Squashed entanglement, k-extendibility, quantum Markov chains, and recovery maps

Ke Li\textsuperscript{1,2} and Andreas Winter\textsuperscript{3}

\textsuperscript{1}IBM T J Watson Research Center, Yorktown Heights, NY 10598, USA
\textsuperscript{2}Center for Theoretic Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
\textsuperscript{3}ICREA \& Física Teòrica: Informació i Fènoms Quàntics, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain

(Dated: 6 November 2014)

Squashed entanglement \cite{Christandl/Winter, J. Math. Phys. 45(3):829-840 (2004)} is a monogamous entanglement measure, which implies that highly extendible states have small value of the squashed entanglement. Here, invoking a recent inequality for the quantum conditional mutual information \cite[Fawzi/Renner, arXiv:1410.0664]{Fawzi/Renner}, we show the converse, that a small value of squashed entanglement implies that the state is close to a highly extendible state. As a corollary, we establish an alternative proof of the faithfulness of squashed entanglement \cite[Brando\textquotesingle/Christandl/Yard, Commun. Math. Phys. 306:805-830 (2011)]{Brando\textquotesingle/Christandl/Yard}.

We briefly discuss the pre-history of the Fawzi-Renner bound and related conjectures, and close by advertising a potentially far-reaching generalization to the monotonicity of the relative entropy.

**Squashed entanglement.**—One of the core goals in the theory of entanglement is its quantification, for which purpose a large number of either operationally or mathematically/axiomatically motivated entanglement measures and monotonies have been introduced and studied intensely since the 1990s \cite{Horodecki/etal}. In this paper we will discuss one specific such measure, the so-called squashed entanglement \cite{Christandl/Winter}, defined as

\[
\text{E}_\text{sq}(\rho_{AB}^F) := \inf \frac{1}{2} I(A : B | E) \text{ s.t. } \text{Tr}_E \rho_{ABE} = \rho_{AB},
\]

where \(I(A : B | E) = S(AE) + S(BE) - S(E) - S(ABE)\) is the (quantum) conditional mutual information, which by strong subadditivity of the von Neumann entropy is always non-negative \cite{Nielsen/Chuang}, and \(\rho_{ABE}\) as above is called an extension of \(\rho_{AB}\). This definition appears to have been put forward first in \cite{Brandao/Christandl/Yard}, where it was also remarked that by restricting the extension of \(\rho_{AB}\) to have the form \(\rho_{ABE} = \sum_i p_i |\varphi_i\rangle \langle \varphi_i|_{AB} \otimes |i\rangle \langle i|\), the minimization reduces to the well-known entanglement of formation \cite{Wootters/Winter}.

\[
\text{E}_F(\rho_{AB}) = \min \sum_i p_i S(\varphi_i^A) \text{ s.t. } \sum_i p_i |\varphi_i\rangle \langle \varphi_i| = \rho. \tag{2}
\]

While it is fairly straightforward to see from their definitions that both \(\text{E}_\text{sq}\) and \(\text{E}_F\) are convex functions of the state, the former has many properties that the latter lacks, among them additivity and monogamy \cite{Christandl/Winter, Brandao/Christandl/Yard, Horodecki/etal}, cf. \cite{Horodecki/etal}. Abbreviating \(\text{E}_\text{sq}(\rho_{AB}) = \text{E}_\text{sq}(A : B)\),

\[
\text{E}_\text{sq}(A : B_1B_2) \geq \text{E}_\text{sq}(A : B_1) + \text{E}_\text{sq}(A : B_2). \tag{3}
\]

In particular, if \(\rho_{AB}^k\) is \(k\)-extendible, meaning that there exists a state \(\rho_{AB_1\ldots B_k}\) such that \(\rho_{iAB}^k = \rho_{AB_i}\) for all \(i\) (and that w.l.o.g. is symmetric with respect to permutations of the \(B\)-systems), then

\[
\text{E}_\text{sq}(A : B) \leq \frac{1}{k} \log |A|. \tag{4}
\]

While clearly \(E_{\text{sq}} \leq E_F\), in the other direction, squashed entanglement is an upper bound on the distillable entanglement and indeed on the distillable secret key in a state \cite{Christandl/Winter, Brandao/Christandl/Yard}, which makes it very useful to the theory of state distillation and channel capacities, cf. \cite{Winter/Consani}. One of the properties much desirable for a quantitative entanglement measure is faithfulness, i.e. the fact that it is zero if and only if the state is separable, and otherwise strictly positive. To be truly useful, such a statement ought to come in the form of a relationship between the value of the entanglement measure, and a suitably chosen distance from the set of separable states. Such a statement was finally obtained a couple of years ago by Brandão et al. \cite{Brandao/Christandl/Yard}, and later improved by us \cite{Li/Winter2013}. In the present paper, we will reproduce this finding in a conceptually simple and appealing way, by first showing a relation between the value of squashed entanglement and the distance from \(k\)-extendible states, and then invoking a suitable de Finetti theorem to bound the distance from separable states. (That in the limit of \(k \to \infty\) the state has to be separable was known for some time \cite{Dupuis/etal}, but we shall use more recent, quantitative, versions.)

