R}S}^{\otimes n}(\bar{U}_R^\dagger)^{\otimes n}\big) \geq F\big(\bar{\rho}_{R^n \bar{R}^n S^n E^n}, \bar{U}_R^{\otimes n} \bar{\sigma}_{R\bar{R}SE}^{\otimes n}(\bar{U}_R^\dagger)^{\otimes n}\big)$$
$$= F\big(\bar{\rho}_{R^n \bar{R}^n S^n E^n}, \tilde{\sigma}_{R\bar{R}SE}^{\otimes n}\big). \tag{98}$$

Finally, by the definition of $\bar{\rho}_{R^n \bar{R}^n S^n E^n}$ and $\bar{\sigma}_{R\bar{R}SE}$, and using Lemma B.6 we have

$$F\big(\bar{\rho}_{R^n \bar{R}^n S^n}, \bar{U}_R^{\otimes n} \bar{\sigma}_{R\bar{R}S}^{\otimes n}(\bar{U}_R^\dagger)^{\otimes n}\big) = F\big(\rho_{R^n S^n}, U_R^{\otimes n} \sigma_{RS}^{\otimes n}(U_R^\dagger)^{\otimes n}\big) \tag{99}$$

where $U_R = C^\dagger(\bar{U}_R \otimes \mathrm{id}_{\bar{R}})C$. Combining this with (89), (96), (97), and (98) we obtain again inequality (88). Furthermore, by construction, U_R is diagonal in the basis $\{|r\rangle\}_r$ and satisfies $U_R^\dagger U_R \leq \mathrm{id}_R$.

Remark 4.4. If $\rho_{R^n S^n}$ has product form $\rho_{RS}^{\otimes n}$, the statement of Lemma 4.2 can be rewritten as

$$F(\rho_{RS}, U_R \sigma_{RS} U_R^\dagger) \geq \sqrt[n]{(n+1)^{-d^2} F(W_{R^n} \rho_{RS}^{\otimes n} W_{R^n}^\dagger, \sigma_{RS}^{\otimes n})}. \tag{100}$$

Hence, for a family $\{W_{R^n}\}_{n \in \mathbb{N}}$ of permutation-invariant non-negative operators such that $\|W_{R^n}\|_\infty \leq 1$ we have

$$\sup_{U_R} F(\rho_{RS}, U_R \sigma_{RS} U_R^\dagger) \geq \limsup_{n \to \infty} \sqrt[n]{F(W_{R^n} \rho_{RS}^{\otimes n} W_{R^n}^\dagger, \sigma_{RS}^{\otimes n})}. \tag{101}$$

5. Main Result and Proof

Theorem 5.1. *For any density operator ρ_{ABC} on $A \otimes B \otimes C$, where A, B, and C are separable Hilbert spaces, there exists a trace-preserving completely positive map $\mathcal{T}_{B \to BC}$ from the space of operators on B to the space of operators on $B \otimes C$ such that[9]*

$$2^{-\frac{1}{2} I(A:C|B)_\rho} \leq F(\rho_{ABC}, (\mathcal{I}_A \otimes \mathcal{T}_{B \to BC})(\rho_{AB})). \tag{102}$$

Furthermore, if A, B, and C are finite-dimensional then $\mathcal{T}_{B \to BC}$ has the form

$$X_B \mapsto V_{BC} \rho_{BC}^{\frac{1}{2}} (\rho_B^{-\frac{1}{2}} U_B X_B U_B^\dagger \rho_B^{-\frac{1}{2}} \otimes \mathrm{id}_C) \rho_{BC}^{\frac{1}{2}} V_{BC}^\dagger \tag{103}$$

on the support of ρ_B, where U_B and V_{BC} are unitaries on B and $B \otimes C$, respectively.

Proof. We first note that, by Remark 5.2 below, it is sufficient to prove the statement for the case where A, B, and C are finite-dimensional. Let $\delta > 0$, $\delta' > 0$, and $\delta'' > 0$. Let $n \in \mathbb{N}$ and let $\{\Pi_b\}_{b \in \bar{B}_n}$ and $\{\Pi_d\}_{d \in \bar{D}_n}$ be the families of projectors onto the eigenspaces of $\rho_B^{\otimes n}$ and $\rho_{BC}^{\otimes n}$, labelled by their eigenvalues $b \in \bar{B}_n$ and $d \in \bar{D}_n$, respectively. Furthermore, let $\bar{B}_n^{\delta'}$ and $\bar{D}_n^{\delta''}$ be the subsets of $\bar{B}_n$ and $\bar{D}_n$ defined by Lemma 2.5 and define

$$\Pi_{B^n} = \sum_{b \in \bar{B}_n^{\delta'}} \Pi_b \quad \text{and} \quad \Pi_{B^n C^n} = \sum_{d \in \bar{D}_n^{\delta''}} \Pi_d. \tag{104}$$

Note that for any $\eta > 0$ we have

$$\mathrm{tr}(\Pi_{B^n} \rho_B^{\otimes n}) \geq 1 - \eta \quad \text{and} \quad \mathrm{tr}(\Pi_{B^n C^n} \rho_{BC}^{\otimes n}) \geq 1 - \eta \tag{105}$$

for n sufficiently large. Define the mapping on $(A \otimes B \otimes C)^{\otimes n}$

$$\mathcal{W}_n : X_{A^n B^n C^n} \mapsto (\mathrm{id}_{A^n} \otimes \Pi_{B^n C^n})(\mathrm{id}_{A^n} \otimes \Pi_{B^n} \otimes \mathrm{id}_{C^n})$$
$$\times X_{A^n B^n C^n}(\mathrm{id}_{A^n} \otimes \Pi_{B^n} \otimes \mathrm{id}_{C^n})(\mathrm{id}_{A^n} \otimes \Pi_{B^n C^n}). \tag{106}$$

as well as the abbreviation

$$\Gamma_{A^n B^n C^n} = \mathcal{W}_n(\rho_{ABC}^{\otimes n}) = \Pi_{B^n C^n} \Pi_{B^n} \rho_{ABC}^{\otimes n} \Pi_{B^n} \Pi_{B^n C^n}. \tag{107}$$

---

[9] $\mathcal{I}_A$ denotes the identity map on the space of operators on A. We include it here in our notation to stress that the map leaves the A system unaffected, but will usually omit it when it is clear from the context.

It is easily seen that the map $\mathcal{W}_n$ is trace non-increasing and completely positive. Furthermore, because of (105), we always have $\text{tr}(\mathcal{W}_n(\rho_{ABC}^{\otimes n})) = \text{tr}(\Gamma_{A^n B^n C^n}) > 2/3$ for η sufficiently small. One way to see this is using the gentle measurement lemma (see e.g., [54, Lemma 9.4.2]), which implies that

$$\text{tr}(\Pi_{B^n C^n} \Pi_{B^n} \rho_{ABC}^{\otimes n} \Pi_{B^n}) \geq \text{tr}(\Pi_{B^n C^n} \rho_{ABC}^{\otimes n}) - 2\sqrt{1 - \text{tr}(\Pi_{B^n} \rho_{ABC}^{\otimes n})} \geq 1 - \eta - 2\sqrt{\eta}. \tag{108}$$

Lemma 2.3 then tells us that, for n sufficiently large,[10]

$$\begin{aligned} D\big(\Gamma_{A^n B^n C^n} \| \Pi_{B^n C^n} \Pi_{B^n} \rho_{AB}^{\otimes n} \Pi_{B^n} \Pi_{B^n C^n}\big) \\ = D\big(\mathcal{W}_n(\rho_{ABC}^{\otimes n}) \| \mathcal{W}_n((\rho_{AB} \otimes \text{id}_C)^{\otimes n})\big) \\ \leq n\big(D(\rho_{ABC} \| \rho_{AB} \otimes \text{id}) + \frac{\delta}{2}\big) = n(-H(C|AB) + \frac{\delta}{2}), \end{aligned} \tag{109}$$

where the last equality is the definition of the conditional entropy, $H(C|AB) = -\text{tr}(\rho_{ABC} \log_2 \rho_{ABC}) + \text{tr}(\rho_{AB} \log_2 \rho_{AB}) = -D(\rho_{ABC} \| \rho_{AB} \otimes \text{id}_C)$. The relation between the fidelity and the relative entropy (Lemma B.2) now allows us to conclude that

$$\frac{1}{\text{tr}(\Gamma_{A^n B^n C^n})} F(\Gamma_{A^n B^n C^n}, \Pi_{B^n C^n} \Pi_{B^n} \rho_{AB}^{\otimes n} \Pi_{B^n} \Pi_{B^n C^n}) \geq 2^{\frac{1}{2} n(H(C|AB) - \frac{\delta}{2})}. \tag{110}$$

We now use Lemma B.6 to remove the projector $\Pi_{B^n C^n}$ from the second argument and note that the factor $\text{tr}(\Gamma_{A^n B^n C^n}) > 2/3$ can be absorbed by another factor $2^{-\frac{1}{4} n \delta}$ for n sufficiently large. This shows that

$$F(\Gamma_{A^n B^n C^n}, \Pi_{B^n} \rho_{AB}^{\otimes n} \Pi_{B^n}) \geq 2^{\frac{1}{2} n(H(C|AB) - \delta)}. \tag{111}$$

Because $\sum_{b \in \bar{B}_n} \Pi_b = \text{id}_{B^n}$ we can apply Lemma B.7, which gives

$$\begin{aligned} F(\Gamma_{A^n B^n C^n}, \Pi_{B^n} \rho_{AB}^{\otimes n} \Pi_{B^n}) &\leq \sum_{b \in \bar{B}_n} F(\Gamma_{A^n B^n C^n}, \Pi_b \Pi_{B^n} \rho_{AB}^{\otimes n} \Pi_{B^n} \Pi_b) \\ &= \sum_{b \in \bar{B}_n^{\delta'}} F(\Gamma_{A^n B^n C^n}, \Pi_b \rho_{AB}^{\otimes n} \Pi_b) \\ &\leq |\bar{B}_n^{\delta'}| \max_{b \in \bar{B}_n^{\delta'}} F(\Gamma_{A^n B^n C^n}, \Pi_b \rho_{AB}^{\otimes n} \Pi_b), \end{aligned} \tag{112}$$

where the equality follows from

$$\Pi_b \Pi_{B^n} = \begin{cases} \Pi_b & \text{if } b \in \bar{B}_n^{\delta'} \\ 0 & \text{otherwise.} \end{cases} \tag{113}$$

Hence, there exists $b \in \bar{B}_n^{\delta'}$ such that

$$\begin{aligned} F(\Gamma_{A^n B^n C^n}, \Pi_{B^n} \rho_{AB}^{\otimes n} \Pi_{B^n}) &\leq |\bar{B}_n^{\delta'}| F(\Gamma_{A^n B^n C^n}, \Pi_b \rho_{AB}^{\otimes n} \Pi_b) \\ &\leq \text{poly}(n) F(\Gamma_{A^n B^n C^n}, \Pi_b \rho_{AB}^{\otimes n} \Pi_b), \end{aligned} \tag{114}$$

