Limitations on Quantum Key Repeaters

Stefan Bäuml,[1, 2, *] Matthias Christandl,[3, †] Karol Horodecki,[4, 5, ‡] and Andreas Winter[6, 2, 1, §]

[1] *Department of Mathematics, University of Bristol, Bristol BS8 1TW, UK*

[2] *Física Teòrica: Informació i Fenomens Quàntics,*
Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain

[3] *Institute for Theoretical Physics, ETH Zürich,*
Wolfgang-Pauli-Str. 27, 8093 Zürich, Switzerland

[4] *Institute of Informatics, University of Gdańsk, 80-952 Gdańsk, Poland*

[5] *National Quantum Information Centre of Gdańsk, 81-824 Sopot, Poland*

[6] *ICREA - Institució Catalana de Recerca i Estudis Avançats, ES-08010 Barcelona, Spain*

(Dated: 24 February 2014)

A main application of quantum communication is the distribution of entangled particles for use in quantum key distribution (QKD). Due to unavoidable noise in the communication line, QKD is in practice limited to a distance of a few hundred kilometers and can only be extended to longer distances by use of a future quantum repeater, a small-scale quantum computer which performs iterated entanglement distillation and quantum teleportation. The existence of entangled particles that are undistillable but nevertheless useful for QKD raises the question for a *quantum key repeater* which works beyond the limits of entanglement distillation. In this work we show that any such apparatus is severely limited in its performance; in particular, we exhibit entanglement suitable for QKD but unsuitable for the most general quantum key repeater protocol. The mathematical techniques we develop can be viewed as a step towards opening the theory of entanglement measures to networks of communicating parties.

I. SUMMARY

When a signal is passed from a sender to a receiver, it inevitably degrades due to the noise present in any realistic communication channel (e.g. a cable or free space). The degradation of the signal is typically exponential in the length of the communication line. When the signal is classical, degradation can be counteracted by use of an amplifier that measures the degraded signal and,

[*] Electronic address: `stefan.baeuml@bristol.ac.uk`

[†] Electronic address: `christandl@phys.ethz.ch`

[‡] Electronic address: `khorodec@inf.ug.edu.pl`

[§] Electronic address: `andreas.winter@uab.cat`

depending on a threshold, replaces it by a stronger signal. When the signal is quantum mechanical (e.g. encoded in non-orthogonal polarisations of a single photon), such an amplifier cannot work anymore, since the measurement inevitably disturbs the signal [1]. Sending a quantum mechanical signal, however, is the basis of quantum key distribution (QKD), a method to distribute a key which can later be used for perfectly secure communication between sender and receiver [2]. The degradation of sent quantum signals therefore seems to place a fundamental limit on the distance at which secure communication is possible thereby severely limiting its applicability in the internet [3, 4].

A way around this limitation is the use of an entanglement-based quantum key distribution scheme [5, 6] in conjunction with a so-called quantum repeater [7]. Here, many entangled pairs of particles are being distributed between the sender (Alice) and an intermediate station (Charlie), and between Charlie and the receiver (Bob) (see Fig. 1). Charlie, who plays the role of an untrusted telecom provider, for instance, prepares n Einstein-Podolsky-Rosen (EPR) entangled pairs of photons and sends one photon of each pair to Alice. In the same fashion he distributes n pairs between himself and Bob. Noise, of course, will degrade the quality of the EPR pairs during the transmission process. If the distances between Charlie and Alice/Bob are small enough, however, the noisy pairs remain distillable, this means that they can be transformed into $\approx E_D \times n$ perfect EPR pairs, where E_D is known as the distillable entanglement of the quantum state of the noisy pair. The EPR pairs between Charlie and Bob are then used to quantum teleport the state of Charlie's other particles to Bob. This process, known as entanglement swapping, results in EPR pairs between Alice and Bob [8]. When Alice and Bob make appropriate measurements on the EPR pairs they obtain a sequence of secret key bits, that is, an identical but random sequence of bits that is uncorrelated with the rest of the universe (including Charlie's systems). This secure key can later be used for perfectly secure communication. The described scheme with one intermediate station effectively doubles the distance over which QKD can be carried out and more repeater stations can be inserted to efficiently extend the distance arbitrarily. The implementation of quantum repeaters is therefore one of the focal points of the experimental quantum information science community [9, 10].

As this explanation illustrates, and as it was realised early on in quantum cryptography, entanglement distillation and QKD are tightly connected [11, 12]; and indeed it is clear that entanglement is a necessary prerequisite for privacy [13]. It therefore came as a surprise to many researchers in 2005 that there are undistillable entangled states (so-called bound entangled states that have $E_D = 0$) from which secret key can be obtained [14]. With the help of a quantum repeater as

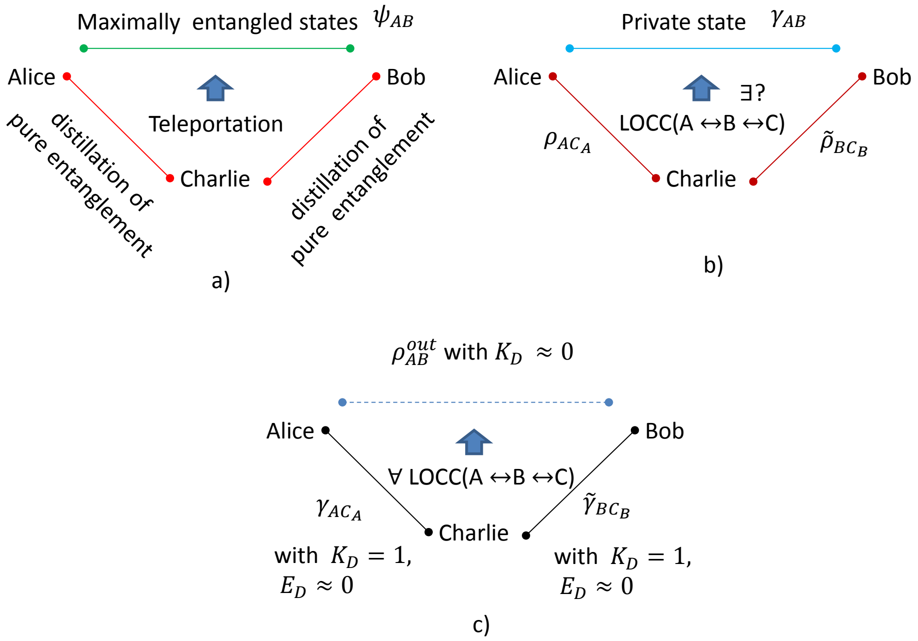
Maximally entangled states ψ_{AB}
Alice
Bob
Teleportation
Charlie
distillation of pure entanglement
distillation of pure entanglement
a)
Private state γ_{AB}
Alice
Bob
∃?
LOCC(A ↔B ↔C)
ρ_{AC_A}
$\tilde{\rho}_{BC_B}$
Charlie
b)
ρ_{AB}^{out} with $K_D \approx 0$
Alice
Bob
∀ LOCC(A ↔B ↔C)
γ_{AC_A}
$\tilde{\gamma}_{BC_B}$
Charlie
with $K_D = 1$, $E_D \approx 0$
with $K_D = 1$, $E_D \approx 0$
c)

FIG. 1: a) Quantum repeater distributing maximally entangled states ψ_{AB}. b) Quantum key repeater distributing general private states (not necessarily maximally entangled ones). c) States containing privacy (e.g. p-bit) which cannot be successfully used in a quantum key repeater.

described above, however, the secret key contained in such states cannot be extended to larger distances, as the states do not allow for the distillation of EPR pairs.

The present paper raises the question of whether there may be other ways to extend the secret key to arbitrary distances than by distillation and swapping of entanglement, other *quantum key repeaters*. More formally, we analyse the quantum key repeater rate $K_{A\leftrightarrow C\leftrightarrow B}$ at which a protocol is able to extract private bits between Alice and Bob from entangled states which they shared with Charlie by local operations and classical communication (LOCC) among the three of them. Note that just as the definition of the distillable key [14, 15], the definition of this rate is information-theoretic in nature. By using quantum tomography, the post-selection technique [16], error correction and privacy amplification [17], this rate can be made robust in a cryptographic sense, therefore leading to unconditional security of the obtained secret key. For unconditional security in relation to pbits see also [18, 19]. By a private bit we mean an entangled state representing a unit of privacy paralleling the EPR pair as a unit of entanglement [14, 20]. Mathematically, private bits are entangled states of the form

$$\rho_{AA'BB'} = \frac{1}{2}\begin{bmatrix} \sqrt{XX^\dagger} & 0 & 0 & X \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ X^\dagger & 0 & 0 & \sqrt{X^\dagger X} \end{bmatrix}, \tag{1}$$

where A and B are qubits that contain the key bits, corresponding to the rows and columns in the matrix. A' and B' are each a d-dimensional systems, called the shield. X is a d^2-by-d^2 matrix with $\|X\|_1 = 1$. In the following we will describe our main results which demonstrate that the performance of quantum key repeaters beyond the use of entanglement distillation is severely limited.