After that, we make a comparison with the faithfulness of entanglement of formation. Then, we put the technical result of Fawzi and Renner \cite[Thm. 5.1]{Fawzi/Renner}, on which our proof crucially relies, in the context of other conjectured inequalities; motivated by a much more general observation in classical probability, we propose as an open problem to find the right quantum generalization.

**Main result.**—Now we show that the monogamy bound \cite{Christandl/Winter} has a partial converse:

**Theorem 1** Consider a state \(\rho_{AB}\) with \(\text{E}_\text{sq}(\rho) \leq \epsilon\). Then, for every integer \(k\), there exists a \(k\)-extendible state \(\sigma^{AB}\) such that \(||\rho - \sigma|| \leq (k-1)\sqrt{2 \ln 2} \sqrt{\epsilon}\). In particular, \(\rho\) is \(O(\sqrt{\epsilon})\)-close to a \(\Omega\left(\frac{1}{\sqrt{\epsilon}}\right)\)-extendible state.
Corollary 2  For every state \( \rho^{AB} \) with \( E_{sq}(\rho) \leq \epsilon \), there exists a separable state \( \sigma \) with
\[
\|\rho - \sigma\|_1 \leq 3.1 |B| \sqrt{\epsilon}.
\]
In particular, squashed entanglement is faithful: \( E_{sq}(\rho) = 0 \) if and only if the state \( \rho \) is separable.

For comparison, the earlier result of Brandão et al. \cite{Brandao}, Cor. 1] yields
\[
\|\rho - \sigma\|_1 \leq \sqrt{\det(A)} \|\rho - \sigma\|_2 \leq 12 \sqrt{\det(A)} \|B\| \sqrt{\epsilon}.
\] (5)
The Hilbert-Schmidt (2-)norm bound seems not available with our techniques, but the trace (1-)norm behaviour is qualitatively reproduced here, albeit with a worse polynomial dependence on \( \epsilon \) but with a slightly better constant. In particular, it is of interest if we can extend our bound in Corollary \cite{Brandao} only the dimensionality of one of the two systems appears (cf. however \cite{Brandao} Eq. (66)).

The proof of this theorem relies essentially on a very recent result by Fawzi and Renner \cite{Fawzi}, stating that for every tripartite state \( \rho^{AEB} \) there exists a cptp map \( \tilde{R} : \mathcal{L}(E) \rightarrow \mathcal{L}(EB) \) such that
\[
-\log F(\rho^{AEB}, (\text{id}_{A} \otimes \tilde{R})\rho^{AEB})^2 \leq I(A : B | E)_\rho.
\] (6)
with the fidelity \( F \) of two states \( \alpha \) and \( \beta \) defined as \( F(\alpha, \beta) = \sqrt{\text{tr}(\sqrt{\alpha \beta})} \).

Proof  Choose an extension \( \rho^{AEB} \) for \( \rho^{AB} \), and use the map \( \tilde{R} \) from Eq. (6). Now we employ a basic inequality from \cite{Brandao} Thm. 1, saying
\[
1 - F(\alpha, \beta) \leq 1/2 \|\alpha - \beta\|_1 \leq \sqrt{1 - F(\alpha, \beta)^2},
\] (7)
hence, from Eq. (6),
\[
t := \sqrt{4 \ln 2 \cdot I(A : B | E)} \geq \|\rho^{AEB} - (\text{id}_A \otimes \tilde{R})\rho^{AE}\|_1.
\]

But since \( (\text{id}_A \otimes \tilde{R})\rho^{AE} \approx \rho^{AEB} \), we may apply the same map again, say \( k - 1 \) times, always to the \( E \) system of \( \rho^{AEB} \), arriving at a state
\[
\omega^{AB_1 \ldots B_k} = (\text{id}_A \otimes \tilde{R}^{E \rightarrow EB_1} \circ \ldots \circ \tilde{R}^{E \rightarrow EB_k})\rho^{AEB_1 \ldots B_k},
\]
which has the property that for each \( i, \|\omega^{AB_i} - \rho^{AB_i}\|_1 \leq (i - 1)t \), by the triangle inequality and the contravariant property of the trace norm under cptp maps. Hence, tracing out \( E \) and considering the symmetrization of the \( B \) systems, i.e.
\[
\Omega^{AB_1 \ldots B_k} = \frac{1}{k!} \sum_{\pi \in S_k} (\mathbb{1} \otimes U^\pi) \omega^{AB_1 \ldots B_k} (\mathbb{1} \otimes U^\pi)^\dagger,
\]
we have that it is manifestly permutation symmetric on the \( B \) systems, and for all \( i, \)
\[
\|\Omega^{AB_i} - \rho^{AB_i}\|_1 \leq \frac{k - 1}{2} t.
\] (8)
Minimizing over all extensions as required by the definition of squashed entanglement, allowing \( I(A : B | E) \) to get arbitrarily close to \( 2\epsilon \), concludes the proof of the theorem.