---

[10] We note that a similar conclusion may be obtained from the inequality given in Remark 2.4.

where the second inequality follows from Remark 2.6. By the definition of Π_b we also have

$$\Pi_b = \sqrt{b}(\rho_B^{-\frac{1}{2}})^{\otimes n}\Pi_b, \tag{115}$$

where b is the eigenvalue of $\rho_B^{\otimes n}$ corresponding to Π_b. By the definition of $\bar{B}_n^{\delta'}$ we also have $\sqrt{b} \leq 2^{-\frac{1}{2}n(H(B)-\delta')}$ and, hence,

$$
\begin{aligned}
F(\Gamma_{A^n B^n C^n}, \Pi_b \rho_{AB}^{\otimes n}\Pi_b) &= \sqrt{b}\,F\big(\Gamma_{A^n B^n C^n}, (\rho_B^{-\frac{1}{2}})^{\otimes n}\Pi_b\rho_{AB}^{\otimes n}\Pi_b(\rho_B^{-\frac{1}{2}})^{\otimes n}\big) \\
&\leq 2^{-\frac{1}{2}n(H(B)-\delta')}F\big(\Gamma_{A^n B^n C^n}, (\rho_B^{-\frac{1}{2}})^{\otimes n}\Pi_b\rho_{AB}^{\otimes n}\Pi_b(\rho_B^{-\frac{1}{2}})^{\otimes n}\big) \\
&= 2^{-\frac{1}{2}n(H(B)-\delta')}F\big(\Pi_b(\rho_B^{-\frac{1}{2}})^{\otimes n}\Gamma_{A^n B^n C^n}(\rho_B^{-\frac{1}{2}})^{\otimes n}\Pi_b, \rho_{AB}^{\otimes n}\big),
\end{aligned}
\tag{116}
$$

where the equality follows from Lemma B.6, which we will use repeatedly in the following. Furthermore, by Lemma 4.2, there must exist a unitary U_B on B such that

$$
\begin{aligned}
&F\big(\Pi_b(\rho_B^{-\frac{1}{2}})^{\otimes n}\Gamma_{A^n B^n C^n}(\rho_B^{-\frac{1}{2}})^{\otimes n}\Pi_b, \rho_{AB}^{\otimes n}\big) \\
&\leq \mathrm{poly}(n)F\big((\rho_B^{-\frac{1}{2}})^{\otimes n}\Gamma_{A^n B^n C^n}(\rho_B^{-\frac{1}{2}})^{\otimes n}, U_B^{\otimes n}\rho_{AB}^{\otimes n}(U_B^{\otimes n})^{\dagger}\big) \\
&= \mathrm{poly}(n)F\big(\Gamma_{A^n B^n C^n}, (\rho_B^{-\frac{1}{2}})^{\otimes n}U_B^{\otimes n}\rho_{AB}^{\otimes n}(U_B^{\otimes n})^{\dagger}(\rho_B^{-\frac{1}{2}})^{\otimes n}\big).
\end{aligned}
\tag{117}
$$

Combining now (111), (114), (116), and (117) we obtain

$$
\begin{aligned}
2^{\frac{1}{2}n(H(C|AB)+H(B)-\delta-\delta')} &\leq \mathrm{poly}(n)F\big(\Gamma_{A^n B^n C^n}, (\rho_B^{-\frac{1}{2}})^{\otimes n}U_B^{\otimes n}\rho_{AB}^{\otimes n}(U_B^{\otimes n})^{\dagger}(\rho_B^{-\frac{1}{2}})^{\otimes n}\big)\big) \\
&= \mathrm{poly}(n)F(\Pi_{B^n C^n}\Pi_{B^n}\rho_{ABC}^{\otimes n}\Pi_{B^n}\Pi_{B^n C^n}, \gamma_{ABC}^{\otimes n}),
\end{aligned}
\tag{118}
$$

where $\gamma_{ABC} = \rho_B^{-\frac{1}{2}}U_B\rho_{AB}U_B^{\dagger}\rho_B^{-\frac{1}{2}}$.

Next we use that $\sum_{d\in\bar{D}_n}\Pi_d = \mathrm{id}_{B^n C^n}$ and apply again Lemma B.7 to obtain

$$
\begin{aligned}
&F(\Pi_{B^n C^n}\Pi_{B^n}\rho_{ABC}^{\otimes n}\Pi_{B^n}\Pi_{B^n C^n}, \gamma_{ABC}^{\otimes n}) \\
&\leq \sum_{d\in\bar{D}_n} F(\Pi_d\Pi_{B^n C^n}\Pi_{B^n}\rho_{ABC}^{\otimes n}\Pi_{B^n}\Pi_{B^n C^n}\Pi_d, \gamma_{ABC}^{\otimes n}) \\
&= \sum_{d\in\bar{D}_n^{\delta''}} F(\Pi_d\Pi_{B^n}\rho_{ABC}^{\otimes n}\Pi_{B^n}\Pi_d, \gamma_{ABC}^{\otimes n}) \\
&\leq |\bar{D}_n^{\delta''}| \max_{d\in\bar{D}_n^{\delta''}} F(\Pi_d\Pi_{B^n}\rho_{ABC}^{\otimes n}\Pi_{B^n}\Pi_d, \gamma_{ABC}^{\otimes n}),
\end{aligned}
\tag{119}
$$

where $|\bar{D}_n^{\delta''}| \leq \mathrm{poly}(n)$ by Remark 2.6. Hence, there exists $d \in \bar{D}_n^{\delta''}$ such that

$$F(\Pi_{B^n C^n}\Pi_{B^n}\rho_{ABC}^{\otimes n}\Pi_{B^n}\Pi_{B^n C^n}, \gamma_{ABC}^{\otimes n}) \leq \mathrm{poly}(n)F(\Pi_d\Pi_{B^n}\rho_{ABC}^{\otimes n}\Pi_{B^n}\Pi_d, \gamma_{ABC}^{\otimes n}). \tag{120}$$

By the definition of $\bar{D}_n^{\delta''}$ we have

$$\Pi_d = \frac{1}{\sqrt{d}}(\rho_{BC}^{\frac{1}{2}})^{\otimes n}\Pi_d \tag{121}$$

with $d \geq 2^{-n(H(BC)+\delta'')}$. This implies

$$
\begin{aligned}
& F(\Pi_d \Pi_{B^n} \rho_{ABC}^{\otimes n} \Pi_{B^n} \Pi_d, \gamma_{ABC}^{\otimes n}) \\
&= \sqrt{\frac{1}{d}} F\big((\rho_{BC}^{\frac{1}{2}})^{\otimes n} \Pi_d \Pi_{B^n} \rho_{ABC}^{\otimes n} \Pi_{B^n} \Pi_d (\rho_{BC}^{\frac{1}{2}})^{\otimes n}, \gamma_{ABC}^{\otimes n}\big) \\
&\leq 2^{\frac{1}{2} n(H(BC)+\delta'')} F\big((\rho_{BC}^{\frac{1}{2}})^{\otimes n} \Pi_d \Pi_{B^n} \rho_{ABC}^{\otimes n} \Pi_{B^n} \Pi_d (\rho_{BC}^{\frac{1}{2}})^{\otimes n}, \gamma_{ABC}^{\otimes n}\big) \\
&= 2^{\frac{1}{2} n(H(BC)+\delta'')} F\big(\Pi_d \Pi_{B^n} \rho_{ABC}^{\otimes n} \Pi_{B^n} \Pi_d, (\rho_{BC}^{\frac{1}{2}})^{\otimes n} \gamma_{ABC}^{\otimes n} (\rho_{BC}^{\frac{1}{2}})^{\otimes n}\big). \quad (122)
\end{aligned}
$$

We use again Lemma 4.2, which asserts that there must exist a unitary V_{BC} on $B \otimes C$ such that

$$
\begin{aligned}
& F\big(\Pi_d \Pi_{B^n} \rho_{ABC}^{\otimes n} \Pi_{B^n} \Pi_d, (\rho_{BC}^{\frac{1}{2}})^{\otimes n} \gamma_{ABC}^{\otimes n} (\rho_{BC}^{\frac{1}{2}})^{\otimes n}\big) \\
&\leq \operatorname{poly}(n) F\big(\rho_{ABC}^{\otimes n}, V_{BC}^{\otimes n} (\rho_{BC}^{\frac{1}{2}})^{\otimes n} \gamma_{ABC}^{\otimes n} (\rho_{BC}^{\frac{1}{2}})^{\otimes n} (V_{BC}^{\otimes n})^{\dagger}\big). \quad (123)
\end{aligned}
$$

Combining this with (118), (120) and (122) yields

$$
\begin{aligned}
& 2^{\frac{1}{2} n(H(C|AB)+H(B)-H(BC)-\delta-\delta'-\delta'')} \\
&\leq \operatorname{poly}(n) F\big(\rho_{ABC}^{\otimes n}, V_{BC}^{\otimes n} (\rho_{BC}^{\frac{1}{2}})^{\otimes n} \gamma_{ABC}^{\otimes n} (\rho_{BC}^{\frac{1}{2}})^{\otimes n} (V_{BC}^{\otimes n})^{\dagger}\big). \quad (124)
\end{aligned}
$$

We take the nth root, use $H(BC) - H(B) - H(C|AB) = I(A : C|B)$, and insert the expression for γ_{ABC} to rewrite this as

$$
\begin{aligned}
2^{-\frac{1}{2} I(A:C|B)-\delta-\delta'-\delta''} &\leq \sqrt[n]{\operatorname{poly}(n)} F(\rho_{ABC}, V_{BC} \rho_{BC}^{\frac{1}{2}} \rho_B^{-\frac{1}{2}} U_B \rho_{AB} U_B^{\dagger} \rho_B^{-\frac{1}{2}} \rho_{BC}^{\frac{1}{2}} V_{BC}^{\dagger}) \\
&\leq \sqrt[n]{\operatorname{poly}(n)} \max_{U_B, V_{BC}} F(\rho_{ABC}, V_{BC} \rho_{BC}^{\frac{1}{2}} \rho_B^{-\frac{1}{2}} U_B \rho_{AB} U_B^{\dagger} \rho_B^{-\frac{1}{2}} \\
&\quad \times \rho_{BC}^{\frac{1}{2}} V_{BC}^{\dagger}), \quad (125)
\end{aligned}
$$

where the maximum is take over unitary transformations U_B and V_{BC}. As this maximised fidelity is now independent of n and because $\sqrt[n]{\operatorname{poly}(n)}$ approaches 1 for n large, we conclude that

$$
2^{-\frac{1}{2} I(A:C|B)-\delta-\delta'-\delta''} \leq \max_{U_B, V_{BC}} F(\rho_{ABC}, V_{BC} \rho_{BC}^{\frac{1}{2}} \rho_B^{-\frac{1}{2}} U_B \rho_{AB} U_B^{\dagger} \rho_B^{-\frac{1}{2}} \rho_{BC}^{\frac{1}{2}} V_{BC}^{\dagger}). \quad (126)
$$

Inequality (102) now follows because $\delta > 0$, $\delta' > 0$, and $\delta'' > 0$ were arbitrary.