Our first result takes its starting point in the observation that there are private bits that are almost indistinguishable from separable states by local operations and classical communication. An example is the choice $X = \frac{1}{d\sqrt{d}} \sum_{ij} u_{ij} |i\rangle\langle j| \otimes |j\rangle\langle i|$, where the u_{ij} are the entries in the quantum Fourier transform in dimension d. This can be easily seen as the LOCC distinguishability of two states is upper bounded by the distinguishability under operations that preserve the positivity of the partial transpose, and the latter is bounded by $\|X^\Gamma\|_1 = \frac{1}{\sqrt{d}}$, where Γ indicates the transpose of one of the systems [21]. Imagine that such private bits are the entangled states that are distributed between Alice and Charlie, and between Bob and Charlie, and that a quantum repeater protocol using local operations and classical communication between Alice, Bob and Charlie, transforms them successfully into a private bit between Alice and Bob. Then, by Alice and Bob joining their labs, they can distinguish this resulting state from a separable state using a measurement (this is done by untwisting the shield A'B' to obtain an EPR pair which can be distinguished, for instance by a Bell measurement, from separable states [14]). This would imply an LOCC procedure for Alice-Bob (when they join their labs) and Charlie to distinguish the initial private bits $\rho \otimes \rho$ from separable states: first run the quantum key repeater protocol and then perform the measurement. This, however, is in contradiction to the assumption that the private state ρ (and hence $\rho \otimes \rho$) is almost indistinguishable from separable states under LOCC. In conclusion this shows that such private bits cannot be successfully extended to a private bit between Alice and Bob by any quantum key repeater protocol. A direct mathematical formulation of this explanation is given in Section III B, but applies only to protocols acting on single copies of the states $\rho \otimes \rho$ and therefore does not give a bound on $K_{A\leftrightarrow C\leftrightarrow B}$, which allows joint operations on an arbitrary number of copies.

The language of entanglement measures allows us to formulate this argument asymptotically as a rigorous distinguishability bound on $K_{A\leftrightarrow C\leftrightarrow B}$ (see Section III C):

$$K_{A\leftrightarrow C\leftrightarrow B}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}) \leq D^{\infty}_{C\leftrightarrow AB}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}), \tag{2}$$

where the RHS is the regularised LOCC-restricted relative entropy distance to the closest separable state [22]. Arguably, it is difficult if not impossible to compute this expression (there is a regularisation, a maximisation over LOCC measurements and a minimisation over separable states). But noting that this bound is invariant under partial transposition of the C system, we can easily upper bound the quantity for all known bound entangled states (these are the ones with positive partial transpose) in terms of the relative entropy of entanglement of the partially transposed state ρ^{Γ}: $E_R(\rho^{\Gamma})+E_R(\tilde{\rho}^{\Gamma})$; if we restrict to forward communication from Charlie and $\rho_{AC_A}=\tilde{\rho}_{BC_B}$, some more effort shows that squashed entanglement provides a bound: $K_{A\leftarrow C\rightarrow B}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}) \leq 4E_{sq}(\rho^{\Gamma})$. As we show, D^{∞} is a robust quantity in that it does not decrease by too much when a qubit is lost (it is not lockable). This fact can be used to extend our results to states that are not exactly but only close to states having a positive partial transpose. We also use this fact to improve the squashed entanglement bound to the *reduced* squashed entanglement [23]. Both the relative entropy bound and the reduced squashed entanglement bound can be regularised.

Extracting the algebraic content of the idea of the partial transposition of the C system we are bound to loose the intuition behind our results, but are able to circumvent the quantity D and directly obtain for PPT states ρ and $\tilde{\rho}$ (see Section III A):

$$K_{A\leftrightarrow C\leftrightarrow B}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}) \leq K_D(\rho^{\Gamma}), \tag{3}$$

where K_D is the key rate, the rate at which two parties can extract key from ρ. This bound (and its similar version where we swap the roles of A and B) leads to improved relative entropy and squashed entanglement bounds, as these entanglement measures are the well-known upper bounds on the key rate K_D [14, 24]. We leave open the question of whether the reduced squashed entanglement bound can be obtained and improved in the same way.

We will now give an example of a state $\rho_{AC_A}=\tilde{\rho}_{BC_B}$ for which the key rate is large, but the bounds, and hence the quantum key repeater rate are arbitrarily small. Guided by our intuition, we would like to consider the private bit from above whose partial transpose is close to a separable state. The only caveat here is that the state is not PPT (no private bit can be PPT [14]). Fortunately, our state turns into a PPT state ρ under adding a little bit of noise. Since its partial transpose ρ^{Γ} is almost separable, the key rate, the relative entropy of entanglement and the squashed entanglement are close to zero and we find $K_{A\leftrightarrow C\leftrightarrow B}(\rho\otimes\rho)\approx 0$. Since a small amount of noise can easily be removed [15], the state continues to have almost one bit of key: $K_D(\rho)\approx 1$. This leads us to the main conclusion of our paper: there exist entangled quantum states that are useful for quantum key distribution at small distances but that are virtually useless for long-distance quantum key distribution.

There is another type of bound, based on the direct analysis of the entanglement of a concrete output state of a quantum repeater protocol (see Section III D). More precisely, by considering the state that Alice and Bob have conditioned on Charlie's measurement, we find

$$K_{A\leftarrow C\leftrightarrow B}(\rho\otimes\tilde{\rho})\leq\frac{1}{2}E_C(\rho)+\frac{1}{2}E_D(\tilde{\rho}). \tag{4}$$

In contrast to the bounds presented above, which mainly apply to PPT states and are able to deliver maximal limitations, this bound appears weaker but applies to all quantum states ρ and $\tilde{\rho}$. The bound is maximally strong for ρ an EPR pair, since then $K_{A\leftarrow C\leftrightarrow B}$ is bounded by $\frac{1}{2}$ regardless of how much key the bound entangled state $\tilde{\rho}$ contains. The multiplicative constants in (4) are tight which can be seen by inserting for both ρ and $\tilde{\rho}$ an EPR pair. We also apply this bound to states ρ and $\tilde{\rho}$ that are locally equivalent to their partial transposition, thereby giving non-trivial limitations in the regime where the bounds based on the partial transposition fail to deliver non-trivial results. Note that in the case of PPT states, one may also partially transpose the states appearing on the right hand side since $K_{A\leftarrow C\leftrightarrow B}$ is invariant under partial transposition.

The proof of this result is obtained by upper bounding the squashed entanglement of the output state of the protocol using a manipulation of entropies resulting in the RHS of (4). The squashed entanglement in turn upper bounds the distillable key of the output state (which upper bounds the LHS) [24]. This raises the question, of whether there are other inequalities relating the output state of such protocols by entanglement measures of the input states. In the context of algebro-geometric measures of entanglement, this question has been raised and relations among the concurrence of input and output states have been found [25–28]. In our context, we have focused on operational entanglement measures and we may ask in particular, whether our result can be made stronger by replacing the LHS by the entanglement cost of the output state. Based on a random construction we show that this is not true, therefore giving a further indication of the tightness of our result. When restricting the attention to PPT states, one may ask whether Alice and Bob's post-measurement state conditioned on any measurement by Charlie is always separable [29–31]. If this was true, the quantum key repeater rate would vanish for all PPT states. The upper bounds presented in this work may therefore be seen as information-theoretic evidence for the truth of this PPT^2

conjecture. Reaching even further, and consistent with our findings, we may speculate that perhaps the only “transitive” entanglement in quantum states, i.e. entanglement suitable for repeaters, is their distillable entanglement.

With this paper we initiate a study of long-distance quantum communication and cryptography beyond the use of entanglement distillation. Even though the reported results provide limitations rather than new possibilities, we hope that this work will lead to a rethinking of the currently used protocols resulting in procedures for long-distance quantum communication that are both more efficient and that can operate in noisier environments. More abstractly, our results can be viewed as a step towards an entanglement theory for networks of communicating parties with inequalities relating initial and final entanglements.

Contents

II. PRELIMINARIES

In this section we first formally recall the definition of a private bit, of the secret key rate and of the distillable entanglement. We will then introduce the distillation of secure key with an intermediate station and formally introduce the corresponding information theoretic rate of secure key.