To show the corollary, we use \cite[Thm. 2 & Cor. 5]{Brandao} or alternatively \cite[Thm. II.7]{Brandao}, which say that a \( k \)-extendible state is at trace distance at most \( \frac{2|B|^2}{k} \) from a separable state. To use the former result, which requires Bose-symmetric extensions, we have to go from the permutation symmetric \( \Omega^{AB_1 \ldots B_k} \) to a permutation invariant purification
\[
|\Psi\rangle^{AA'B_1 B_1' \ldots B_k B_k'} = \left( \sqrt{\Omega^{AB_1 \ldots B_k} \otimes \mathbb{1}} \right) |\Phi\rangle^{AA'} |\Phi\rangle^{B_1 B_1'} \ldots |\Phi\rangle^{B_k B_k'},
\]
with the non-normalized maximally entangled state \( |\Phi\rangle = \sum_i |i\rangle |i\rangle \). The choice
\[
k = \left[ \frac{\sqrt{2} \ |B|}{\ln 2 \ \sqrt{\epsilon}} \right],
\]
then does the rest. \( \square \)

Comparison with entanglement of formation.—It is instructive to compare the monogamy relation Eq. (4) and its “converse”, Theorem 1 for the squashed entanglement, with the analogous statements for the entanglement of formation:

Proposition 3  In a bipartite system \( AB \), if the state \( \rho^{AB} \) is \( \delta \)-close in trace norm to a separable state \( \sigma^{AB} \), with \( \delta \leq \frac{1}{\sqrt{2}} \), then
\[
E_F(\rho) \leq 5 \log(\|A\| B) \sqrt{\delta} + \sqrt{\delta} \log \delta.
\] (9)
Conversely, if \( E_F(\rho) \leq \epsilon \), then this implies that there is a separable state \( \sigma \) such that \( \|\rho - \sigma\|_1 \leq \sqrt{\epsilon} \).

Proof  The first part is due to Nielsen \cite{Nielsen}. For the second part, consider an optimal decomposition \( \rho = \sum_i p_i |\varphi_i\rangle \langle \varphi_i| \), such that
\[
\epsilon \geq \sum_i p_i \frac{1}{2} I(A : B | \varphi_i) \geq \sum_i p_i \frac{1}{4 \ln 2} \|\varphi_i^{AB} - \varphi_i^{A} \otimes \varphi_i^{B}\|_1^2
\]
\[
\geq \frac{1}{4 \ln 2} \|\rho - \sum_i p_i |\varphi_i\rangle \langle \varphi_i| \otimes |\varphi_i\rangle \langle \varphi_i| \|_1^2,
\]
and the right hand state inside the trace norm is manifestly separable. \( \square \)

In other words, while entanglement of formation is essentially about the distance from separable states, squashed entanglement is about the distance from highly extendible states (up to log-dimensionality factors and
polynomial relation of \( \epsilon \) and \( \delta \). Note that squashed entanglement, like the entanglement of formation, is asymptotically continuous \( \text{[13]} \). Alcstri and Fannes \( \text{[2]} \) showed that for \( \| \rho^{AB} - \sigma^{AB} \|_1 \leq \epsilon \leq 1 \),

\[
|E_{\text{sq}}(\rho) - E_{\text{sq}}(\sigma)| \leq 8\epsilon \log |A| + 4H_2(\epsilon),
\]
where \( H_2(x) = -x \log x - (1-x) \log(1-x) \) is the binary entropy.

This explains the occurrence of states such as the \( d \times d \) fully antisymmetric state \( \alpha \), which is at trace distance 1 from the separable states for all \( d \), but has \( E_{\text{sq}}(\alpha) \leq \frac{d}{2} \) very small \( \text{[14]} \). Indeed, this state is \((d - 1)\)-extendible, so by monogamy of \( E_{\text{sq}} \) it has to have small squashed entanglement. And by Theorem \( \text{[4]} \) this is the only way in which a state can have small squashed entanglement. On the other hand, the large distance from separable, and the dimension-dependent constants in Corollary \( \text{[2]} \) and Eq. \( \text{(5)} \), are entirely due to the fact that in large dimension, highly extendible states can be far away from being separable.