It remains to verify that the map $\mathcal{T}_{B \to BC}$ is trace-preserving. But this follows from the observation that

$$
\begin{aligned}
\operatorname{tr}_C(U_B^{\dagger} \rho_B^{-\frac{1}{2}} \rho_{BC}^{\frac{1}{2}} V_{BC}^{\dagger} V_{BC} \rho_{BC}^{\frac{1}{2}} \rho_B^{-\frac{1}{2}} U_B) &= \operatorname{tr}_C(U_B^{\dagger} \rho_B^{-\frac{1}{2}} \rho_{BC} \rho_B^{-\frac{1}{2}} U_B) \\
&= U_B^{\dagger} \rho_B^{-\frac{1}{2}} \rho_B \rho_B^{-\frac{1}{2}} U_B = \operatorname{id}_B. \quad (127)
\end{aligned}
$$

$\square$

Remark 5.2. Any proof of the main claim of Theorem 5.1,

$$2^{-\frac{1}{2}I(A:C|B)_\rho} \leq \sup_{\mathcal{T}_{B\to BC}} F\big(\rho_{ABC}, (\mathcal{I}_A \otimes \mathcal{T}_{B\to BC})(\rho_{AB})\big), \tag{128}$$

which uses the assumption that A, B, and C are finite-dimensional Hilbert spaces, implies that the claim also holds under the less restrictive assumption that these spaces are separable.

To see this, let $\{P_A^k\}_{k\in\mathbb{N}}$, $\{P_B^k\}_{k\in\mathbb{N}}$, and $\{P_C^k\}_{k\in\mathbb{N}}$ be sequences of finite-rank projectors on A, B, and C which converge to id_A, id_B, and id_C, respectively, with respect to the weak (and, hence, also the strong) operator topology (see, e.g., Definition 2 of [25]). Define furthermore the density operators

$$\rho_{ABC}^{k,k'} = \frac{(P_A^k \otimes P_B^{k'} \otimes P_C^k)\rho_{ABC}(P_A^k \otimes P_B^{k'} \otimes P_C^k)}{\mathrm{tr}\big((P_A^k \otimes P_B^{k'} \otimes P_C^k)\rho_{ABC}\big)} \tag{129}$$

and

$$\rho_{ABC}^{k} = \frac{(P_A^k \otimes \mathrm{id}_B \otimes P_C^k)\rho_{ABC}(P_A^k \otimes \mathrm{id}_B \otimes P_C^k)}{\mathrm{tr}\big((P_A^k \otimes \mathrm{id}_B \otimes P_C^k)\rho_{ABC}\big)}. \tag{130}$$

We note that, for any $k \in \mathbb{N}$, the sequence $\{\rho_{ABC}^{k,k'}\}_{k'\in\mathbb{N}}$ converges to ρ_{ABC}^k in the trace norm (see, e.g., Corollary 2 of [25]). Also, $\{\rho_{ABC}^k\}_{k\in\mathbb{N}}$ converges to ρ_{ABC} in the trace norm.

Let us first consider the left hand side of (128). Because, for any fixed finite dimension of system A, the conditional mutual information $I(A : C|B)_\rho = H(A|B)_\rho - H(A|BC)_\rho$ is continuous in ρ with respect to the trace norm [1], we have

$$\lim_{k'\to\infty} I(A : C|B)_{\rho^{k,k'}} = I(A : C|B)_{\rho^k} \tag{131}$$

for any $k \in \mathbb{N}$. In addition, using the fact that local projectors applied to the subsystems A and C can only decrease $I(A : C|B)_\rho$, provided we scale by the probability of such a projector,

$$\mathrm{tr}\big((P_A^k \otimes \mathrm{id}_B \otimes P_C^k)\rho_{ABC}\big)I(A : C|B)_{\rho^k} \leq I(A : C|B)_\rho \tag{132}$$

holds for any $k \in \mathbb{N}$.

Because $\lim_{k\to\infty} \mathrm{tr}\big((P_A^k \otimes \mathrm{id}_B \otimes P_C^k)\rho_{ABC}\big) = \mathrm{tr}(\rho) = 1$, we find

$$\limsup_{k\to\infty} I(A : C|B)_{\rho^k} \leq I(A : C|B)_\rho. \tag{133}$$

The combination of this statement with (131) yields

$$2^{-\frac{1}{2}I(A:C|B)_\rho} \leq \liminf_{k\to\infty} \lim_{k'\to\infty} 2^{-\frac{1}{2}I(A:C|B)_{\rho^{k,k'}}}. \tag{134}$$

We now consider the right hand side of (128). Let $\delta > 0$ and note that, for sufficiently large k and k', we have

$$\big\|\rho_{ABC}^{k,k'} - \rho_{ABC}\big\|_1 < (\delta/2)^2. \tag{135}$$

Because the trace norm is monotonically non-increasing under trace-preserving completely positive maps, we also have

$$\left\| \mathcal{T}_{B\to BC}(\rho_{AB}^{k,k'}) - \mathcal{T}_{B\to BC}(\rho_{AB}) \right\|_1 < (\delta/2)^2 \tag{136}$$

for any $\mathcal{T}_{B\to BC}$. Lemma B.9 then implies that

$$F\left(\rho_{ABC}^{k,k'}, \mathcal{T}_{B\to BC}(\rho_{AB}^{k,k'})\right) < F\left(\rho_{ABC}, \mathcal{T}_{B\to BC}(\rho_{AB})\right) + \delta. \tag{137}$$

But because this holds for any $\mathcal{T}_{B\to BC}$, we have

$$\sup_{\mathcal{T}_{B\to BC}} F\left(\rho_{ABC}^{k,k'}, \mathcal{T}_{B\to BC}(\rho_{AB}^{k,k'})\right) \leq \sup_{\mathcal{T}_{B\to BC}} F\left(\rho_{ABC}, \mathcal{T}_{B\to BC}(\rho_{AB})\right) + \delta. \tag{138}$$

Because this holds for all $\delta > 0$ and sufficiently large k and k', we find that

$$\limsup_{k\to\infty} \limsup_{k'\to\infty} \sup_{\mathcal{T}_{B\to BC}} F\left(\rho_{ABC}^{k,k'}, \mathcal{T}_{B\to BC}(\rho_{AB}^{k,k'})\right) \leq \sup_{\mathcal{T}_{B\to BC}} F\left(\rho_{ABC}, \mathcal{T}_{B\to BC}(\rho_{AB})\right). \tag{139}$$

To conclude the argument, we observe that if the inequality (128) is valid for finite-dimensional spaces A, B, and C we have in particular

$$\liminf_{k\to\infty} \lim_{k'\to\infty} 2^{-\frac{1}{2}I(A:C|B)_{\rho^{k,k'}}} \leq \limsup_{k\to\infty} \limsup_{k'\to\infty} \sup_{\mathcal{T}_{B\to BC}} F\left(\rho_{ABC}^{k,k'}, \mathcal{T}_{B\to BC}(\rho_{AB}^{k,k'})\right). \tag{140}$$

Combining this with (134) and (139) then proves the claim that the inequality holds for arbitrary separable spaces A, B, and C.

Remark 5.3. By Remark 4.3, the unitary U_B chosen in (117) may be replaced by an operator which commutes with ρ_B and satisfies $\|U_B\|_\infty \leq 1$. Analogously, the unitary V_{BC} chosen in (123) may be replaced by an operator of the form $V_{BC} = V'_B V''_{BC}$ where V'_B commutes with ρ_B and V''_{BC} commutes with ρ_{BC}, and where $\|V'_B\|_\infty \leq 1$ and $\|V''_{BC}\|_\infty \leq 1$. Similarly to (127) one can see that the resulting recovery map $\mathcal{T}_{B\to BC}$ is trace non-increasing. Furthermore, we have

$$\begin{aligned}
\mathcal{T}_{B\to BC}(\rho_B) &= V_{BC}\rho_{BC}^{\frac{1}{2}}(\rho_B^{-\frac{1}{2}}U_B\rho_B U_B^\dagger \rho_B^{-\frac{1}{2}} \otimes \mathrm{id}_C)\rho_{BC}^{\frac{1}{2}}V_{BC}^\dagger \\
&= V_{BC}\rho_{BC}^{\frac{1}{2}}(U_B U_B^\dagger \otimes \mathrm{id}_C)\rho_{BC}^{\frac{1}{2}}V_{BC}^\dagger \\
&\leq V_{BC}\rho_{BC}^{\frac{1}{2}}\rho_{BC}^{\frac{1}{2}}V_{BC}^\dagger = V'_B\rho_{BC}^{\frac{1}{2}}V''_{BC}(V''_{BC})^\dagger \rho_{BC}^{\frac{1}{2}}(V'_B)^\dagger \leq V'_B\rho_{BC}(V'_B)^\dagger.
\end{aligned} \tag{141}$$

In particular, we have

$$\mathrm{tr}_C\left(\mathcal{T}_{B\to BC}(\rho_B)\right) \leq \rho_B \quad \text{and} \quad \mathrm{tr}_B\left(\mathcal{T}_{B\to BC}(\rho_B)\right) \leq \rho_C. \tag{142}$$

This implies that one can always choose a recovery map that exactly reproduces the marginal on B and the marginal on C.

Appendices

A. One-Shot Relative Entropies

In this appendix, we briefly review the *generalised relative entropy* introduced in [21,53] and the *smooth max-relative entropy* introduced in [19] (we will use a slightly modified variant defined in [49,50]).