A private state can be constructed from a maximally entangled state $|\Psi_{AB}\rangle$ by tensoring some state $\sigma_{A'B'}$ and performing a so-called "twisting" operation. A twisting operation is a controlled unitary of the form $U^{\text{twist}} = \sum_{ij} |ij\rangle\langle ij|_{AB} \otimes U^{(ij)}_{A'B'}$ that spreads the entanglement over the enlarged Hilbert space. Formally

$$\gamma_m = U^{\text{twist}} \left(|\Psi^{(2^m)}\rangle\langle\Psi^{(2^m)}|_{AB} \otimes \sigma_{A'B'} \right) U^{\text{twist}\dagger} \tag{5}$$

$$= \frac{1}{2^m} \sum_{ij=0}^{2^m-1} |ii\rangle\langle jj|_{AB} \otimes U^{(ii)} \sigma_{A'B'} U^{(jj)\dagger}. \tag{6}$$

It has been shown that even if Eve is in possession of the entire purification of γ_m, Alice and Bob will still be able to obtain m bits of perfect key by measuring the AB subsystem in the computational basis, while keeping the $A'B'$ part away from Eve. As all the correlation the key has with the outside world is contained in $A'B'$, it is called the "shield part", whereas AB is called the "key part". For $m = 1$, γ_1 is also called a "private bit" or "p-bit" which can alternatively be represented in the form of (1). As the twisting operations can be non-local, not every private state can be obtained from a single rank 2^m maximally entangled state via LOCC. This shows that privacy is a truly different property of a quantum state than its distillable entanglement, motivating the introduction of a quantity known as "distillable key" [14]

$$K_D(\rho) = \inf_{\epsilon>0} \limsup_{n\to\infty} \sup_{\Lambda_n \text{ LOCC},\gamma_m} \left\{ \frac{m}{n} : \Lambda_n(\rho^{\otimes n}) \approx_\epsilon \gamma_m \leq \epsilon \right\}, \tag{7}$$

in analogy to the distillable entanglement

$$E_D(\rho) = \inf_{\epsilon>0} \limsup_{n\to\infty} \sup_{\Lambda_n \text{ LOCC}} \left\{ \frac{\log d}{n} : \Lambda_n(\rho^{\otimes n}) \approx_\epsilon |\Psi^{(d)}\rangle\langle\Psi^{(d)}| \leq \epsilon \right\}. \tag{8}$$

With $\alpha \approx_\epsilon \beta$ we mean $\|\alpha - \beta\|_1 \leq \epsilon$. Clearly $K_D(\gamma_m) \geq m$. As every rank 2^m-dimensional maximally entangled state is a private state, $K_D \geq E_D$.

Surprisingly there exist bound entangled states that are arbitrarily close to private states in trace distance [14]. A natural question arising now is how such nearly bound entangled private states can be distributed between distant parties. Of course it would be possible to distribute maximal entanglement using a conventional repeater and then distill the state needed. This would, however, have no advantage over using the maximal entanglement directly for QKD. Here, we deal with the question whether there are other, not maximally entangled, possibly even bound entangled states that could be initially distributed between the nodes and then swapped yielding a state useful for cryptography. In order to study this question, we introduce the following quantity. For input states ρ_{AC_A} and $\tilde{\rho}_{C_BB}$ we call

$$K_{A\leftrightarrow C\leftrightarrow B}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}) = \inf_{\epsilon>0}\limsup_{n\to\infty}\sup_{\Lambda_n \mathrm{LOCC},\gamma_m}\left\{\frac{m}{n} : \mathrm{Tr}_C\Lambda_n\left((\rho_{AC_A}\otimes\tilde{\rho}_{C_BB})^{\otimes n}\right)\approx_\epsilon \gamma_{\lfloor m\rfloor}\right\} \quad (9)$$

the *quantum key repeater rate of ρ and $\tilde{\rho}$ with respect to arbitrary LOCC operations among Alice, Bob and Charlie.* If we restrict the protocols to one-way communication from Charlie to Alice we write $K_{A\leftarrow C\leftrightarrow B}$ and if all communication is one-way from Charlie we write $K_{A\leftarrow C\rightarrow B}$. It is the goal of this paper to find significant upper bounds on this quantity.

III. BOUNDS ON THE QUANTUM KEY REPEATER RATE

This section is structured into four parts. First, we will explain the partial transpose idea which is mathematically straightforward and delivers strong bounds on the key rate for PPT states. Then, we will explain the distinguishability idea (for single copy and multiple copy repeaters), which is more intuitive but also technically more involved. Finally, we present the entanglement measures idea, which analyses the output state of a protocol without reference to partial transposition. All sections contain examples illustrating and comparing the different bounds.

A. Partial Transpose Idea

Let us assume that Alice shares a PPT state ρ with Charlie and that Bob shares a PPT state $\tilde{\rho}$ with Charlie and that they apply an LOCC operation Λ among the three of them at the end of which Charlie traces out his part of the system. It is the observation of this section that they obtain the identical output state had they applied the LOCC operation Λ^Γ (the operation where Charlie's Kraus operators are complex conjugated) to the partially transposed states ρ^Γ and $\tilde{\rho}^\Gamma$ instead. As a consequence, the quantum key repeater rate is invariant under partial transposition: $K_{A\leftrightarrow C\leftrightarrow B}(\rho\otimes\tilde{\rho}) = K_{A\leftrightarrow C\leftrightarrow B}(\rho^\Gamma\otimes\tilde{\rho}^\Gamma)$. The invariance remains true when restricting partially or fully to one-way communication. In the following, we make this statement precise and use it to find upper bounds. We then give examples illustrating the power of the idea and comparing the obtained bounds.

1. Bounds by Key, Relative Entropy of Entanglement and Squashed Entanglement

We start with the above mentioned invariance property.

Lemma 1 *Let ρ and $\tilde{\rho}$ be PPT. Then*

$$K_{A\leftrightarrow C\leftrightarrow B}(\rho\otimes\tilde{\rho}) = K_{A\leftrightarrow C\leftrightarrow B}(\rho^{\Gamma}\otimes\tilde{\rho}^{\Gamma}), \tag{10}$$

where the transpose is taken w.r.t. Charlie's subsystems.

Proof Note that every LOCC protocol can be implemented by many rounds of local POVMs and classical communication. If Charlie uses the complex conjugate of all of his Kraus operators $S_C^{(k)}$, we have another valid LOCC protocol. Since

$$\mathrm{Tr}_C\left[\left(\ldots\otimes(S_C^{(1)*}\cdots S_C^{(r)*})\otimes\ldots\right)\rho^{\Gamma}_{AC_A}\otimes\tilde{\rho}^{\Gamma}_{C_BB}\left(\ldots\otimes({S_C^{(r)*}}^{\dagger}\cdots{S_C^{(1)*}}^{\dagger})\otimes\ldots\right)\right] \tag{11}$$

$$= \mathrm{Tr}_C\left[\left(\ldots\otimes(S_C^{(1)}\cdots S_C^{(r)})\otimes\ldots\right)\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}\left(\ldots\otimes({S_C^{(r)}}^{\dagger}\cdots{S_C^{(1)}}^{\dagger})\otimes\ldots\right)\right], \tag{12}$$

every protocol applied to copies of $\rho\otimes\tilde{\rho}$ has the same output as when the protocol with complex conjugated Kraus operators is applied to $\rho^{\Gamma}\otimes\tilde{\rho}^{\Gamma}$. Consequently, we find

$$K_{A\leftrightarrow C\leftrightarrow B}(\rho\otimes\tilde{\rho}) = K_{A\leftrightarrow C\leftrightarrow B}(\rho^{\Gamma}\otimes\tilde{\rho}^{\Gamma}). \tag{13}$$

Recall that this statement only makes sense for PPT states ρ and $\tilde{\rho}$. □

By the monotonicity of distillable key, we have $K_{A\leftrightarrow C\leftrightarrow B}(\rho\otimes\tilde{\rho}) \leq K_D(\rho_{AC_A})$. Since the relative entropy of entanglement and squashed entanglement are upper bounds on the key rate [14, 24], i.e. the RHS, we obtain the following bounds

Theorem 2 *Let ρ and $\tilde{\rho}$ be PPT. Then*

$$K_{A\leftrightarrow C\leftrightarrow B}(\rho\otimes\tilde{\rho}) \leq \min\left\{K_D(\rho^{\Gamma}), K_D(\tilde{\rho}^{\Gamma})\right\} \leq \min\left\{E_R^{\infty}(\rho^{\Gamma}), E_R^{\infty}(\tilde{\rho}^{\Gamma}), E_{sq}(\rho^{\Gamma}), E_{sq}(\tilde{\rho}^{\Gamma})\right\}, \tag{14}$$

where the transpose is taken w.r.t. Charlie's subsystems.

The relative entropy of entanglement [32] is given by

$$E_R(\rho) = \inf_{\sigma\in\mathrm{SEP}} D(\rho\|\sigma), \tag{15}$$

where SEP denotes the set of separable states. Since it is subadditive, it upper bounds its regularised version

$$E_R^{\infty}(\rho) = \lim_{n\to\infty}\frac{1}{n}E_R(\rho^{\otimes n}). \tag{16}$$

The *squashed entanglement* [33, 34] is given by

$$E_{sq}(\rho_{AB}) = \inf_{\rho_{ABE}}\frac{1}{2}I(A:B|E)_{\rho_{ABE}}, \tag{17}$$

where ρ_{ABE} is an arbitrary extension of ρ_{AB}.