**Recovery maps and related facts & conjectures.**—
The form \( \text{[6]} \) of the Fawzi-Renner bound \( \text{[12]} \) was arrived at in a succession of speculative steps. The initial insight is no doubt Petz’s \( \text{[24]} \), who showed a general statement on the relative entropy

\[
D(\rho||\sigma) = \text{Tr} \rho (\log \rho - \log \sigma).
\]

Indeed, while for any two states \( \rho \) and \( \sigma \) on a system \( H \) and a ctp map \( T : \mathcal{L}(H) \rightarrow \mathcal{L}(K) \),

\[
D(\rho||\sigma) = \|T(\rho) - T(\sigma)\|_1 \geq \|D(T||T)\|_1
\]

– this is equivalent to strong subadditivity \( \text{[21]} \) –, Petz showed that equality holds if and only if there exists a ctp map \( R \) such that \( RT = \sigma \) and \( RT\rho = \rho \). What is more, this map can be constructed in a unified way from \( T \) and \( \sigma \) alone, as the transpose channel, or Petz recovery map \( R = R(T, \sigma) \), given by

\[
R(\xi) = \sqrt{\sigma} T^* (T\xi)^{1/2} \xi (T\sigma)^{-1/2} \sqrt{\sigma},
\]
where \( T^* \) is the adjoint map to \( T \), at least in the finite dimensional case (cf. \( \text{[3]} \)).

The above problem involving the conditional mutual information is recovered by letting \( T = \text{Tr}_B \), \( \rho = \rho^{AEB} \) and \( \sigma = \rho^A \otimes \rho^{EB} \), where it can be checked that

\[
I(A : B|E) = I(A : EB) - I(A : E) = D(\rho^{AEB}||\rho^A \otimes \rho^{EB}) - D(\rho^{AE}||\rho^A \otimes \rho^E).
\]

In this case, the Petz recovery map reads

\[
R(\xi) = \sqrt{\rho^{EB}} \left( \sqrt{\rho^{E-1}} \xi \sqrt{\rho^{E-1}} \otimes \mathbb{I}^B \right) \sqrt{\rho^{EB}},
\]
and the recovered state from \( \rho^{AE} \) is

\[
\omega^{AE} = (\text{id}^A \otimes R^{E\rightarrow B}) \rho^{AE} \\
= \sqrt{\rho^{EB}} \left( \sqrt{\rho^{E-1}} \rho^{AE} \sqrt{\rho^{E-1}} \otimes \mathbb{I}^B \right) \sqrt{\rho^{EB}}.
\]

This map was used to elucidate the structure of \( \rho^{AEB} \text{[14]} \). The result is that there has to exist a decomposition \( E = \bigoplus_j \epsilon^j_j \otimes \rho^j_B \) of \( E \) as a direct sum of tensor products, such that

\[
\rho^{AEB} = \bigoplus_j p_j \sigma_j^A \otimes \tau_j^B.
\]
(In particular, \( \rho^{AB} \) is separable.) Such states were called “quantum Markov chains” \( \text{[1]} \).

The recovery map of Fawzi and Renner \( \text{[12]} \) looks very similar to the form \( \text{(11)} \):

\[
\tilde{R}(\xi) = \sqrt{\rho^{EB}} \left( \sqrt{\rho^{E-1}} U \xi U^T \sqrt{\rho^{E-1}} \otimes \mathbb{I}^B \right) \sqrt{\rho^{EB}} V^T,
\]
with certain unitaries \( U \) (on \( E \)) and \( V \) (on \( EB \)).

The near-equality case of Petz’s theorem seems to have attracted little attention until recently, for instance as shown here in the context of squashed entanglement, or in the approach of Brandão and Harrow to finite quantum de Finetti theorems \( \text{[7]} \), or potentially in considerations of many-body physics \( \text{[18]} \).

The conjecture that the Petz recovery map \( R \) in Eq. \( \text{(11)} \) might yield \( \omega^{AB} \approx \rho^{AB} \) in trace norm was first formulated by Kim \( \text{[17]} \):

\[
I(A : B|E) \geq \Omega \left( \|\rho^{AEB} - (\text{id} \otimes R)\rho^{AE}\|_1^2 \right).
\]
See also Zhang \( \text{[30]} \) for this, who suggested the generalized version

\[
D(\rho||\sigma) - D(T\rho||T\sigma) \geq \Omega \left( \|\rho - RT\rho\|_1^2 \right).
\]

Berta et al. \( \text{[3]} \) then proposed the more natural conjecture with \( -\log F(\rho^{AEB}, R_\rho^{AE})^2 \) at the right hand side of \( \text{(13)} \), motivated by the observation that the latter is a Rényi conditional mutual information:

\[
I(A : B|E) \geq -\log F(\rho^{AEB}, (\text{id} \otimes R)\rho^{AE})^2.
\]