Definition A.1. For any two non-negative operators ρ and σ and for any $\epsilon \in [0, \text{tr}(\rho)]$, the *generalised relative entropy* is defined by[11]

$$2^{-D_H^\epsilon(\rho\|\sigma)} = \inf_{\substack{0 \leq Q \leq 1 \\ \text{tr}(Q\rho) \geq \epsilon}} \text{tr}(Q\sigma)/\epsilon \,, \tag{143}$$

where the optimisation is over operators Q. For ρ a density operator,[12] the ϵ-*smooth max-relative entropy* is defined by

$$2^{-D_{\max}^\epsilon(\rho\|\sigma)} = \sup_{\substack{\mu\bar{\rho} \leq \sigma \\ 1-F(\bar{\rho},\rho)^2 \leq \epsilon^2}} \mu, \tag{144}$$

where the optimisation is over non-negative operators $\bar{\rho}$ with $\text{tr}(\bar{\rho}) \leq 1$.[13]

Remark A.2. The second argument, σ, of the two one-shot entropy measures of Definition A.1 may be rescaled easily because

$$D_H^\epsilon(\rho\|\lambda\sigma) = D_H^\epsilon(\rho\|\sigma) - \log_2(\lambda) \tag{145}$$
$$D_{\max}^\epsilon(\rho\|\lambda\sigma) = D_{\max}^\epsilon(\rho\|\sigma) - \log_2(\lambda) \tag{146}$$

holds for any $\lambda > 0$.

The generalised relative entropy may be expressed equivalently as follows.

Lemma A.3. *For any two non-negative operators ρ and σ,*

$$2^{-D_H^\epsilon(\rho\|\sigma)} = \sup_{\substack{\mu(\rho-Y) \leq \sigma \\ Y \geq 0 \\ \mu \geq 0}} \mu(1 - \text{tr}(Y)/\epsilon). \tag{147}$$

where the optimisation is over operators Y and reals μ.

Proof. As shown in [21], the expression on the right hand side of (143) is a semidefinite program whose dual has the form

$$2^{-D_H^\epsilon(\rho\|\sigma)} = \sup_{\substack{\mu\rho \leq \sigma+X \\ X \geq 0 \\ \mu \geq 0}} \mu - \text{tr}(X)/\epsilon, \tag{148}$$

---

[11] For $\epsilon = 0$ the quantity is defined via continuous extension, i.e., $D_H^0(\rho\|\sigma) = \lim_{\epsilon\downarrow 0} D_H^\epsilon(\rho\|\sigma)$.

[12] We note that the definition proposed in [49,50] applies more generally to any ρ with $\text{tr}(\rho) \leq 1$.

[13] The expressions on the right hand side of (143) and (144) may be equal to 0, in which case the corresponding relative entropy is defined to be equal to ∞.

where the optimisation is over operators X and reals μ. Replacing X by μY we can rewrite this as

$$2^{-D_H^\epsilon(\rho\|\sigma)} = \sup_{\substack{\mu(\rho-Y)\leq\sigma \\ \mu Y\geq 0 \\ \mu\geq 0}} \mu(1 - \mathrm{tr}(Y)/\epsilon). \tag{149}$$

To conclude the proof, we note that the condition $\mu Y \geq 0$ can be replaced by $Y \geq 0$ because $\mu \geq 0$ and because for $\mu = 0$ the value of Y is irrelevant. $\qquad\square$

Remark A.4. It is obvious from this representation that $D_H^\epsilon(\rho\|\sigma)$ is a monotonically non-increasing function in ϵ.

The following lemma provides an upper bound on $D_H^\epsilon(\rho\|\sigma)$, expressed in terms of the trace distance of ρ to an operator $\bar\rho$ that satisfies a simple operator inequality.

Lemma A.5. *Let ρ, $\bar\rho$, and σ be non-negative operators and suppose that $\bar\rho \leq \lambda\sigma$ for some $\lambda > 0$. Then*

$$D_H^\epsilon(\rho\|\sigma) \leq \log_2(\lambda) - \log_2\big(1 - \Delta(\rho, \bar\rho)/\epsilon\big). \tag{150}$$

Proof. Let $Y^+ \geq 0$ and $Y^- \geq 0$ be the positive and negative parts of $\rho - \bar\rho$, respectively, so that $\rho - \bar\rho = Y^+ - Y^-$ and $\mathrm{tr}(Y^+) \leq \Delta(\rho, \bar\rho)$ (see Eq. 18). We then have, in particular, $\rho - Y^+ \leq \bar\rho$ and, using the assumption that $\bar\rho \leq \lambda\sigma$,

$$\rho - Y^+ \leq \lambda\sigma. \tag{151}$$

This means that $\mu = \frac{1}{\lambda}$ and $Y = Y^+$ fulfill the constraints of the maximisation in (147) for $D_H^\epsilon(\rho\|\sigma)$ and, hence,

$$2^{-D_H^\epsilon(\rho\|\sigma)} \geq \frac{1}{\lambda}\Big(1 - \frac{\mathrm{tr}(Y^+)}{\epsilon}\Big) \geq \frac{1}{\lambda}\Big(1 - \frac{\Delta(\rho, \bar\rho)}{\epsilon}\Big). \tag{152}$$

Taking the negative logarithm on both sides of the inequality yields the claim. $\qquad\square$

Although we are not using this for our argument, we note that Lemma A.5 can be extended to a relation between the smooth relative max-entropy and the generalised relative entropy.

Lemma A.6. *Let ρ be a density operator, let σ be a non-negative operator, and let $\epsilon > \epsilon' \geq 0$. Then*

$$D_H^\epsilon(\rho\|\sigma) \leq D_{\max}^{\epsilon'}(\rho\|\sigma) + \log_2\frac{\epsilon}{\epsilon - \epsilon'}. \tag{153}$$

Proof. Let $\mu = 2^{-D_{\max}^{\epsilon'}(\rho\|\sigma)}$ and let $\bar\rho$ be such that the expression on the right hand side of (144) is satisfied for $D_{\max}^{\epsilon'}(\rho\|\sigma)$. That is, we have $\bar\rho \leq \sigma/\mu$ as well as $\sqrt{1 - F(\bar\rho, \rho)^2} \leq \epsilon'$, which, by Lemma B.1, implies $\Delta(\bar\rho, \rho) \leq \epsilon'$. Hence, by Lemma A.5,

$$D_H^\epsilon(\rho\|\sigma) \leq -\log_2(\mu) - \log_2\big(1 - (\epsilon'/\epsilon)\big) = D_{\max}^{\epsilon'}(\rho\|\sigma) + \log_2\frac{\epsilon}{\epsilon - \epsilon'}. \tag{154}$$

$\qquad\square$

The following claim about the smooth relative max-entropy for product states is known as the *Quantum Asymptotic Equipartition Property* [50]. (Note that this is a strictly more general statement than the "classical" Asymptotic Equipartition Property stated as Lemma 2.5.) While the proof in [50] applies to the case where σ is a density operator, the slightly extended claim provided here follows directly from Remark A.2 (see also Footnote 9 of [50] as well as Chapter 6 of [49]).

Lemma A.7. *For any density operator ρ, for any non-negative operator σ, for any $\epsilon \in (0, 1)$, and for sufficiently large $n \in \mathbb{N}$,*

$$\frac{1}{n} D_{\max}^{\epsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) < D(\rho \| \sigma) + c \sqrt{\frac{\log_2(2/\epsilon^2)}{n}}, \tag{155}$$

where $c = c(\rho, \sigma)$ is independent of n and ϵ.

Because of Lemma A.6, almost the same upper bound also holds for $D_H^{\epsilon}(\cdot \| \cdot)$. In fact, as a consequence of the Quantum Stein's Lemma [29,42], the statement holds asymptotically with equality [21].

Lemma A.8. *Let ρ be a density operator, let σ be a non-negative operator, and let $\epsilon \in (0, 1)$. Then*

$$\lim_{n \to \infty} \frac{1}{n} D_H^{\epsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho \| \sigma). \tag{156}$$

B. General Facts About the Fidelity

In the literature, the definition and discussion of the fidelity $F(\rho, \sigma)$ is often restricted to the case where its arguments, ρ and σ, are density operators (see, e.g., Chapter 9 of [41]). In this work, however, we need the fidelity for general non-negative operators. Recall that we defined $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$. Fortunately, most established properties of the fidelity are still valid in this more general case. For completeness, we state them in the following.

Lemma B.1. *For any two non-negative operators ρ and σ with $\mathrm{tr}(\rho) \geq \mathrm{tr}(\sigma)$, the trace distance is upper bounded by*

$$\Delta(\rho, \sigma) \leq \sqrt{\mathrm{tr}(\rho)^2 - F(\rho, \sigma)^2}. \tag{157}$$

The following lemma relates the relative entropy to the fidelity (see also Sect. 5.4 of [27]).

Lemma B.2. *For any non-negative operators ρ and σ*

$$D(\rho \| \sigma) \geq -2 \log_2 \frac{F(\rho, \sigma)}{\mathrm{tr}(\rho)}. \tag{158}$$

Proof. Let $D_\alpha(\cdot\|\cdot)$ be the α-*Quantum Rényi Divergence* as defined in [40,55]. As shown in these papers, for $\alpha = 1$ it is identical to the relative entropy, i.e.,

$$D_1(\rho\|\sigma) = D(\rho\|\sigma). \tag{159}$$

For $\alpha = 1/2$, it is related to the fidelity via

$$D_{\frac{1}{2}}(\rho\|\sigma) = -2\log_2\frac{F(\rho,\sigma)}{\mathrm{tr}(\rho)}. \tag{160}$$

Finally, $\alpha \mapsto D_\alpha(\rho\|\sigma)$ is a monotonically non-decreasing function in α. Combining these statements, we find

$$-2\log_2 F(\rho,\sigma) = D_{\frac{1}{2}}(\rho\|\sigma) \le D_1(\rho\|\sigma) = D(\rho\|\sigma). \tag{161}$$

$\square$

Next we recall a statement that is known as Uhlmann's theorem [52].

Lemma B.3. *Let $\rho_{DR} = |\psi\rangle\langle\psi|_{DR}$ and $\sigma_{DR} = |\phi\rangle\langle\phi|_{DR}$ be purifications of non-negative operators $\rho = \rho_D$ and $\sigma = \sigma_D$, respectively. Then*

$$F(\rho,\sigma) = \sup_{U_R}\big|\langle\psi|(\mathrm{id}_D \otimes U_R)|\phi\rangle\big|, \tag{162}$$

where the maximisation is over all unitaries U_R on R.

The following lemma is a direct consequence of Lemma B.3. It asserts that the fidelity is monotonically non-decreasing when a partial trace is applied to both arguments.

Lemma B.4. *For any two non-negative operators ρ_{DE} and σ_{DE} we have*

$$F(\rho_D, \sigma_D) \ge F(\rho_{DE}, \sigma_{DE}). \tag{163}$$

Using the Stinespring dilation theorem, the statement can be brought into the following more general form.

Lemma B.5. *For any trace-preserving completely positive map $\mathcal{T}$ we have*

$$F(\mathcal{T}(\rho), \mathcal{T}(\sigma)) \ge F(\rho,\sigma). \tag{164}$$

The next few claims allow us to keep track of the change of the fidelity when we apply operators to its arguments.