2. Example: PPT state close to p-bit

In the following we exhibit an example, where the RHSs of our bounds are very small, but where the state itself has a high key rate. The idea here is simple, we find PPT states that have high key but whose partial transpose is close to a separable state [35]. More precisely, we present a family of states $\{\rho_{d_s}\}_s$ of increasing dimension which asymptotically reach the gap of 1 between $K_D(\rho_{d_s})$ and $K_{A\leftrightarrow C\leftrightarrow B}(\rho_{d_s}^{\otimes 2})$. Their construction is based on [21]; there, two private bits were mixed to give a PPT key distillable state. Here we take only one of the p-bits and admix the block-diagonal part of the second one. Alternatively, one may use the family of PPT key distillable states introduced in [14, 20], but we omit this argument, since it is more involved.

Proposition 3 *There are PPT states $\rho_{d_s} \in \mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^{d_s} \otimes \mathbb{C}^{d_s})$, obtained by admixing a p_s-fraction of a separable state to a p-bit, such that $\rho_{d_s}^\Gamma$ is p_s-close to a separable state in trace norm. Furthermore, $p_s = \frac{1}{\sqrt{d_s}+1}$ and $d_s \to \infty$ for large s.*

Proof Our construction of ρ_{d_s} is based on [21]. Consider

$$\rho_{d_s} = \frac{1}{2}\begin{bmatrix} (1-p)\sqrt{XX^\dagger} & 0 & 0 & (1-p)X \\ 0 & p\sqrt{YY^\dagger} & 0 & 0 \\ 0 & 0 & p\sqrt{Y^\dagger Y} & 0 \\ (1-p)X^\dagger & 0 & 0 & (1-p)\sqrt{X^\dagger X} \end{bmatrix}, \tag{18}$$

with

$$X = \frac{1}{d_s\sqrt{d_s}} \sum_{i,j=1}^{d_s} u_{ij} |ij\rangle\langle ji| \tag{19}$$

and

$$Y = \sqrt{d_s} X^\Gamma = \frac{1}{d_s} \sum_{i,j=1}^{d_s} u_{ij} |ii\rangle\langle jj|. \tag{20}$$

Here, $p_s = \frac{1}{\sqrt{d_s}+1}$ and u_{ij} are the matrix elements of some (arbitrary) unitary matrix U acting on $\mathbb{C}^{d_s}$ that satisfies $|u_{ij}| = 1/\sqrt{d_s}$ for all i, j. For example, we may set U to be quantum Fourier transform

$$U|k\rangle = \sum_{j=1}^{d_s} \sqrt{\frac{1}{d_s}} e^{2\pi i jk/d_s} |j\rangle. \tag{21}$$

Note that ρ_{d_s} is a mixture of private state (defined by X) with probability $1-p$ and a with separable state $\frac{1}{2}[|0\rangle\langle 0| \otimes |1\rangle\langle 1| \otimes \sqrt{YY^\dagger} + |1\rangle\langle 1| \otimes |0\rangle\langle 0| \otimes \sqrt{Y^\dagger Y}]$ with probability p. It is easy to see that the state is PPT, as $(1-p)X^\Gamma = pY$. So after partial transposition of BB':

$$\rho_{d_s}^\Gamma = \frac{1}{2}\begin{bmatrix} (1-p)\sqrt{XX^\dagger} & 0 & 0 & 0 \\ 0 & p\sqrt{YY^\dagger} & pY & 0 \\ 0 & pY^\dagger & p\sqrt{Y^\dagger Y} & 0 \\ 0 & 0 & 0 & (1-p)\sqrt{X^\dagger X} \end{bmatrix}, \tag{22}$$

which is evidently non-negative, as $\sqrt{XX^\dagger}$ and $\sqrt{X^\dagger X}$ are non-negative by definition, and the middle block is (up to normalisation factor p) a private bit defined by operator Y [20].

Consider now the state ρ_{d_s} dephased on the first qubit of Alice's system (this state is also known as "key attacked state"). It reads:

$$\sigma_{d_s} = \frac{1}{2}\begin{bmatrix} (1-p)\sqrt{XX^\dagger} & 0 & 0 & 0 \\ 0 & p\sqrt{YY^\dagger} & 0 & 0 \\ 0 & 0 & p\sqrt{Y^\dagger Y} & 0 \\ 0 & 0 & 0 & (1-p)\sqrt{X^\dagger X} \end{bmatrix}, \tag{23}$$

and is clearly separable. It is easy to see that

$$\|\rho_{d_s}^\Gamma - \sigma_{d_s}^\Gamma\|_1 = \|(1-p)X^\Gamma\|_1 = \|pY\|_1 = p = \frac{1}{\sqrt{d_s}+1}. \tag{24}$$

This concludes the proof. □

Since the states ρ_s are obtained by admixing a small fraction of a separable state to a p-bit, the key rate of the state is high: Alice and Bob's mutual information in fact equals $1 - h(p_s)$ and the quantum mutual information of Alice and Eve is bounded by $h(p_s)$. Hence, by [15], $K(\rho) \geq 1 - 2h(p_s)$. On the other hand, ρ^Γ is almost separable which implies that $K(\rho^\Gamma)$, $E_R(\rho^\Gamma)$ and $E_{sq}(\rho^\Gamma)$ are small. A particularly good bound is obtained with help of the following lemma.

Lemma 4 *Let $\rho_{ABA'B'} \in \mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^d \otimes \mathbb{C}^d)$ be a PPT($AA' : BB'$) state and assume that its key attacked version $\sigma_{ABA'B'} = \sum_i (|i\rangle\langle i|_A \otimes \mathbb{1})\rho(|i\rangle\langle i|_A \otimes \mathbb{1})$ is separable. Then if $\epsilon = \|\rho^\Gamma - \sigma^\Gamma\|_1 < \frac{1}{3}$, we have*

$$E_R^\infty(\rho^\Gamma) \leq 2\epsilon \log 2d + \eta(\epsilon), \tag{25}$$

where $\eta(\epsilon) = -\epsilon \log \epsilon$.

Proof We start by noting that σ and hence σ^Γ are separable, therefore we have

$$E_R^\infty(\rho^\Gamma) \leq E_R(\rho^\Gamma) \leq D(\rho^\Gamma \| \sigma^\Gamma) \tag{26}$$

We write out the RHS

$$D(\rho^\Gamma \| \sigma^\Gamma) = \operatorname{tr} \rho^\Gamma \log \rho^\Gamma - \operatorname{tr} \rho^\Gamma \log \sigma^\Gamma. \tag{27}$$

and find, since $\operatorname{tr} \rho^\Gamma \log \sigma^\Gamma = \operatorname{tr} \sigma^\Gamma \log \sigma^\Gamma$ (due to the fact that σ is block diagonal) that

$$D(\rho^\Gamma \| \sigma^\Gamma) = H(\sigma^\Gamma) - H(\rho^\Gamma). \tag{28}$$

An application of Fannes' inequality [36] gives the result. □

Theorem 5 *There are PPT states $\rho_{d_s} \in \mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^{d_s} \otimes \mathbb{C}^{d_s})$, satisfying $K_D(\rho_{d_s}) = 1 - 2h(p_s)$ with $p = \frac{1}{\sqrt{d_s}+1}$ and h the binary Shannon entropy, such that $K_{A \leftrightarrow C \leftrightarrow B}(\rho_{d_s} \otimes \rho_{d_s}) \leq 2p \log(2d_s) + \eta(p)$ where $\eta(p) = -p \log p$. In summary, there exist states with*

$$1 \approx K_D(\rho) > K_{A \leftrightarrow C \leftrightarrow B}(\rho \otimes \rho) \approx 0. \tag{29}$$

3. Comparison of the Bounds: Werner States

In the following we show that the bound by the squashed entanglement can be smaller than the one by the relative entropy of entanglement. Recall that it was previously known that squashed entanglement of the antisymmetric Werner state is smaller than its relative entropy of entanglement [24, 37]. Since the antisymmetric Werner state is not PPT, however, this example does not apply directly to our situation. Using a related PPT state from [38], we are able to obtain our goal. We leave open the question of whether the relative entropy of entanglement can be smaller than squashed entanglement. This, however, seems very plausible, as squashed entanglement is lockable [39], and the relative entropy is not [23]. The challenge therefore remains to show locking of squashed entanglement for a PPT state.