By the well-known relations connecting fidelity and trace norm, this would imply Kim’s conjecture \( \text{[13]} \). While all of the above conjectures remain open, Fawzi-Renner’s \( \text{[6]} \) proves a variant of the last inequality, with \( \tilde{R} \) instead of \( R \). The crucial point of course is that this new map still only acts on \( E \), and as the identity on \( A \). Similarly, Seddarean et al. \( \text{[25]} \), Conj. 26 & Sect. 6.1] suggested the following most general form extending \( \text{[14]} \), encompassing all of the above:

\[
D(\rho||\sigma) - D(T\rho||T\sigma) \geq -\log F(\rho, RT\rho)^2,
\]
again motivated by a way of writing both sides of the above as Rényi relative entropies or variants thereof.

**The classical case.**—It is well-known that for classical random variables, \( XYZ \), the conditional mutual independence of information, \( I(X : Z|Y) = 0 \), implies that
Theorem 4 If \( I(X : Z | Y) = \epsilon \) for a distribution \( P(XYZ) \), then there exists a Markov chain of the same alphabet, with distribution \( Q(XYZ) = P(XY)P(Z|Y) \), such that the relative entropy distance between \( P \) and \( Q \) is small: \( D(P_{XYZ} || Q) = \epsilon \). By Pinsker’s inequality, this implies \( \| P_{XYZ} - Q \|_1 \leq \sqrt{2\ln 2} \epsilon \).

This is a special case of the following more general

Theorem 5 For any two probability distributions \( P \) and \( Q \) on the same set \( X \), and a stochastic map \( T : X \rightarrow \mathcal{U} \), there exists another stochastic map \( R \), called the transpose channel, and which depends only on \( Q \) and \( T \), such that \( RTQ = Q \) and

\[
D(P || Q) - D(TP || TQ) \geq D(P || RTP).
\]

Furthermore, this is an identity if \( T \) is deterministic.

The transpose channel is defined by the property that \( T(u|z)Q(x) = TQ(z|u)(Q(u|x)) \), and this is the classical case of Petz’s recovery map.

We relegate the proof, which is essentially an application of log-concavity, to Appendix A.

Observe that the inequality (17) implies the conjectures [13], [14], [15] and [16] in the classical case, because of \( D(P || Q) \geq -\log F(P, Q)^2 \).

It is known, by numerical counterexamples, that (17) is false in the quantum case, already in qubits, and also restricting to the case \( T = \text{Tr}_B \), \( \rho = \rho^{AEB} \) and \( \sigma = \rho^A \otimes \rho^{EB} \). However, it is possible that with a variant of the Fawzi-Renner map, say \( \hat{R} \), we might have

\[
I(A : B|E) \geq D(\rho^{AEB} (\text{id} \otimes \hat{R})\rho^{AE}) \geq D(\rho || \hat{R}T\rho),
\]

which would also imply (14).

Discussion.—We have shown how Fawzi and Renner’s recent breakthrough in the characterization of small quantum conditional mutual information has consequences for the faithfulness of squeezed entanglement. We believe that the same approach can be used also to address the faithfulness of the multi-party squeezed entanglement [29], however technical issues remain, which are explained in Appendix B.

The result of [12] also finally clarifies the “right” robust version of quantum Markov chains, which are equivalently given by \( I(A : B|E) \approx 0 \) and the existence of a recovery map such that \( \rho^{AEB} \approx (\text{id}_A \otimes \hat{R})\rho^{AE} \), cf. [3, Prop. 35]. For classical probability distributions, yet another way of expressing this is to say that there exists a Markov chain close to the given density, but this is not the case in the quantum analogue [10, 16], at least if one does not want to introduce strong dimensional dependence.

To conclude, looking back at the conjectures reviewed above and contrasting it with the clear picture emerging from the classical case, we wish to suggest a target for further investigation, which takes us in a direction different from the conjecture [19] and its descendants.

Namely, the question is, whether it is possible to define a recovery map \( \hat{R} = \hat{R}(T, \sigma) \) for every pair of a cptp map \( T \) and a state \( \sigma \) in its domain, such that \( \hat{R}T\sigma = \sigma \) and

\[
D(\rho || \sigma) - D(T\rho || T\sigma) \geq D(\rho || \hat{R}T\rho),
\]

and such that the following functoriality properties hold:

- **Normalization**: To the identity map \( \text{id} \) and any state (of full rank), the identity map is associated: \( \hat{R}(\text{id}, \tau) = \text{id} \).