Lemma B.6. *For any non-negative operators ρ and σ and any operator W on the same space we have*

$$F(\rho, W\sigma W^\dagger) = F(W^\dagger\rho W, \sigma). \tag{165}$$

Proof. Let $W_D = W$ and let $|\psi\rangle\langle\psi|_{DR}$ and $|\phi\rangle\langle\phi|_{DR}$ be purifications of $\rho_D = \rho$ and $\sigma_D = \sigma$, respectively. Then, by Uhlmann's theorem (Lemma B.3),

$$F(\rho_D, W_D\sigma_D W_D^\dagger) = \sup_{U_R}\big|\langle\psi|(W_D \otimes U_R)|\phi\rangle\big| = \sup_{U_R}\big|\langle\phi|(W_D^\dagger \otimes U_R^\dagger)|\psi\rangle\big|$$
$$= F(\sigma_D, W_D^\dagger\rho_D W_D), \tag{166}$$

where the maximisation is taken over the set of unitaries U_R on R. $\square$

Lemma B.7. *Let ρ and σ be non-negative operators and let $\{W_d\}_{d \in D}$ be a family of operators such that $\sum_{d \in D} W_d = \text{id}$. Then*

$$\sum_{d \in D} F(W_d^\dagger \rho W_d, \sigma) \geq F(\rho, \sigma). \tag{167}$$

Proof. Let $|\psi\rangle\langle\psi|_{DR}$ and $|\phi\rangle\langle\phi|_{DR}$ be purifications of $\rho_D = \rho$ and $\sigma_D = \sigma$, respectively. By Uhlmann's theorem (Lemma B.3), there exists a unitary U_R on R such that

$$F(\rho, \sigma) = \left| \langle\psi|(\text{id}_D \otimes U_R)|\phi\rangle \right| = \left| \sum_{d \in D} \langle\psi|(W_d \otimes U_R)|\phi\rangle \right| \leq \sum_{d \in D} \left| \langle\psi|(W_d \otimes U_R)|\phi\rangle \right|. \tag{168}$$

The assertion follows because, again by Uhlmann's theorem,

$$\left| \langle\psi|(W_d \otimes U_R)|\phi\rangle \right| \leq F(W_d^\dagger \rho W_d, \sigma) \tag{169}$$

holds for any $d \in D$. $\quad\square$

Lemma B.8. *Let ρ_{DE} and σ_{DE} be non-negative operators on $D \otimes E$ and let $\mathcal{W}_E$ be a trace non-increasing completely positive map on E. Then*

$$F\big(\rho_{DE}, (\mathcal{I}_D \otimes \mathcal{W}_E)(\sigma_{DE})\big) \leq F(\rho_D, \sigma_D). \tag{170}$$

Proof. Let $X_E \mapsto \sum_e W_e X W_e^\dagger$ be an operator-sum representation of $\mathcal{W}_E$. The second argument of the fidelity on the left hand side of (170), $\sigma'_{DE} = (\mathcal{I}_D \otimes \mathcal{W}_E)(\sigma_{DE})$, may then be written as

$$\sigma'_{DE} = \sum_e (\text{id}_D \otimes W_e) \sigma_{DE} (\text{id}_D \otimes W_e^\dagger). \tag{171}$$

Because, by assumption, $\sum_e W_e^\dagger W_e \leq \text{id}_E$, we have

$$\sigma'_D = \text{tr}_E(\sigma'_{DE}) = \text{tr}_E\Big(\sum_e (\text{id}_D \otimes W_e^\dagger W_e) \sigma_{DE}\Big) \leq \text{tr}_E(\sigma_{DE}) = \sigma_D. \tag{172}$$

Together with the fact that the square root is operator monotone (cf. Theorem V.1.9 of [5]), this implies

$$F(\rho_D, \sigma'_D) = \text{tr}\big(\sqrt{\sqrt{\rho_D}\sigma'_D\sqrt{\rho_D}}\big) \leq \text{tr}\big(\sqrt{\sqrt{\rho_D}\sigma_D\sqrt{\rho_D}}\big) = F(\rho_D, \sigma_D). \tag{173}$$

The claim then follows from Lemma B.4, which asserts that $F(\rho_{DE}, \sigma'_{DE}) \leq F(\rho_D, \sigma'_D)$. $\quad\square$

We also recall that the fidelity is continuous in its arguments with respect to the trace norm.

Lemma B.9. *Let ρ, ρ', and σ be non-negative operators. Then*

$$\left| F(\rho, \sigma) - F(\rho', \sigma) \right|^2 \leq \|\rho - \rho'\|_1 \text{tr}(\sigma). \tag{174}$$

Proof. Let $|\phi\rangle_{DR}$ be a purification of $\sigma_D = \sigma$. Furthermore, let $|\psi\rangle_{DR}$, $|\psi'\rangle_{DR}$ be purifications of $\rho_D = \rho$, $\rho'_D = \rho'$ such that $F(\rho, \rho') = |\langle\psi|\psi'\rangle|$ (cf. Lemma B.3) and assume without loss of generality that $\langle\psi|\psi'\rangle \geq 0$. We have

$$F(\rho, \sigma) - F(\rho', \sigma) = \sup_U |\langle\psi|(\mathrm{id}_D \otimes U_R)|\phi\rangle| - \sup_{U'} |\langle\psi'|(\mathrm{id}_D \otimes U'_R)|\phi\rangle| \quad (175)$$

$$\leq \sup_U |\langle\psi|(\mathrm{id}_D \otimes U_R)|\phi\rangle| - |\langle\psi'|(\mathrm{id}_D \otimes U_R)|\phi\rangle| \quad (176)$$

$$\leq \sup_U |(\langle\psi| - \langle\psi'|)(\mathrm{id}_D \otimes U_R)|\phi\rangle| \quad (177)$$

$$\leq \sup_U \||\psi\rangle - |\psi'\rangle\|_2 \|(\mathrm{id}_D \otimes U_R)|\phi\rangle\|_2 \quad (178)$$

$$= \||\psi\rangle - |\psi'\rangle\|_2 \||\phi\rangle\|_2, \quad (179)$$

where we have used the Cauchy–Schwarz inequality. The claim then follows from $\||\phi\rangle\|_2^2 = \mathrm{tr}(\sigma)$ and

$$\||\psi\rangle - |\psi'\rangle\|_2^2 = \mathrm{tr}(\rho) + \mathrm{tr}(\rho') - 2\langle\psi|\psi'\rangle = \mathrm{tr}(\rho) + \mathrm{tr}(\rho') - 2F(\rho, \rho') \leq \|\rho - \rho'\|_1, \quad (180)$$

where the inequality is known as the Fuchs–van de Graaf inequality [24].[14]   □

Finally, we provide a lemma (Lemma B.11) that simplifies the evaluation of the fidelity between permutation-invariant operators. It may be seen as a generalisation of a known result on symmetric purifications, which we state as Lemma B.10 (see, e.g., Lemma II.5 of [14] for a proof). Specifically, Lemma B.11 may be seen as a combination of this result and Uhlmann's theorem (Lemma B.3). We also note that the lemma may be generalised to other symmetry groups (other than the symmetric group).

Lemma B.10. *For any permutation-invariant operator ρ_{D^n} on $D^{\otimes n}$ and any space R with $\dim(R) \geq \dim(D)$ there exists a permutation-invariant purification $\rho_{D^n R^n}$ on $(D \otimes R)^{\otimes n}$.*

Lemma B.11. *Let ρ_{D^n} and σ_{D^n} be permutation-invariant non-negative operators on $D^{\otimes n}$ and let $\rho_{D^n R^n}$ be a permutation-invariant purification of ρ_{D^n}. Then there exists a permutation-invariant purification $\sigma_{D^n R^n}$ of σ_{D^n} such that*

$$F(\rho_{D^n}, \sigma_{D^n}) = F(\rho_{D^n R^n}, \sigma_{D^n R^n}). \quad (181)$$

Proof. The proof of this lemma essentially follows the lines of the standard proof of Uhlmann's theorem (see, e.g., Chapter 9 of [41]), while keeping track of the permutation invariance the relevant operators.

For the following, we assume without loss of generality that ρ_{D^n} and σ_{D^n} are invertible. (The claim for the cases where this assumption does not hold may be obtained by considering the operators $\rho_{D^n} + \epsilon\,\mathrm{id}_{D^n}$ and $\sigma_{D^n} + \epsilon\,\mathrm{id}_{D^n}$ for $\epsilon > 0$ and then taking the limit $\epsilon \to 0$.)

---

[14] The Fuchs–van de Graaf inequality is usually formulated for normalised ρ and ρ' as $2 - 2F(\rho, \rho') \leq \|\rho - \rho'\|_1$. However, it is obvious from the proof in [24] that the more general inequality (180) holds for any non-negative ρ and ρ'.

Let $|\Psi\rangle_{D^n R^n}$ be a vector in $(D \otimes R)^{\otimes n}$ such that $\rho_{D^n R^n} = |\Psi\rangle\langle\Psi|_{DR}$ and define

$$|\Omega\rangle_{D^n R^n} = (\rho_{D^n}^{-\frac{1}{2}} \otimes \mathrm{id}_{R^n})|\Psi\rangle_{D^n R^n}. \tag{182}$$

Note that $\mathrm{tr}_{R^n}(|\Omega\rangle\langle\Omega|_{D^n R^n}) = \mathrm{id}_{D^n}$. It thus follows from the Schmidt decomposition that $|\Omega\rangle_{D^n R^n}$ has the form

$$|\Omega\rangle_{D^n R^n} = \sum_x |d_x\rangle_{D^n} \otimes |r_x\rangle_{R^n}, \tag{183}$$

where $\{|d_x\rangle_{D^n}\}_x$ and $\{|r_x\rangle_{R^n}\}_x$ are orthonormal bases of $D^{\otimes n}$ and $R^{\otimes n}$, respectively.

Let U_{D^n} be the unitary operator in the left polar decomposition of $\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}}$, i.e.,

$$\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}} = Q_{D^n} U_{D^n}, \tag{184}$$

where

$$Q_{D^n} = \sqrt{(\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}})(\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}})^\dagger} = \sqrt{\sqrt{\rho_{D^n}}\sigma_{D^n}\sqrt{\rho_{D^n}}} \tag{185}$$

is non-negative. We now define the purification $\sigma_{D^n R^n} = |\Phi\rangle\langle\Phi|_{D^n R^n}$ by

$$|\Phi\rangle = (\sqrt{\sigma_{D^n}} U_{D^n}^\dagger \otimes \mathrm{id}_{R^n})|\Omega\rangle. \tag{186}$$

It is readily verified that this is indeed a purification of σ_{D^n}.