Let $\tau_\pm$ be the symmetric and antisymmetric Werner state. In [38] it is shown that

$$\rho^n := w\tau_-^{\otimes n} + (1-w)\tau^{\otimes n} \tag{30}$$

is PPT for $w = 1/(1+z^n)$ for $z = (d+2)/d$, $p = (d+1)/(d+2)$ and $\tau := (1-p)\tau_- + p\tau_+$. Note that

$$E_{sq}(\rho^n) \leq nE_{sq}(\tau_-), \tag{31}$$

since τ is separable. By a result of [24], $E_{sq}(\tau_-) \leq O(1/d)$ hence we find

$$E_{sq}(\rho^n) \leq O(n/d). \tag{32}$$

Let us now derive a lower bound on the regularised relative entropy of this state. Since the relative entropy is not lockable we find

$$E_R((\rho^n)^{\otimes k}) \geq \sum_j \binom{k}{j} w^j (1-w)^{k-j} E_R(\tau_-^{\otimes jn} \otimes \tau^{\otimes (k-j)n}) - kh(w) \tag{33}$$

$$= \sum_j \binom{k}{j} w^j (1-w)^{k-j} E_R(\tau_-^{\otimes jn}) - kh(w) \tag{34}$$

$$\approx E_R(\tau_-^{\otimes wkn}) - kh(w), \tag{35}$$

where we used the separability of τ in the first equality and the law of large numbers in the second. Taking the large k limit we find

$$E_R^\infty(\rho^n) \geq wnE_R^\infty(\tau_-) - h(w). \tag{36}$$

By [24], $E_R^\infty(\tau_-)$ is lower bounded by a constant independent of d. Setting $n = O(d)$ we find $w = O(1)$ (which can be made arbitrarily small) and hence $E_R^\infty(\rho^n) \geq O(n)$. From the bound above $E_{sq}(\rho^n) \leq O(1)$. Hence there are PPT states $\hat{\rho}$ for which

$$E_{sq}(\hat{\rho}) \ll E_R^\infty(\hat{\rho}). \tag{37}$$

Since $\rho := \hat{\rho}^\Gamma$ is again a PPT state we also find that there are PPT states ρ for which

$$E_{sq}(\rho^\Gamma) \ll E_R^\infty(\rho^\Gamma). \tag{38}$$

This shows that the squashed entanglement bound may be stronger than the regularised relative entropy bound.

B. Distinguishability Idea: Single Copy

1. *Trace Norm Bound*

The distinguishability bound that we present below is based on the notion of distinguishing entangled states from separable states by means of restricted measurements (e.g. LOCC measurements). Let us briefly describe the derivation of the bound. Consider a state, $\rho_{in} = \rho_{AC_A} \otimes \tilde{\rho}_{BC_B}$, and suppose ρ_{in} is highly indistinguishable by LOCC operations between C and AB from some triseparable state σ_{in}. Examples of states ρ_{in} with this property were given in [35]: the states are in fact identical private bits $\rho_{AC_A} = \tilde{\rho}_{BC_B} = \rho$ ($K_D(\rho) = 1$) and σ_{in} is of the form $\sigma_{AC_A} \otimes \tilde{\sigma}_{BC_B}$ with $\sigma_{AC_A} = \sigma_{BC_B}$ identical and separable. One may think of them as states that hide entanglement.

Consider now any quantum key repeater protocol Λ. Since Λ is an LOCC operation (between C and A and B), its output when acting on ρ_{in} has to be highly indistinguishable by *arbitrary* CPTP quantum operations from its output when acting on σ_{in}. But this means that ρ_{out} and σ_{out} are close in trace norm. Since σ_{out} is separable this means that ρ_{out} is close to separable and therefore contains almost no key (and is certainly no p-bit).

To show the above reasoning formally, we first recall the notion of maximal probability of discrimination between two states ρ and σ, using some set S of two-outcome POVMs $\{E^0, E^1 = \mathbb{1} - E^0\}$ [35, 40]. By definition we have:

$$p^S(\rho, \sigma) = \sup_{\{E^0, E^1\} \in S} \frac{1}{2} \operatorname{tr} E^0 \rho + \frac{1}{2} \operatorname{tr} E^1 \sigma. \tag{39}$$

In what follows we will consider several sets of operations: LOCC, SEP, PPT and ALL. The set ALL is the set of all two-outcome POVMs. PPT consists only of elements that have a positive partial transpose and SEP contains only separable elements, whereas LOCC are those POVMs that can be implemented by an LOCC protocol. Note that LOCC $\subset$ SEP $\subset$ PPT $\subset$ ALL.

Lemma 6 *For any two states $\rho, \tilde{\rho}$, two separable states $\sigma, \tilde{\sigma}$ and any $\Lambda \in LOCC(A : C : B)$,*

$$\|\hat{\rho} - \hat{\sigma}\|_1 \leq \|(\rho_{AC_A} \otimes \tilde{\rho}_{BC_B})^\Gamma - (\sigma_{AC_A} \otimes \tilde{\sigma}_{BC_B})^\Gamma\|_1, \tag{40}$$

where $\hat{\rho} = \mathrm{Tr}_C \Lambda(\rho_{AC_A} \otimes \tilde{\rho}_{BC_B})$ and $\hat{\sigma} = \mathrm{Tr}_C \Lambda(\sigma_{AC_A} \otimes \tilde{\sigma}_{BC_B})$ are the AB outputs of the protocol.

Proof Since Λ is LOCC, it is a tri-separable map, i.e. has its Kraus representation $\Lambda(\rho) = \sum_i M_A^i \otimes M_B^i \otimes M_C^i(\rho) M_A^{i\dagger} \otimes M_B^{i\dagger} \otimes M_C^{i\dagger}$. In particular it is separable in the cut $AB : C$, which will be crucial in what follows. Moreover, upon input of any two separable states $\sigma_{AC_A} \otimes \sigma_{B_C B}$, the map outputs a state ρ_{ABC} with $\mathrm{Tr}_C \rho_{ABC}$ separable. We now prove the following chain of (in)equalities and comment on them below:

$$\begin{aligned}
1+\frac{1}{2}\|\hat{\rho}-\hat{\sigma}\|_1 &= 2p^{\mathrm{ALL}}(\hat{\rho},\hat{\sigma}) \qquad (41)\\
&= \sup_{\{E^j\}\in \mathrm{ALL}} [\operatorname{tr} E^0\hat{\rho} + \operatorname{tr} E^1\hat{\sigma}] \qquad (42)\\
&= \sup_{\{E^j_{AB}\}\in \mathrm{ALL}} [\operatorname{tr} E^0_{AB} \operatorname{tr}_C \Lambda(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B}) + \operatorname{tr} E^1_{AB} \operatorname{tr}_C \Lambda(\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B})] \qquad (43)\\
&= \sup_{\{E^j_{AB}\}\in \mathrm{ALL}} [\operatorname{tr}(E^0_{AB}\otimes I_C)\Lambda(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B}) + \operatorname{tr}(E^1_{AB}\otimes I_C)\Lambda(\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B})]\\
&\qquad (44)\\
&= \sup_{\{E^j_{AB}\}\in \mathrm{ALL}} \Bigg[\sum_j \operatorname{tr}(M_A^{j\dagger}\otimes M_B^{j\dagger}\otimes M_C^{j\dagger}(E^0_{AB}\otimes I_C)M_A^j\otimes M_B^j\otimes M_C^j(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B}))\\
&\qquad + \sum_j \operatorname{tr}(M_A^{j\dagger}\otimes M_B^{j\dagger}\otimes M_C^{j\dagger}(E^1_{AB}\otimes I)M_A^j\otimes M_B^j\otimes M_C^j(\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B}))\Bigg] \qquad (45)\\
&\leq p^{\mathrm{SEP}(AB:C)}(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B}, \sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B}) \qquad (46)\\
&\leq p^{\mathrm{PPT}(AB:C)}(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B}, \sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B}) \qquad (47)\\
&= \sup_{\{F^j\geq 0, \sum_j F^j=\mathbb{1}, (F^j)^\Gamma\geq 0\}} [\operatorname{tr} F^0(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B}) + \operatorname{tr} F^1(\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B})] \qquad (48)\\
&= \sup_{\{F^j\geq 0, \sum_j F^j=\mathbb{1}, (F^j)^\Gamma\geq 0\}} [\operatorname{tr} {F^0}^\Gamma(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B})^\Gamma + \operatorname{tr} {F^1}^\Gamma(\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B})^\Gamma] \qquad (49)\\
&\leq \sup_{\{\sum_j F^j=\mathbb{1}, (F^j)^\Gamma\geq 0\}} [\operatorname{tr} {F^0}^\Gamma(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B})^\Gamma + \operatorname{tr} {F^1}^\Gamma(\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B})^\Gamma] \qquad (50)\\
&= 2p^{\mathrm{ALL}}((\rho_{AC_A}\otimes\tilde{\rho}_{BC_B})^\Gamma, (\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B})^\Gamma) \qquad (51)\\
&= 1+\frac{1}{2}\|(\rho_{AC_A}\otimes\tilde{\rho}_{BC_B})^\Gamma - (\sigma_{AC_A}\otimes\tilde{\sigma}_{BC_B})^\Gamma\|_1. \qquad (52)
\end{aligned}$$

The first equality is the well known Helstrom formula for optimally distinguishing two quantum states. Subsequently, we simply insert the definitions step by step. Inequality (45) follows from the fact that Λ is a tri-separable map. In the next inequality we use SEP $\subset$ PPT. Then we write this explicitly out and partially transpose all the C systems. Then we drop the positivity constraint on the POVM elements and see that the remaining maximisation extends over all POVMs. Using Helstrom once again concludes the calculation. □