- **Tensor**: If \( \hat{R}_i = \hat{R}(T_i, \sigma_i) \) is associated to maps \( T_i \) and states \( \sigma_i \), then the map associated to \( T_1 \otimes T_2 \) and state \( \sigma_1 \otimes \sigma_2 \) is \( \hat{R}(T_1 \otimes T_2, \sigma_1 \otimes \sigma_2) = \hat{R}_1 \otimes \hat{R}_2 \).

This would clearly imply the inequality [18]. Note that the Petz map quite evidently obeys the functoriality properties, in fact in addition also another one:

- **Composition**: For cptp maps \( T \) on suitable space, such that we can form their composition \( T_2 \circ T_1 \), and a state \( \sigma \) such that we have associated maps \( \hat{R}_1 = \hat{R}(T_1, \sigma) \) and \( \hat{R}_2 = \hat{R}(T_2, T_1\sigma) \), we have \( \hat{R}(T_2 \circ T_1, \sigma) = \hat{R}_1 \circ \hat{R}_2 \).

Can all these constraints be satisfied simultaneously? And if so, what would be the applications of such a result? Note that the Petz recovery map is a very useful tool in “pretty good” state discrimination and quantum error correction [3]; the functoriality above along with [18] is meant to preserve these good properties.

Acknowledgments.—Since the first formulation of the proof idea of Theorem 1 in 2008, we have enjoyed conversations and the keen interest of many people, among them Fernando Brandão, Matthias Christandl, Paul Erker, Omar Fawzi, Aram Harrow, Isaac Kim, Renato Renner and Mark Wilde.

When this work was started, KL was affiliated with the Centre for Quantum Technologies (CQT), National University of Singapore; AW was affiliated with the School of Mathematics, University of Bristol and had a part-time appointment with CQT. KL is supported by NSF Grant CCF-1110941 and CCF-1111382. AW is supported by the ERC Advanced Grant “IRQUAT”, the European Commission (STREP “RAQUEL”), and the Spanish MINECO (project FIS2008-01236) with the support of FEDER funds.
[1] L. Accardi, A. Frigerio, “Markovian cocycles”, Proc. Roy. Irish Acad. 83A(2):251–263 (1983).
[2] R. Alicki, M. Fannes, “Continuity of quantum conditional information”, Phys. Lett. A 37:155-157 (2004).
[3] H. Barnum, E. Knill, “Reversing quantum dynamics with near-optimal quantum and classical fidelity”, J. Math. Phys. 43(5):2097-2106 (2002).
[4] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, W. K. Wootters, “Mixed-state entanglement and quantum error correction”, Phys. Rev. A 54(5):3824-3851 (1996).
[5] M. Berta, K. P. Seshadreesan, M. M. Wilde, “Rényi generalizations of the conditional quantum mutual information”, arXiv:quant-ph/1403.6102 (2014).
[6] F. G. S. L. Brandão, M. Christandl, J. T. Yard, “Faithful Squashed Entanglement”, Commun. Math. Phys. 306:805-830 (2011).
[7] F. G. S. L. Brandão, A. W. Harrow, “Quantum de Finetti Theorems under Local Measurements with Applications”, in: Proc. 45th ACM Symp. Theory Comput. (STOC 2013), pp. 861-870. arXiv:quant-ph/1210.6367 (2012).
[8] M. Christandl, The Structure of Bipartite Quantum States: Insights from Group Theory and Cryptography, PhD thesis, University of Cambridge (February 2006).
[9] M. Christandl, R. König, G. Mitchison, R. Renner, “One-and-a-Half Quantum de Finetti Theorems”, Commun. Math. Phys. 273:473-498 (2007).
[10] M. Christandl, N. Schuch, A. Winter, “Highly Entangled States with Almost No Secrecy”, Phys. Rev. Lett. 104:240405 (2010). “Entanglement of the Antisymmetric State”, Commun. Math. Phys. 311(2):397-422 (2012).
[11] M. Christandl, A. Winter, “Squashed entanglement: An additive entanglement measure”, J. Math. Phys. 45(3):829-840 (2004).
[12] O. Fawzi, R. Renner “Quantum conditional mutual information and approximate Markov chains”, arXiv:quant-ph/1410.0664 (2014).
[13] C. A. Fuchs, J. van de Graaf, “Cryptographic Distiguishability Measures for Quantum-Mechanical States”, IEEE Trans. Inf. Theory 45(4):1216-1227 (1999).
[14] P. Hayden, R. Jozsa, D. Petz, A. Winter, “Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality”, Commun. Math. Phys. 246(2):359-374 (2004).
[15] R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, “Quantum entanglement”, Rev. Mod. Phys. 81(2):865-942 (2009).
[16] B. Ibinson, N. Linden, A. Winter, “Robustness of Quantum Markov Chains”, Commun. Math. Phys. 277(2):289-304 (2008).
[17] I. H. Kim, personal communication, July 2012.
[18] I. H. Kim, “On the informational completeness of local observables”, arXiv:quant-ph/1405.0137 (2014).
[19] M. Koashi, A. Winter, “Monogamy of quantum entanglement and other correlations”, Phys. Rev. A 69:022309 (2004).
[20] K. Li, A. Winter, “Relative Entropy and Squashed Entanglement”, Commun. Math. Phys. 326(1):63-80 (2014).
[21] E. H. Lieb, M.-B. Ruskai, “Proof of the strong subadditivity of quantum-mechanical entropy. With an appendix by B. Simon”, J. Math. Phys. 14(12):1938-1941 (1973).
[22] M. Navascués, M. Owari, M. B. Plenio, “Power of symmetric extensions for entanglement detection”, Phys. Rev. A 80:052306 (2009).
[23] M. A. Nielsen, “Continuity bounds for entanglement”, Phys. Rev. A 61:064301 (2000).
[24] D. Petz, “Sufficient subalgebras and the relative entropy of states of a von Neumann algebra”, Commun. Math. Phys. 105(1):123-131 (1986). “Sufficiency of channels over von Neumann algebras”, Quart. J. Math. Oxford, Ser. 2 39(1):97-108 (1988).
[25] K. P. Seshadreesan, M. Berta, M. M. Wilde, “Rényi squashed entanglement, discord, and relative entropy differences”, arXiv:quant-ph/1410.1443 (2014).
[26] E. Stormer, “Symmetric states of infinite tensor products of C*-algebras”, J. Func. Anal. 7(1):48-68 (1969). G. A. Raggio, R. F. Werner, “Quantum statistical mechanics of general mean field systems”, Helv. Phys. Acta 62:980-990 (1989). M. Fannes, J. T. Lewis, A. Verbeure, “Symmetric states of composite systems”, Lett. Math. Phys. 15:255-260 (1988).
[27] M. Takeoka, S. Guha, M. M. Wilde, “The squashed entanglement of a quantum channel”, IEEE Trans. Inf. Theory 60(8):4987-4998 (2014).
[28] R. R. Tucci, “Quantum entanglement and conditional information transmission”, arXiv:quant-ph/9909041 (1999). “Enganglement of distillation and conditional mutual information”, arXiv:quant-ph/0202144 (2002).
[29] D. Yang, K. Horodecki, M. Horodecki, J. Oppenheim, W. Song, “Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof”, IEEE Trans. Inf. Theory 55(7):3375-3387 (2009).
[30] L. Zhang, “Conditional Mutual Information and Commutator”, Int. J. Theor. Phys. 52(6):2112-2117 (2013).
APPENDIX