The fidelity between the purifications is given by

$$F(\rho_{D^n R^n}, \sigma_{D^n R^n}) = |\langle\Psi|\Phi\rangle| = |\langle\Omega|(\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}} U^\dagger \otimes \mathrm{id}_{R^n})|\Omega\rangle|. \tag{187}$$

Exploiting now the particular form (183) of $|\Omega\rangle$ as well as (184), this can be rewritten as

$$F(\rho_{D^n R^n}, \sigma_{D^n R^n}) = \Big|\sum_x \langle d_x|\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}} U_{D^n}^\dagger|d_x\rangle\Big| = |\mathrm{tr}(\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}} U_{D^n}^\dagger)| = \mathrm{tr}(Q_{D^n}). \tag{188}$$

The claim (181) then follows by inserting the explicit expression for Q_{D^n}, i.e.,

$$\mathrm{tr}(Q_{D^n}) = \|\sqrt{\rho_{D^n}}\sqrt{\sigma_{D^n}}\|_1 = F(\rho_{D^n}, \sigma_{D^n}). \tag{189}$$

To verify that $\sigma_{D^n R^n}$ is permutation-invariant, we first note that for any permutation-invariant Hermitian operator X on an n-fold product space and for any real function f the operator $f(X)$ is also permutation-invariant. (To see this, consider the decomposition $X = \sum_i x_i \Pi_i$, where Π_i are the projectors onto the eigenspaces of X and x_i are the corresponding eigenvalues. Because $[X, \pi] = 0$ for any permutation π, we also have $[\Pi_i, \pi] = 0$ for any i. Using now that $f(X) = \sum_i f(x_i)\Pi_i$, we conclude that $[f(X), \pi] = 0$.) We therefore know, in particular, that

$$[\rho_{D^n}^{\frac{1}{2}}, \pi] = 0 \quad \text{and} \quad [\rho_{D^n}^{-\frac{1}{2}}, \pi] = 0 \quad \text{and} \quad [\sigma_{D^n}^{\frac{1}{2}}, \pi] = 0 \quad \text{and} \quad [\sigma_{D^n}^{-\frac{1}{2}}, \pi] = 0 \tag{190}$$

for any permutation π. Furthermore, it follows from the explicit expression for Q_{D^n} that this operator is also permutation-invariant. Similarly, since $U_{D^n}^\dagger$ can be written as

$$U_{D^n}^\dagger = \sigma_{D^n}^{-\frac{1}{2}} \rho_{D^n}^{-\frac{1}{2}} Q_{D^n}, \tag{191}$$

it is also permutation-invariant. By assumption, we also have $\pi|\Psi\rangle_{D^n R^n} = |\Psi\rangle_{D^n R^n}$. Because $|\Phi\rangle_{D^n R^n}$ is obtained by multiplying permutation-invariant operators to $|\Psi\rangle_{D^n R^n}$, we conclude that $\pi|\Phi\rangle_{D^n R^n} = |\Phi\rangle_{D^n R^n}$, i.e., $\sigma_{D^n R^n}$ is invariant under permutations. $\square$

C. On the Schur–Weyl Duality

The following lemma follows immediately from the considerations in Chapter 6 of [26] (see, in particular, Eq. 6.25).

Lemma C.1. *Let D and E be Hilbert spaces with $\dim(D) = \dim(E) = d$ and let $n \in \mathbb{N}$. Furthermore, let $\Lambda_{n,d}$ be the set of Young diagrams of size n with at most d rows, and, for any $\lambda \in \Lambda_{n,d}$, let U_λ and V_λ be the corresponding irreducible representations of the unitary group $U(d)$ and the symmetric group S_n, respectively, so that, according to the Schur–Weyl duality (see, e.g., Theorem 1.10 of* [13]*)*

$$D^{\otimes n} \cong \bigoplus_{\lambda \in \Lambda_{n,d}} U_{D,\lambda} \otimes V_{D,\lambda} \tag{192}$$

$$E^{\otimes n} \cong \bigoplus_{\lambda \in \Lambda_{n,d}} U_{E,\lambda} \otimes V_{E,\lambda}. \tag{193}$$

Then there exists a family $\{|\psi_\lambda\rangle_{V_{D,\lambda} V_{E,\lambda}}\}_{\lambda \in \Lambda_{n,d}}$ of maximally entangled normalised vectors on $V_{D,\lambda} \otimes V_{E,\lambda}$ such that any vector $|\Omega\rangle \in \mathrm{Sym}^n(D \otimes E)$ in the symmetric subspace of $(D \otimes E)^{\otimes n}$ can be decomposed as

$$|\Omega\rangle = \sum_\lambda |\phi_\lambda\rangle_{U_{D,\lambda} U_{E,\lambda}} \otimes |\psi_\lambda\rangle_{V_{D,\lambda} V_{E,\lambda}}, \tag{194}$$

where $\{|\phi_\lambda\rangle_{U_{D,\lambda} U_{E,\lambda}}\}_{\lambda \in \Lambda_{n,d}}$ is a family of (not necessarily normalised) vectors on $U_{D,\lambda} \otimes U_{E,\lambda}$.

Proof. According to (192) and (193), the space $(D \otimes E)^{\otimes n}$ decomposes as

$$(D \otimes E)^{\otimes n} \cong \Big(\bigoplus_{\lambda \in \Lambda_{n,d}} U_{D,\lambda} \otimes V_{D,\lambda} \Big) \otimes \Big(\bigoplus_{\lambda' \in \Lambda_{n,d}} U_{E,\lambda'} \otimes V_{E,\lambda'} \Big). \tag{195}$$

Any vector $|\Omega\rangle \in (D \otimes E)^{\otimes n}$ can therefore always be written as

$$|\Omega\rangle = \sum_{\lambda,\lambda' \in \Lambda_{n,d}} \sum_i |\phi_{\lambda,\lambda',i}\rangle_{U_{D,\lambda} U_{E,\lambda'}} \otimes |\psi_{\lambda,\lambda',i}\rangle_{V_{D,\lambda} V_{E,\lambda'}}, \tag{196}$$

where, for any $\lambda, \lambda' \in \Lambda_{n,d}$, $\{|\phi_{\lambda,\lambda',i}\rangle_{U_{D,\lambda} U_{E,\lambda'}}\}_i$ and $\{|\psi_{\lambda,\lambda',i}\rangle_{V_{D,\lambda} V_{E,\lambda'}}\}_i$ are families of vectors in $U_{D,\lambda} \otimes U_{E,\lambda'}$ and $V_{D,\lambda} \otimes V_{E,\lambda'}$, respectively.

For any $\lambda \in \Lambda_{n,d}$, let $\{|v_k\rangle_{V_{D,\lambda}}\}_k$ and $\{|\bar{v}_k\rangle_{V_{E,\lambda}}\}_k$ be orthonormal bases of $V_{D,\lambda}$ and $V_{E,\lambda}$, respectively, with respect to which the representations of the symmetric group S_n are given by the same real-valued matrices. (Such bases always exist, see, e.g., [33].) We then define the maximally entangled vector $|\psi_\lambda\rangle_{V_{D,\lambda} V_{E,\lambda}}$ on $V_{D,\lambda} \otimes V_{E,\lambda}$ by

$$|\psi_\lambda\rangle_{V_{D,\lambda} V_{E,\lambda}} = \sqrt{\frac{1}{\dim V_\lambda}} \sum_k |v_k\rangle_{V_{D,\lambda}} \otimes |\bar{v}_k\rangle_{V_{E,\lambda}}. \tag{197}$$

Now, to prove the claim (194) for any permutation-invariant $|\Omega\rangle$, it suffices to show that the vectors on $V_{D,\lambda} \otimes V_{E,\lambda'}$ in (196) satisfy

$$|\psi_{\lambda,\lambda',i}\rangle_{V_{D,\lambda} V_{E\lambda'}} = \begin{cases} |\psi_\lambda\rangle_{V_{D,\lambda} V_{E,\lambda}} & \text{if } \lambda = \lambda' \\ 0 & \text{otherwise} \end{cases} \tag{198}$$

for all $\lambda, \lambda' \in \Lambda_{n,d}$ and for all i.

For any permutation π, let $V_\lambda(\pi)$ be its action on the irreducible space V_λ. Using that, by definition, the matrix elements $\langle v_k | V_{D,\lambda}(\pi) | v_{k'} \rangle = \langle \bar{v}_k | V_{E,\lambda}(\pi) | \bar{v}_{k'} \rangle$ are real-valued, it is easily verified that

$$(V_{D,\lambda}(\pi) \otimes V_{E,\lambda}(\pi)) | \psi_\lambda \rangle_{V_{D,\lambda} V_{E,\lambda}} = | \psi_\lambda \rangle_{V_{D,\lambda} V_{E,\lambda}} \quad (\forall \pi). \tag{199}$$

Furthermore, the Schur–Weyl duality (cf. Theorem 1.10 of [13]) states that π acts on $\bigoplus_\lambda U_\lambda \otimes V_\lambda$ as

$$V(\pi) = \bigoplus_{\lambda \in \Lambda_{n,d}} \mathrm{id}_{U_\lambda} \otimes V_\lambda(\pi), \tag{200}$$

Using this and that the vector $|\Omega\rangle$ is by assumption invariant under the action of π, we find that

$$(V_{D,\lambda}(\pi) \otimes V_{E,\lambda'}(\pi)) | \psi_{\lambda,\lambda',i} \rangle_{V_{D,\lambda} V_{E,\lambda'}} = | \psi_{\lambda,\lambda',i} \rangle_{V_{D,\lambda} V_{E,\lambda'}} \quad (\forall \pi) \tag{201}$$

holds for all $\lambda, \lambda' \in \Lambda_{n,d}$ and for all i for which the corresponding term in the sum (196) is nonzero.