The above lemma shows that the trace norm distance between the output states of any quantum key repeater protocol is upper bounded by the trace norm distance of the partially transposed input states of it. Combining this result with asymptotic continuity of relative entropy of entanglement gives the following theorem:

Theorem 7 *Consider any two states $\rho, \tilde{\rho}$, and separable states $\sigma, \tilde{\sigma}$ in $\mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$ such that $\|\rho^\Gamma - \sigma^\Gamma\|_1 \leq \epsilon$ and $\|\tilde{\rho}^\Gamma - \tilde{\sigma}^\Gamma\|_1 \leq \epsilon$, Then, if $\mu := \min\{\|\rho^\Gamma\|_1, \|\tilde{\rho}^\Gamma\|_1\}$ satisfies $\epsilon' := \epsilon(\mu+1) \leq \frac{1}{3}$, we have*

$$K_{A\leftrightarrow C\leftrightarrow B}^{\textit{single copy}}(\rho \otimes \tilde{\rho}) \leq 4(1+\log d)\epsilon' + 2\eta(\epsilon'), \tag{53}$$

with $\eta(x) = -x\log x$. Here, $K_{A\leftrightarrow C\leftrightarrow B}^{\textit{single copy}}$ is the quantum key repeater rate, i.e. the repeater is restricted to act on single copies $\rho \otimes \tilde{\rho}$ only.

Proof Let us consider $\|(\rho\otimes\tilde{\rho})^\Gamma - (\sigma\otimes\tilde{\sigma})^\Gamma\|_1$. By adding and subtracting either $(\rho\otimes\tilde{\sigma})^\Gamma$ or $(\sigma\otimes\tilde{\rho})^\Gamma$, and by triangle inequality, we obtain

$$\|(\rho \otimes \tilde{\rho})^\Gamma - (\sigma \otimes \tilde{\sigma})^\Gamma\|_1 \leq (\min\{\|\rho^\Gamma\|_1, \|\tilde{\rho}^\Gamma\|_1\} + 1)\epsilon. \tag{54}$$

By Lemma 6 and the asymptotic continuity of the relative entropy of entanglement [41] we find

$$|E_R(\hat{\rho}) - E_R(\hat{\sigma})| \leq 4(1+\log d)\|\hat{\rho} - \hat{\sigma}\|_1 + 2\eta(\|\hat{\rho} - \hat{\sigma}\|_1), \tag{55}$$

which, by separability of $\hat{\sigma}$ implies

$$E_R(\hat{\rho}) \leq 4(1+\log d)\epsilon' + 2\eta(\epsilon'). \tag{56}$$

Since $K_D \leq E_R$ [14, 20] we have proven the claim. □

2. Example: p-bit with X = SWAP

Since the single copy quantum key repeater rate is upper bounded by the general quantum key repeater rate, the example from Section III A 2 can also be used to illustrate the above theorem. We therefore choose to provide an example in this section, which, we believe, is not amenable to the bounds from Section III A nor the techniques we present later on in this paper [42].

We consider $\rho = \tilde{\rho} = \gamma_V$, where γ_V is the private state from [14], shown to be entanglement hiding in [35]. It is defined by (1) for $X = \frac{V}{d_s^2}$ with $V = \sum_{i,j=0}^{d_s-1} |ij\rangle\langle ji|$ the swap operator. Note, that for any private bit described by operator X as in (1), we have $\|\gamma^\Gamma\|_1 = 1 + \|X^\Gamma\|_1$ (see proof of Theorem 6.5 of [35]). Now, following [35], as a state which is separable and highly indistinguishable from γ_V, we take γ_V dephased on the key part of Alice: $\sigma := \tilde{\sigma} := \frac{1}{2}[|0\rangle\langle 0| \otimes |1\rangle\langle 1| \otimes \sqrt{XX^\dagger} + |1\rangle\langle 1| \otimes |0\rangle\langle 0| \otimes \sqrt{X^\dagger X}]$. Then $\|\gamma_V^\Gamma - \sigma^\Gamma\|_1 = \|X^\Gamma\|_1$ and $\|X^\Gamma\|_1 = \|\frac{V^\Gamma}{d_s^2}\|_1 = \|\frac{d_s P_+}{d_s^2}\|_1 = \frac{1}{d_s}$ where $P_+ = \frac{1}{d_s}\sum_{i,j=0}^{d_s-1} |ii\rangle\langle jj|$. Thus, $\|\gamma_V^\Gamma - \sigma^\Gamma\|_1 = \frac{1}{d_s}$, which for $d_s \geq 7$ by Theorem 7 (with $\epsilon' = \frac{2d_s+1}{d_s^2}$) implies that

$$K_{A\leftrightarrow C\leftrightarrow B}^{\text{single copy}}(\gamma_V \otimes \gamma_V) \leq \frac{4(2d_s+1)(\log d_s + 1)}{d_s^2} + 2\eta\left(\frac{2d_s+1}{d_s^2}\right). \tag{57}$$

Note that the RHS of the above inequality vanishes with large d_s. It cannot be exactly zero, though, because perfect p-bits always have some non-zero, albeit sometimes small, distillable entanglement [43]. This means that γ_V, although being a private bit ($K_D(\gamma_V) \geq 1$ by definition), in fact with $K_D(\gamma_V) = 1$ [20], cannot be extended by a single copy quantum key repeater for large enough d_s.

C. Distinguishability Idea: Many Copies

1. *Restricted Relative Entropy Bound*

In this section we derive an asymptotic version of the distinguishability bound, that is, one that upper bounds $K_{A\leftrightarrow C\leftrightarrow B}$. The quantity which upper bounds the quantum key repeater rate measures the distinguishability of the state to the next separable state in terms of the relative entropy distance of the probability distributions that can be obtained by LOCC. The bound almost allows to recover the relative entropy and squashed entanglement bounds. Important is the fact that it does not decrease too much when tracing over a qubit at Alice's or Bob's side, which allows us to extend the results to states that are not exactly PPT but only close to it. It also allows us to derive a reduced squashed entanglement bound.

Let LOCC$(A : B)$ be the set of POVMs which can be implemented with local operations and classical communication. We think of an element of this class as the corresponding CPTP map, i.e. instead of a POVM given by $\{M_i\}$ we consider the CPTP map $M : X \mapsto \sum_i (\operatorname{tr} M_i X)|i\rangle\langle i|$. Note that $M(\rho)$ is a probability distribution for ρ a density operator. Our first bound on the quantum key repeater rate is given in terms of the following quantities:

$$D_{C\leftrightarrow AB}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) := \inf_{\sigma\in\mathrm{SEP}(A:C_A:C_B:B)} \sup_{M\in\mathrm{LOCC}(C:AB)} D(M(\rho\otimes\tilde{\rho})\|M(\sigma)), \tag{58}$$

$$D_{C\rightarrow AB}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) := \inf_{\sigma\in\mathrm{SEP}(A:C_A:C_B:B)} \sup_{M\in\mathrm{LOCC}(C\rightarrow AB)} D(M(\rho\otimes\tilde{\rho})\|M(\sigma)). \tag{59}$$

We denote by D^∞ the regularised versions of the above quantities. Note that for trivial $\tilde{\rho}$, the measures reduce to the measures defined in [22]. Sometimes, we omit the minimisation over separable states in which case we write $D_{C\leftrightarrow AB}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}\|\sigma_{ACB})$.

Before we prove the bound we need an easy lemma that shows that D_{ALL} (as defined by Piani [22]) is normalised to (at least) m on private states γ_m [14, 20] containing at least m bits of pure privacy.

Lemma 8 *For $\tilde{\gamma}_m \approx_\epsilon \gamma_m$ and σ separable we have*

$$D_{ALL}(\tilde{\gamma}_m\|\sigma) \geq (1-\epsilon)m - h(\epsilon). \tag{60}$$

Proof Recall that γ_m is of the form $UP_m \otimes \rho_{A'B'}U^\dagger$ for P_m the projector onto the maximally entangled state in dimension 2^m on systems AB and U a controlled unitary with control A and target $A'B'$. $\rho_{A'B'}$ is arbitrary. We calculate:

$$\begin{aligned} D_{\mathrm{ALL}}(\tilde{\gamma}_m\|\sigma) &\geq D_{\mathrm{ALL}}(\operatorname{tr}_{A'B'} U\tilde{\gamma}_m U^\dagger \| \operatorname{tr}_{A'B'} U\sigma U^\dagger) && (61)\\ &= D_{\mathrm{ALL}}(\tilde{P}_m\|\tilde{\sigma}) && (62)\\ &\geq D(\{\operatorname{tr} P_m\tilde{P}_m, \operatorname{tr}(\mathbb{1}-P_m)\tilde{P}_m\}\|\{\operatorname{tr} P_m\tilde{\sigma}, \operatorname{tr}(\mathbb{1}-P_m)\tilde{\sigma}\}) && (63)\\ &\geq (1-\epsilon)m - h(\epsilon). && (64) \end{aligned}$$

The first inequality holds due to monotonicity of D_{ALL}. Note that $\tilde{P}_m := \operatorname{tr}_{A'B'} U\tilde{\gamma}_m U^\dagger$ is a state ϵ close to P_m. We also defined $\tilde{\sigma} = \operatorname{tr}_{A'B'} U\sigma U^\dagger$. The second inequality is again an application of monotonicity, this time with the measurement map given by the POVM $\{P_m, \mathbb{1}-P_m\}$. The last inequality follows from proof of [20, Lemma 7] which says that $\operatorname{tr} P_m\tilde{\sigma} \leq 1/2^m$ and $\operatorname{tr} P_m\tilde{P}_m \geq 1-\epsilon$, which follows from $\tilde{\gamma}_m \approx_\epsilon \gamma_m$. □

We now come to the main result of this section.