A — Proof of Theorem 5. We have two probability vectors \( P = (p_x)_{x=1}^{|X|} \) and \( Q = (q_x)_{x=1}^{|X|} \), and a stochastic matrix \( T = [t_{ux}]_{u=1}^{|U|}_{x=1}^{|X|} \) (meaning that for all \( x, \sum_{u=1}^{|U|} t_{ux} = 1 \)). The adjoint of ctp map translates into the linear map given by the transpose matrix \( T^t \).

Then,

\[
TP = \left( \sum_{x} t_{ux} p_x \right)_{u=1}^{|U|}, \quad TQ = \left( \sum_{x} t_{ux} q_x \right)_{u=1}^{|U|},
\]

and

\[
RTP = \left( q_x \left( \frac{\sum_{y} t_{uy} p_y}{\sum_{y} t_{uy} q_y} \right) \right)_{x=1}^{|X|} = \left( \sum_{u} t_{ux} \sum_{x'} t_{ux'} p_{x'} \sum_{x'} t_{ux'} q_{x'} \right)_{x=1}^{|X|},
\]

leading to the following expressions for the three relative entropies concerned:

\[
D(P\|Q) = \sum_{x} p_x \log \frac{p_x}{q_x},
\]

\[
D(TP\|TQ) = \sum_{u} \left( \sum_{x} t_{ux} p_x \right) \log \frac{\sum_{x'} t_{ux'} p_{x'}}{\sum_{x'} t_{ux'} q_{x'}}.
\]

The claimed inequality, that the first expression is larger or equal to the sum of the last two, can be rearranged as \( D(P\|Q) - D(P\|RT P) \geq D(TP\|TQ) \), which simplifies to

\[
\sum_{x} p_x \log \left( \sum_{u} t_{ux} \sum_{x'} t_{ux'} p_{x'} \sum_{x'} t_{ux'} q_{x'} \right) \geq \sum_{x} p_x \sum_{u} t_{ux} \log \sum_{x'} t_{ux'} p_{x'} \sum_{x'} t_{ux'} q_{x'}.
\]

However, this is true for any term \( x \), due to the concavity of log and \( \sum_{u} t_{ux} = 1 \).