For any such triple (λ, λ', i) let $H_{\lambda,\lambda',i}$ be the homomorphism between the irreducible representations $V_{E,\lambda'}$ and $V_{E,\lambda}$ defined by

$$\langle \alpha | H_{\lambda,\lambda',i} | \beta \rangle = \langle \psi_{\lambda,\lambda',i} | (\mathrm{id}_{D,\lambda} \otimes |\beta\rangle\langle\alpha|) | \psi_\lambda \rangle, \tag{202}$$

for any $|\alpha\rangle \in V_{E,\lambda}, |\beta\rangle \in V_{E,\lambda'}$. Using (201) and (199) we find that, for any permutation π,

$$\begin{aligned}
\langle \alpha | H_{\lambda,\lambda',i} V_{E,\lambda'}(\pi) | \beta \rangle &= \langle \psi_{\lambda,\lambda',i} | \left(\mathrm{id}_{D,\lambda} \otimes V_{E,\lambda'}(\pi) | \beta \rangle \langle \alpha | \right) | \psi_\lambda \rangle \\
&= \langle \psi_{\lambda,\lambda',i} | \left(V_{D,\lambda}(\pi)^\dagger \otimes |\beta\rangle\langle\alpha| \right) | \psi_\lambda \rangle \\
&= \langle \psi_{\lambda,\lambda',i} | \left(\mathrm{id}_{D,\lambda} \otimes |\beta\rangle\langle\alpha| V_{E,\lambda}(\pi) \right) | \psi_\lambda \rangle \\
&= \langle \alpha | V_{E,\lambda}(\pi) H_{\lambda,\lambda',i} | \beta \rangle.
\end{aligned} \tag{203}$$

This implies that $H_{\lambda,\lambda',i} V_{E,\lambda'}(\pi) = V_{E,\lambda}(\pi) H_{\lambda,\lambda',i}$, i.e., $H_{\lambda,\lambda',i}$ commutes with the action of the symmetry group. Hence, by Schur's lemma (see, e.g., Lemma 0.8 of [13]) and the fact that the representations $V_{E,\lambda}$ and $V_{E,\lambda'}$ are inequivalent for $\lambda \neq \lambda'$, we find

$$H_{\lambda,\lambda',i} = c_{\lambda,i} \delta_{\lambda,\lambda'} \mathrm{id}_{V_{E,\lambda}}, \tag{204}$$

for some appropriately chosen coefficients $c_{\lambda,i}$. Using (202) with $|\alpha\rangle = |\bar{v}_k\rangle_{V_{E,\lambda}}$ and $|\beta\rangle = |\bar{v}_{k'}\rangle_{V_{E,\lambda'}}$ we obtain

$$\begin{aligned}
c_{\lambda,i} \delta_{\lambda,\lambda'} \delta_{k,k'} &= \langle \bar{v}_k | H_{\lambda,\lambda',i} | \bar{v}_{k'} \rangle = \langle \psi_{\lambda,\lambda',i} | (\mathrm{id}_{D,\lambda} \otimes |\bar{v}_{k'}\rangle\langle\bar{v}_k|) | \psi_\lambda \rangle \\
&= \frac{1}{\sqrt{\dim(V_\lambda)}} \langle \psi_{\lambda,\lambda',i} | (|v_k\rangle \otimes |\bar{v}_{k'}\rangle).
\end{aligned} \tag{205}$$

Since this holds for any k, k', we conclude that $| \psi_{\lambda,\lambda',i} \rangle_{V_{D,\lambda} V_{E,\lambda'}}$ is proportional to $| \psi_\lambda \rangle_{V_{D,\lambda} V_{E,\lambda}}$ if $\lambda = \lambda'$ and 0 otherwise. Note that, for $\lambda = \lambda'$, the corresponding proportionality constant can without loss of generality be absorbed in $| \phi_{\lambda,\lambda',i} \rangle_{U_{D,\lambda} U_{E,\lambda'}}$ in the sum (196), so that $| \psi_{\lambda,\lambda,i} \rangle_{V_{D,\lambda} V_{E,\lambda}}$ is normalised. Hence, noting that $| \psi_\lambda \rangle_{V_{D,\lambda} V_{E,\lambda}}$ is also normalised, we have established (198). $\quad\square$

D. Strong Faithfulness of Squashed Entanglement

As an example for how our result can be applied, we present here an argument proposed by Li and Winter. The argument is described in detail in [38]. We summarise it here for convenience.

Squashed entanglement is a measure of entanglement defined for any bipartite state ρ_{AC} as

$$E_{\text{sq}}(\rho_{AC}) = \frac{1}{2} \inf_{\rho_{ACE}} I(A : C|E)_\rho, \tag{206}$$

where the infimum ranges over all non-negative extensions ρ_{ACE} of ρ_{AC} [17]. It is known that squashed entanglement is *faithful*, i.e., strictly positive for any entangled state [7,37]. In other words, $E_{\text{sq}}(\rho_{AC}) = 0$ if and only if the state ρ_{AC} is separable. Theorem 5.1 implies a novel quantitative version of this claim. The main idea is to relate $E_{\text{sq}}(\rho_{AC})$ to the distance between ρ_{AC} and the closest state that is k-extendible (see Footnote 6 for a definition).

Theorem D.1 ([38]). *For any density operator ρ_{AC} on $A \otimes C$ and any $k \in \mathbb{N}$ there exists a k-extendible density operator ω_{AC} such that*[15]

$$\Delta(\rho_{AC}, \omega_{AC}) \leq (k-1)\sqrt{\frac{\ln 2}{2} E_{\text{sq}}(\rho_{AC})}. \tag{207}$$

Proof. Let ρ_{ACE} be a non-negative extension of ρ_{AC}. Theorem 5.1 implies that there exists a trace-preserving completely positive reconstruction map $\mathcal{T}_{E \to CE}$ such that

$$\Delta\big(\rho_{ACE}, (\mathcal{I}_A \otimes \mathcal{T}_{E \to CE})(\rho_{AE})\big) \leq \delta = \sqrt{\ln(2)I(A : C|E)_\rho} \tag{208}$$

(see Eq. 6).

For $i \in \mathbb{N}$, define $\rho^i_{AC_1 \cdots C_i E}$ inductively by

$$\rho^{i+1}_{AC_1 \cdots C_{i+1} E} = (\mathcal{I}_{AC_1 \cdots C_i} \otimes \mathcal{T}_{E \to C_{i+1} E})(\rho^i_{AC_1 \cdots C_i E}), \tag{209}$$

and $\rho^1_{AC_1 E} = \rho_{ACE}$. Because the trace distance cannot increase under the action of $\mathcal{T}_{E \to CE}$ (see Eq. 20) we have

$$\Delta(\rho^i_{AC_i E}, \rho^{i+1}_{AC_{i+1} E}) \leq \Delta(\rho^{i-1}_{AE}, \rho^i_{AE}) \leq \Delta(\rho^{i-1}_{AC_{i-1} E}, \rho^i_{AC_i E}) \tag{210}$$

for $i > 1$. Furthermore, from (208) we have

$$\Delta(\rho^1_{AC_1 E}, \rho^2_{AC_2 E}) = \Delta\big(\rho_{ACE}, (\mathcal{I}_A \otimes \mathcal{T}_{E \to CE})(\rho_{AE})\big) \leq \delta. \tag{211}$$

The combination of these inequalities yields

$$\Delta(\rho^i_{AC_i E}, \rho^{i+1}_{AC_{i+1} E}) \leq \delta \tag{212}$$

for any $i \in \mathbb{N}$. We now apply the triangle inequality to conclude that

$$\Delta(\rho_{ACE}, \rho^j_{AC_j E}) = \Delta(\rho^1_{AC_1 E}, \rho^j_{AC_j E}) \leq \sum_{i=1}^{j-1} \Delta(\rho^i_{AC_i E}, \rho^{i+1}_{AC_{i+1} E}) = (j-1)\delta \tag{213}$$

---

[15] We formulate the claim here for the trace distance $\Delta(\cdot, \cdot)$, but note that it also holds for the purified distance defined in [50].

for any $j \in \mathbb{N}$. Furthermore, because in the definition of the density operators $\rho^i_{AC_1 \cdots C_i E}$ (see Eq. 209) the reconstruction map does not act on the systems $C_1, \ldots C_i$, we have $\rho^j_{AC_1 \ldots C_j} = \rho^k_{AC_1 \ldots C_j}$ for any $j \leq k$, and hence

$$\Delta(\rho_{AC}, \rho^k_{AC_j}) = \Delta(\rho_{AC}, \rho^j_{AC_j}) \leq (j-1) \cdot \delta. \tag{214}$$

Define now the density operator

$$\bar{\omega}_{AC_1 \cdots C_k} = \frac{1}{k!} \sum_\pi \rho^k_{AC_{\pi(1)} \cdots C_{\pi(k)}}, \tag{215}$$

where the sum ranges over all permutations of $\{1, \ldots, k\}$. The density operator $\omega_{AC} = \bar{\omega}_{AC_1}$ is then k-extendible by construction. Using the convexity of the trace distance we find

$$\Delta(\rho_{AC}, \omega_{AC}) = \Delta(\rho_{AC}, \bar{\omega}_{AC_1}) \leq \frac{1}{k!} \sum_\pi \Delta(\rho_{AC}, \rho^k_{AC_{\pi(1)}}) = \frac{1}{k} \sum_{j=1}^k \Delta(\rho_{AC}, \rho^k_{AC_j}). \tag{216}$$

Inserting now the bound (214) we conclude that

$$\Delta(\rho_{AC}, \omega_{AC}) \leq \frac{1}{k} \sum_{j=1}^k (j-1) \cdot \delta \leq \frac{k-1}{2} \cdot \delta = \frac{k-1}{2} \sqrt{\ln(2) I(A:C|E)_\rho}. \tag{217}$$

The claim of the theorem follows because the above holds for any non-negative extension ρ_{ACE} of ρ_{AC}. $\quad\square$

We remark that the bound provided by Theorem D.1 does not depend on the dimension of the two subsystems A and C. As mentioned above, this yields a quantitative claim on the faithfulness of squashed entanglement, which we formulate as Corollary D.3 below. Its proof uses the following statement about the distance of k-extendible states from the set of separable states, which we denote by $S_{A:C}$.

Lemma D.2. *For any k-extendible density operator ω_{AC} on $A \otimes C$*

$$\inf_{\sigma_{AC} \in S_{A:C}} \Delta(\omega_{AC}, \sigma_{AC}) \leq 2 \frac{(\dim C)^2}{k}. \tag{218}$$

Proof. By definition, there exists a density operator $\bar{\omega}_{AC_1 \cdots C_k}$ such that $\omega_{AC} = \bar{\omega}_{AC_i}$ for $i = 1, \ldots, k$. Because this condition still holds if the order of the subsystems $C_1, \ldots, C_n$ is permuted, one can assume without loss of generality that $\bar{\omega}_{AC_1 \ldots C_k}$ is invariant under such permutations. The claim then follows immediately from Theorem II.7′ of [14]. $\quad\square$

Corollary D.3 ([38]). *For any density operator ρ_{AC} on $A \otimes C$*

$$\inf_{\sigma_{AC} \in S_{A:C}} \Delta(\rho_{AC}, \sigma_{AC}) \leq 2 \dim C \sqrt[4]{2 \ln(2) E_{\mathrm{sq}}(\rho_{AC})}. \tag{219}$$

Proof. Let ω_{AC} be a k-extendible density operator that satisfies (207). Using the triangle inequality we can combine this with Lemma D.2 to obtain

$$\inf_{\sigma_{AC} \in S_{A:C}} \Delta(\rho_{AC}, \sigma_{AC}) \leq \Delta(\rho_{AC}, \omega_{AC}) + \inf_{\sigma_{AC} \in S_{A:C}} \Delta(\omega_{AC}, \sigma_{AC})$$

$$\leq (k-1)\sqrt{\frac{\ln 2}{2} E_{\mathrm{sq}}(\rho_{AC})} + 2\frac{(\dim C)^2}{k}. \qquad (220)$$

Inserting $k = \left\lceil \sqrt[4]{\frac{8}{\ln(2) E_{\mathrm{sq}}(\rho_{AC})}} \dim C \right\rceil$ then yields the claim. $\quad\square$

Corollary D.3 quantifies the faithfulness of squashed entanglement in terms of the trace norm. Compared to previously known versions of this claim [7], only the dimension of *one* the two subsystems enters as a factor in the bound (219). (Because of the symmetry of the other involved quantities, one can always choose the lower-dimensional one.) We also note that the example of the totally antisymmetric state on $A \otimes C$ with $\dim A = \dim C = d$ shows that such a factor is necessary. Indeed, the squashed entanglement of this state is of the order $O(1/d)$ [16] whereas its trace distance to the closest separable state cannot be smaller than $\frac{1}{4}$. (To see this, note that for any product state $\sigma_A \otimes \sigma_C$ we have $\mathrm{tr}(\sigma_A \otimes \sigma_C \Pi_{\mathrm{as}}) \leq \frac{1}{2}$ where Π_{as} denotes the projector onto the antisymmetric subspace of $A \otimes C$.)