Theorem 9 *The following inequalities hold for all states ρ and $\tilde{\rho}$:*

$$\begin{aligned} K_{A\leftrightarrow C\leftrightarrow B}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}) &\leq D^\infty_{C\leftrightarrow AB}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}), && (65)\\ K_{A\leftarrow C\rightarrow B}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}) &\leq D^\infty_{C\rightarrow AB}(\rho_{AC_A}\otimes\tilde{\rho}_{C_BB}). && (66) \end{aligned}$$

Proof We will start with proving the first bound. Fix $\epsilon > 0$. Then, there is an n and a $\Lambda \in \mathrm{LOCC}(A^n : C^n : B^n)$ (in the following we will suppress n if obvious from the context), such that $r \geq K_{A\leftrightarrow C\leftrightarrow B}(\rho\otimes\tilde{\rho}) - \epsilon$ and $\tilde{\gamma} := \operatorname{tr}_C \Lambda((\rho_{AC_A}\otimes\tilde{\rho}_{C_BB})^{\otimes n}) \approx_\epsilon \gamma_{\lfloor nr\rfloor}$. For $\sigma \in \mathrm{SEP}(A : C_A : C_B : B)$ we have

$$\begin{aligned} &\max_{M\in\mathrm{LOCC}(C:AB)} D(M(\rho^{\otimes n}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})\|M(\sigma_{ACB})) && (67)\\ &\geq \max_{M\in\mathrm{LOCC}(C:AB)} D(M(\operatorname{tr}_C\Lambda(\rho^{\otimes n}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB}))\|M(\operatorname{tr}_C\Lambda(\sigma_{ACB}))) && (68)\\ &= \max_{M\in\mathrm{ALL}(AB)} D(M(\operatorname{tr}_C\Lambda(\rho^{\otimes n}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB}))\|M(\operatorname{tr}_C\Lambda(\sigma_{ACB}))) && (69)\\ &= \max_{M\in\mathrm{ALL}(AB)} D(M(\tilde{\gamma}_{AB})\|M(\tilde{\sigma}_{AB})). && (70) \end{aligned}$$

The first inequality is true as $M\circ\operatorname{tr}_C\circ\Lambda \in \mathrm{LOCC}(C : AB)$. The first equality follows as the arguments have no system C anymore (or equivalently a one-dimensional system C) and since in this case $\text{LOCC}(C : AB) = \text{ALL}(AB)$. In the last equality we have used the definition of $\tilde{\gamma}$ and introduced $\tilde{\sigma} := \text{tr}_C \Lambda(\sigma)$. Noting that $\tilde{\sigma} \in \text{SEP}(A : B)$ is separable (since $\Lambda \in \text{LOCC}(A : C : B)$ and $\sigma \in \text{SEP}(A : C_A : C_B : B) \subset \text{SEP}(A : C : B)$) and that $\tilde{\gamma} \approx_\epsilon \gamma_{\lfloor nr \rfloor}$ we have from Lemma 8:

$$\max_{M \in \text{ALL}(AB)} D(M(\tilde{\gamma}_{AB}) \| M(\tilde{\sigma}_{AB})) \geq (1 - \epsilon)\lfloor nr \rfloor - h(\epsilon). \tag{71}$$

Combining the bounds, minimizing over σ and taking the limit $n \to \infty$ gives

$$D^\infty_{C \leftrightarrow AB}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) \geq (1 - \epsilon) r \tag{72}$$

Since $r \geq K_{A \leftrightarrow C \leftrightarrow B}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) - \epsilon$ and ϵ was arbitrary we have proven the first claim.

The second claim follows by slight modification: restrict Λ to be in $\text{LOCC}(A \leftarrow C \rightarrow B)$ and note that $M \circ \text{tr}_C \circ \Lambda \in \text{LOCC}(C \to AB)$ and that $\text{LOCC}(C \to AB) = \text{ALL}(AB)$ for trivial system C. Then $K_{A \leftrightarrow C \leftrightarrow B}$ will turn into $K_{A \leftarrow C \rightarrow B}$ and $D_{C \leftrightarrow AB}$ into $D_{C \to AB}$. □

2. Properties of the Restricted Relative Entropy Measure

In this section we present three properties of the distinguishability measure, its invariance under partial transposition of the C system, its non-lockability (i.e. the fact that the measure does not decrease too much when a qubit on Alice's or Bob's side is lost) and its LOCC monotonicity.

Lemma 10 *For all states ρ and $\tilde{\rho}$,*

$$D_{C \leftrightarrow AB}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) = D_{C \leftrightarrow AB}(\rho^\Gamma_{AC_A} \otimes \tilde{\rho}^\Gamma_{C_B B}), \tag{73}$$

$$D_{C \to AB}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) = D_{C \to AB}(\rho^\Gamma_{AC_A} \otimes \tilde{\rho}^\Gamma_{C_B B}). \tag{74}$$

Proof It is sufficient to observe that the sets of measurements which we denote by LOCC as a placeholder for either $\text{LOCC}(C : AB)$ or $\text{LOCC}(C \to AB)$ and the set of separable states are invariant under taking partial transpose of systems C (or AB):

$$\min_{\sigma \in \text{SEP}(A:C_A:C_B:B)} \max_{M \in \text{LOCC}} D(M(\rho \otimes \tilde{\rho}) \| M(\sigma)) \tag{75}$$

$$= \min_{\sigma \in \text{SEP}(A:C_A:C_B:B)} \max_{M \in \text{LOCC}} D(M^\Gamma(\rho^\Gamma \otimes \tilde{\rho}^\Gamma) \| M^\Gamma(\sigma^\Gamma)) \tag{76}$$

$$= \min_{\sigma \in \text{SEP}(A:C_A:C_B:B)} \max_{M \in \text{LOCC}} D(M(\rho^\Gamma \otimes \tilde{\rho}^\Gamma) \| M(\sigma)). \tag{77}$$

□

By the monotonicity of the relative entropy, we can upper bound $D^\infty_{C \leftrightarrow AB}$ by the relative entropy of entanglement and, using the invariance of $D^\infty_{C \leftrightarrow AB}$ under partial transpose of the C system

(Lemma 10), obtain $E_R^\infty(\rho) + E_R^\infty(\tilde{\rho})$ and thereby almost recover the relative entropy bound from Theorem 2.

This lets us also conclude that $D^\infty_{A\leftrightarrow B}(\rho)$, which can similarly be upper bounded by $E_R(\rho^\Gamma)$, can be made strictly smaller than $K_D(\rho)$: simply take the states from Proposition 3. This observation was first made in [44] in order to answer a question posed in [45].

Following [23] we will now prove that D^∞ is not lockable.

Lemma 11 *Let $\mathcal{M} = LOCC(C : AB)$ or $LOCC(C \to AB)$. Then*

$$D^\infty_{\mathcal{M}}(\rho_{A_1A_2C_A} \otimes \tilde{\rho}_{C_BB}) \leq D^\infty_{\mathcal{M}}(\rho_{A_1C_A} \otimes \tilde{\rho}_{C_BB}) + I(A_2 : A_1C_A)_\rho. \tag{78}$$

A similar bound holds when part of C_A is lost. In summary, $D^\infty_{\mathcal{M}}$ is non-lockable.