It can be checked from this that if the channel \( T \) is deterministic, i.e. if for each \( x \in X \) there is only one \( u \in U \) such that \( t_{ux} > 0 \), then equality holds; in particular this is the case where \( T \) is the marginal map from \( \mathcal{X} \times \mathcal{Y} \) to \( \mathcal{X} \).

\[ \square \]

B — Multi-party squashed entanglement. It seems that our approach could also be used to prove faithfulness of the multi-party squashed entanglement \([\text{22}]\),

\[
E_{sq}(\rho^{A_1 \cdots A_n}) = \inf_{\rho^{A_1 \cdots A_n} E} \frac{1}{2} I(A_1 : \cdots : A_n|E),
\]

with \( I(A_1 : \cdots : A_n|E) = \sum_{i=1}^n S(A_i|E) - S(A_1 \cdots A_n|E) \) the conditional multi-information. That is, \( E_{sq}(\rho^{A_1 \cdots A_n}) \) would vanish iff \( \rho \) is n-separable:

\[
\rho^{A_1 \cdots A_n} = \sum_{\lambda} p_{\lambda}^{A_1} \otimes \cdots \otimes p_{\lambda}^{A_n}.
\]

It seems that with the methods of \([\text{20}]\) this cannot be approached.

The idea starts from the identity

\[
I(A_1 : \cdots : A_n|E) = I(A_1 : A_2 \cdots A_n|E) + I(A_2 : \cdots : A_n|E)
\]

\[
= \cdots = \sum_{i=1}^{n-1} I(A_i : A_{i+1} \cdots A_n|E),
\]

showing that \( I(A_1 : \cdots : A_n|E) \leq 2\epsilon \) implies, for all \( i, I(A_i : A_{[n]\setminus i}|E) \leq 2\epsilon \), and more generally \( I(A_i : A_{[n]\setminus i}|E) \leq 2\epsilon \) for all subsets \( I \subset [n] \).

In particular, if \( \epsilon = 0 \), we can use the structure theorem of \([\text{14}]\) to find, for each \( i \), a projective measurement \( \{P^{(i)}_j\} \) on \( E \) that commutes with \( \rho^{A_1 \cdots A_n} E \), such that for all \( \lambda_i \),

\[
\text{Tr}_E \rho^{A_1 \cdots A_n} E_{A_i} = p_{\lambda_i} \sigma^{A_i} \otimes \tau^{A_{[n]\setminus i}},
\]

i.e., conditioned on the measurement outcomes \( \lambda_i, A_i \) and \( A_{[n]\setminus i} \) are in a product state. Performing all these measurements in some fixed order, we thus obtain outcomes \( \lambda = \lambda_1 \cdots \lambda_n \) such that conditioned on \( \lambda \), the state is a product state with respect to all partitions \( i : [n]\setminus i \), which means that conditioned on \( \lambda, A_1, \ldots, A_n \) factorize.

We would like to use the machinery of the recovery maps to extract from \( E \) a large number of \( k \) of approximate copies of each \( A_i \), using approximate recovery maps \( \tilde{R}_i : \mathcal{L}(E) \to \mathcal{L}(E A_i) \) according to Eq. (6). With \( t = \sqrt{8 \ln 2 \sqrt{\epsilon}} \) and tracing out \( E \), we can indeed get a state \( \Omega^{A_1 A_2^{[k]} \cdots A_n^{[k]}} \), with \( A_i^{[k]} = A_i^1 \cdots A_i^k \) consisting of \( k \) copies of \( A_i \), such that

\[
\|\rho^{A_1 \cdots A_n} - \Omega^{A_1 A_2^{2} \cdots A_n^{n}}\|_1 \leq (n-1)(k-1)t \leq nk \sqrt{8 \ln 2 \sqrt{\epsilon}},
\]

for all tuples \( (j_2, \ldots, j_n) \) such that all but at most one \( j_i \) equals 1.

We cannot say easily that this holds for all tuples \((j_2, \ldots, j_n)\), because the different recovery maps may interfere with each other. However, if we could conclude that, we would be done: by symmetrizing the \( k \) copies of each \( A_i \) \((i > 1)\) we would find, as before, that \( \rho \) is \( O(\sqrt{\epsilon}) \)-close to a \( k \)-extendible state, with \( k = \Omega \left( \frac{1}{\sqrt{\epsilon}} \right) \).

We could then again use the results of \([\text{22}]\), now extended to the multi-partite case, to see that \( \Omega^{A_1 A_2^{2} \cdots A_n^{n}} \) is at trace distance at most \( \frac{1}{k} (|A_2|^2 + \cdots + |A_n|^2) \) from a fully separable (i.e. n-separable) state.

We have to leave this problem to the attention of the interested reader.