References

1. Alicki, R., Fannes, M.: Continuity of quantum conditional information. J. Phys. A Math. Gen. **37**(5), L55–L57 (2004). arXiv:quant-ph/0312081
2. Beaudry, N.J., Renner, R.: An intuitive proof of the data processing inequality. Quantum Inf. Comput. **12**(5–6), 432–441 (2012). arXiv:1107.0740
3. Berta, M., Christandl, M., Furrer, F., Scholz, V.B., Tomamichel, M.: Continuous variable entropic uncertainty relations in the presence of quantum memory (2013). arXiv:1308.4527
4. Berta, M., Seshadreesan, K., Wilde, M.M.: Renyi generalizations of the conditional quantum mutual information. J. Math. Phys. **56**, 022205 (2015). arXiv:1403.6102
5. Bhatia, R.: Matrix Analysis. Springer, Berlin (1997)
6. Brandão, F.G.S.L., Harrow, A.W., Oppenheim, J., Strelchuk, S.: Quantum conditional mutual information, reconstructed states, and state redistribution. Phys. Rev. Lett. **115**, 050501 (2015). arXiv:1411.4921
7. Brandão, F.G.S.L., Christandl, M., Yard, J.: Faithful squashed entanglement. Commun. Math. Phys. **306**(3), 805–830 (2011). arXiv:1010.1750
8. Brandão, F.G.S.L., Harrow, A.W.: Product-state approximations to quantum ground states. In: Proceedings of ACM STOC, pp. 871–880. ACM, New York (2013). arXiv:1310.0017
9. Brandão, F.G.S.L., Harrow, A.W.: Quantum de Finetti theorems under local measurements with applications. In: Proceedings of ACM STOC, pp. 861–870. ACM, New York (2013). arXiv:1210.6367
10. Braverman, M.: Interactive information complexity. In: Proceedings of ACM STOC, pp. 505–524. ACM, New York (2012). ECCC:TR11-123

11. Carlen, E., Lieb, E.: Remainder terms for some quantum entropy inequalities. J. Math. Phys. **55**(4), 042201 (2014). arXiv:1402.3840
12. Choi, M.-D.: Completely positive linear maps on complex matrices. Linear Algebra Appl. **10**(3), 285–290 (1975)
13. Christandl, M.: The structure of bipartite quantum states—insights from group theory and cryptography. PhD thesis, University of Cambridge (2006). arXiv:quant-ph/0604183
14. Christandl, M., König, R., Mitchison, G., Renner, R.: One-and-a-half quantum de Finetti theorems. Commun. Math. Phys. **273**(2), 473–498 (2007). arXiv:quant-ph/0602130
15. Christandl, M., König, R., Renner, R.: Postselection technique for quantum channels with applications to quantum cryptography. Phys. Rev. Lett. **102**(2), 020504 (2009). arXiv:0809.3019
16. Christandl, M., Schuch, N., Winter, A.: Entanglement of the antisymmetric state. Commun. Math. Phys. **311**(2), 397–422 (2012). arXiv:0910.4151
17. Christandl, M., Winter, A.: "Squashed entanglement": an additive entanglement measure. J. Math. Phys. **45**(3), 829–840 (2004)
18. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (2005)
19. Datta, N.: Min-and max-relative entropies and a new entanglement monotone. IEEE Trans. Inform. Theory **55**(6), 2816–2826 (2009). arXiv:0803.2770
20. Devetak, I., Yard, J.: Exact cost of redistributing multipartite quantum states. Phys. Rev. Lett. **100**, 230501 (2008). arXiv:quant-ph/0612050
21. Dupuis, F., Krämer, L., Faist, P., Renes, J.M., Renner, R.: Generalized entropies. In: Proceedings of the XVIIth International Congress on Mathematical Physics, pp. 134–153 (2013). arXiv:1211.3141
22. Ekert, A., Renner, R.: The ultimate physical limits of privacy. Nature **507**, 443–447 (2014)
23. Erker, P.: How not to Rényi generalize the quantum conditional mutual information. J. Phys. A Math. Theor. **48**, 275303 (2015). arXiv:1404.3628
24. Fuchs, C.A., van de Graaf, J.: Cryptographic distinguishability measures for quantum-mechanical states. IEEE Trans. Inform. Theory **45**(4), 1216–1227 (1999). arXiv:quant-ph/9712042
25. Furrer, F., Åberg, J., Renner, R.: Min-and max-entropy in infinite dimensions. Commun. Math. Phys. **306**(1), 165–186 (2011). arXiv:1004.1386
26. Harrow, A.W.: Applications of coherent classical communication and the Schur transform to quantum information theory. PhD thesis, Massachusetts Institute of Technology (2005). arXiv:quant-ph/0512255
27. Hayashi, M.: Quantum Information: An Introduction. Springer, Berlin (2006)
28. Hayden, P., Jozsa, R., Petz, D., Winter, A.: Structure of states which satisfy strong subadditivity of quantum entropy with equality. Commun. Math. Phys. **246**(2), 359–374 (2004). arXiv:quant-ph/0304007
29. Hiai, F., Petz, D.: The proper formula for relative entropy and its asymptotics in quantum probability. Commun. Math. Phys. **143**(1), 99–114 (1991)
30. Ibinson, B., Linden, N., Winter, A.: Robustness of quantum Markov chains. Commun. Math. Phys. **277**(2), 289–304 (2008). arXiv:quant-ph/0611057
31. Jain, R., Nayak, A.: The space complexity of recognizing well-parenthesized expressions in the streaming model: the index function revisited. IEEE Trans. Inf. Theory **60**(10) (2014). arXiv:1004.3165
32. Jain, R., Radhakrishnan, J., Sen, P.: A lower bound for the bounded round quantum communication complexity of set disjointness. In: Proceedings of FOCS, pp. 220–229. IEEE (2003). arXiv:quant-ph/0303138
33. James, G., Kerber, A.: The Representation Theory of the Symmetric Group. Addison-Wesley, Reading (1981)
34. Jamiołkowski, A.: Linear transformations which preserve trace and positive semidefiniteness of operators. Rep. Math. Phys. **3**(4), 275–278 (1972)
35. Kerenidis, I., Laplante, S., Lerays, V., Roland, J., Xiao, D.: Lower bounds on information complexity via zero-communication protocols and applications. In: Proceedings of FOCS, pp. 500–509. IEEE (2012). arXiv:1204.1505
36. Kim, I.: Conditional independence in quantum many-body systems. PhD thesis, California Institute of Technology (2013). http://thesis.library.caltech.edu/7697/
37. Li, K., Winter, A.: Relative entropy and squashed entanglement. Commun. Math. Phys. **326**(1), 63–80 (2014). arXiv:1210.3181
38. Li, K., Winter, A.: Squashed entanglement, k-extendibility, quantum Markov chains, and recovery maps (2014). arXiv:1410.4184
39. Lieb, E.H., Ruskai, M.B.: Proof of the strong subadditivity of quantum-mechanical entropy. J. Math. Phys. **14**(12), 1938–1941 (1973)
40. Müller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S., Tomamichel, M.: On quantum Rényi entropies: a new generalization and some properties. J. Math. Phys. **54**(12), 122203 (2013). arXiv:1306.3142
41. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, London (2000)
42. Ogawa, T., Nagaoka, H.: Strong converse and Stein's lemma in quantum hypothesis testing. IEEE Trans. Inf. Theory **46**(7), 2428–2433 (2000). arXiv:quant-ph/9906090

43. Petz, D.: Sufficiency of channels over von Neumann algebras. Q. J. Math. **39**(1), 97–108 (1988)
44. Renner, R.: Security of quantum key distribution. PhD thesis, ETH Zurich (2005). arXiv:quant-ph/0512258
45. Renner, R.: Symmetry of large physical systems implies independence of subsystems. Nat. Phys. **3**, 645–649 (2007). arXiv:quant-ph/0703069
46. Renner, R.: Simplifying information-theoretic arguments by post-selection. Inf. Commun. Secur. **26**, 66–75 (2010)
47. Renner, R., Maurer, U.: About the mutual (conditional) information. In: Proceedings of IEEE ISIT (2002)
48. Seshadreesan, K.P., Wilde, M.M.: Fidelity of recovery and geometric squashed entanglement (2014). arXiv:1410.1441
49. Tomamichel, M.: A framework for non-asymptotic quantum information theory. PhD thesis, ETH Zurich (2012). arXiv:1203.2142
50. Tomamichel, M., Colbeck, R., Renner, R.: A fully quantum asymptotic equipartition property. IEEE Trans. Inform. Theory **55**, 5840–5847 (2009). arXiv:0811.1221
51. Touchette, D.: Quantum information complexity and amortized communication (2014). arXiv:1404.3733
52. Uhlmann, A.: The "transition probability" in the state space of a *-algebra. Rep. Math. Phys. **9**(2), 273–279 (1976)
53. Wang, L., Renner, R.: One-shot classical-quantum capacity and hypothesis testing. Phys. Rev. Lett. **108**, 200501 (2012). arXiv:1007.5456
54. Wilde, M.M.: Quantum Information Theory. Cambridge University Press, London (2014). arXiv:1106.1445
55. Wilde, M.M., Winter, A., Yang, D.: Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. Commun. Math. Phys. **331**(2), 593–622 (2014). arXiv:1306.1586
56. Zhang, L.: Conditional mutual information and commutator. Int. J. Theor. Phys. **52**(6), 2112–2117 (2013). arXiv:1212.5023
57. Zhang, L., Wu, J.: A lower bound of quantum conditional mutual information. J. Phys. A **47**, 415303 (2014). arXiv:1403.1424

Communicated by M. M. Wolf