Proof Let us fix $\epsilon > 0$, a state $\sigma \in \mathrm{SEP}(A : C_A : C_B : B)$ and a POVM given by CPTP map E. By [46, Proposition II.2], there exist $2^{n\delta}$ unitaries $U^{(i)}$, where $\delta = I(A_2 : A_1B) + \epsilon$, such that when applied to $\rho^{\otimes n}_{AC_A}$ they decorrelate A_2 from A_1C_A, i.e.

$$\|\hat{\rho}_{AC_A} - \omega_{A_2} \otimes \omega_{A_1C_A}\|_1 \leq \epsilon \tag{79}$$

for some states ω_{A_2} and $\omega_{A_1C_A}$, where we introduced $\hat{\rho}^{(i)}_{AC_A} = U^{(i)}_{A_2} \otimes I_{A_2C_A}\rho^{\otimes n}_{AC_A}U^{(i)\dagger}_{A_2} \otimes I_{A_1C_A}$ and $\sum_i p_i\hat{\rho}^{(i)}_{AC_A} = \hat{\rho}_{AC_A}$. Since the decorrelation map acts as identity on systems A_1B we find, using the triangle inequality,

$$\|\hat{\rho}_{AC_A} - \omega_{A_2} \otimes \rho_{A_1C_A}\|_1 \leq 2\epsilon. \tag{80}$$

By a theorem from [47] the following holds

$$\textstyle\sum_i p_i D(E(\hat{\rho}^{(i)}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB})) - D(E(\sum_i p_i\hat{\rho}^{(i)}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB})) \tag{81}$$

$$\textstyle\leq H(E(\sum_i p_i\hat{\rho}^{(i)}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{C_BB})) - \sum_i p_i H(E(\hat{\rho}^{(i)}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{C_BB})) \leq H(X), \tag{82}$$

where X is a random variable with distribution $\{p_i\}$. Since we have a bound on the number of unitaries, we can bound $H(X)$ and find

$$\sum_i p_i D(E(\hat{\rho}^{(i)}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB})) \leq D(E(\hat{\rho}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB})) + n\delta. \tag{83}$$

Observe that LHS is an average. Hence, there exists an event i_0 such that

$$D(E(\hat{\rho}^{(i_0)}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{CBB})\|E(\sigma_{ACB})) \leq D(E(\hat{\rho}_{AC_A} \otimes \tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB})) + n\delta. \tag{84}$$

Taking the supremum over E,

$$\sup_{E\in C} D(E(\hat{\rho}^{(i_0)}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB}))\leq \sup_{E\in C} D(E(\hat{\rho}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB}))+n\delta. \tag{85}$$

allows us to shift the unitary $U^{(i_0)}$ from $\hat{\rho}^{(i_0)}_{AC_A}=U^{(i_0)}_{A_2}\otimes I_{A_1C_A}\rho^{\otimes n}_{AC_A}U^{(i_0)\dagger}\otimes I_{A_1C_A}$ to $\hat{\sigma}^{(i_0)}_{AC_A}=U^{(i_0)\dagger}_{A_2}\otimes I_{A_1C_A}\sigma U^{(i_0)}_{A_1}\otimes I_{A_2C_A}$ as it is only applied locally on the A_2 systems

$$\sup_{E\in C} D(E(\rho^{\otimes n}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})\|E(\hat{\sigma}^{i_0}_{ACB}))\leq \sup_{E\in C} D(E(\hat{\rho}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})\|E(\sigma_{ACB}))+n\delta. \tag{86}$$

Taking the infimum over $\sigma\in\mathrm{SEP}(A:C_A:C_B:B)$ on both sides, noting in particular that $\hat{\sigma}^{i_0}\in\mathrm{SEP}(A:C_A:C_B:B)$ we find

$$D_{\mathcal{M}}(\rho^{\otimes n}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})\leq D_{\mathcal{M}}(\hat{\rho}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})+n\delta. \tag{87}$$

Now, by asymptotic continuity of $D_{\mathcal{M}}$ [45, Proposition 3], and inequality (80) this gives

$$\begin{aligned} D_{\mathcal{M}}(\rho^{\otimes n}_{AC_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB}) &\leq D_{\mathcal{M}}(\omega_{A_2}\otimes\rho^{\otimes n}_{A_1C_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})+2\epsilon\log(\frac{6d^n}{\epsilon})+n\delta & (88)\\ &= D_{\mathcal{M}}(\rho^{\otimes n}_{A_1C_A}\otimes\tilde{\rho}^{\otimes n}_{C_BB})+2\epsilon\log(\frac{6d^n}{\epsilon})+n\delta, & (89)\end{aligned}$$

where in the last equation we used the fact that $D_{\mathcal{M}}$ stays unchanged when we add or remove a local tensor product state. Taking the limit $n\to\infty$ and subsequently $\epsilon\to 0$ we have proved the claim. □

We conclude with proving the monotonicity of the bound.

Lemma 12 *Let $\Lambda\in\mathcal{M}$ where $\mathcal{M}$ is one of $LOCC(C_A\leftrightarrow A)$, $LOCC(C_A\to A)$, $LOCC(C_B\leftrightarrow B)$ or $LOCC(C_B\to B)$. Then,*

$$D_{LOCC}(\rho\otimes\tilde{\rho})\geq\sum_i p_iD_{LOCC}(\rho_i\otimes\tilde{\rho}), \tag{90}$$

where $\Lambda(\rho)=\sum_i p_i|i\rangle\langle i|\otimes\rho_i$.

Proof We prove the statements for the → case.

$$\begin{aligned} D_{\mathcal{M}}(\rho\otimes\tilde{\rho}) &= \inf_{\sigma\in\mathrm{SEP}(A:C_A:C_B:B)}\max_{M\in\mathrm{LOCC}(C\to AB)} D(M(\rho\otimes\tilde{\rho})\|M(\sigma)) & (91)\\ &\geq \inf_{\sigma\in\mathrm{SEP}(A:C_A:C_B:B)}\max_{M\in\mathrm{LOCC}(C\to AB)} D(M(\Lambda(\rho\otimes\tilde{\rho}))\|M(\Lambda(\sigma))) & (92)\\ &= \inf_{\sigma\in\mathrm{SEP}(A:C_A:C_B:B)}\max_{M_i\in\mathrm{LOCC}(C\to AB)}\sum_i p_iD(M_i(\rho_i\otimes\tilde{\rho})\|M_i(\sigma_i))+D(p\|q), & (93)\end{aligned}$$

where we used $\Lambda(\sigma)=\sum_i q_i|i\rangle\langle i|\otimes\sigma_i$ and without loss of generality $M=\sum_i|i\rangle\langle i|\otimes M_i$. This is lower bounded by

$$\inf_{\sigma_i\in\mathrm{SEP}(A:C_A:C_B:B)}\max_{M_i\in\mathrm{LOCC}(C\to AB)}\sum_i p_iD(M_i(\rho_i\otimes\tilde{\rho})\|M_i(\sigma_i))=\sum_i p_iD_{\mathcal{M}}(\rho_i\otimes\tilde{\rho}). \tag{94}$$

The other cases are similar. □

It is the goal of this section to derive a bound on the one-way quantum key repeater rate by the reduced squashed entanglement. This will be done in two steps. First, we will prove that the one-way LOCC restricted relative entropy measure is upper bounded by squashed entanglement. Second, we will employ the non-lockability of this measure in order to unlock the squashed entanglement.

For the first step, we need two lemmas in order to prepare for the key lemma, Lemma 15.

Lemma 13 *For any two states ρ_{ABE} and σ_{ABE} and for every $M \in LOCC(A^2 \to B^2)$ with output denoted by X there is a sequence $T_n \in LOCC(A^n \to B^n)$ with cq output $X^n B^n$ such that*

$$\lim_{n\to\infty} \frac{1}{n} D(T_n^c(\rho_{AB}^{\otimes n})^{\otimes 2} \| T_n^c(\sigma_{AB}^{\otimes n})^{\otimes 2}) = D(M(\rho_{AB}^{\otimes 2}) \| M(\sigma_{AB}^{\otimes 2})), \tag{95}$$

$$\lim_{n\to\infty} \| T_n^q \otimes \mathrm{id}_E(\rho_{ABE}^{\otimes n}) - \rho_{BE}^{\otimes n} \|_1 = 0, \tag{96}$$

where we defined $T_n^q = \mathrm{tr}_{X^n} \circ T_n$ and $T_n^c = \mathrm{tr}_{B^n} \circ T_n$.

Proof Apply [45, Lemma 5] to the states $\rho \mapsto \rho^{\otimes 2}$ and $\sigma \mapsto \sigma^{\otimes 2}$. Then manipulate the LHS of their first equation: First, we use the additivity of the relative entropy

$$D(T_n^c(\rho_{AB}^{\otimes 2n}) \otimes T_n^c(\rho_{AB}^{\otimes 2n}) \| T_n^c(\sigma_{AB}^{\otimes 2n}) \otimes T_n^c(\sigma_{AB}^{\otimes 2n})) = 2D(T_n^c(\rho_{AB}^{\otimes 2n}) \| T_n^c(\sigma_{AB}^{\otimes 2n})) \tag{97}$$

in order to conclude

$$\lim_{n\to\infty} \frac{1}{n} D(T_n^c(\rho_{AB}^{\otimes 2n}) \| T_n^c(\sigma_{AB}^{\otimes 2n})) = \lim_{n\to\infty} \frac{1}{2n} D(T_n^c(\rho_{AB}^{\otimes 2n}) \otimes T_n^c(\rho_{AB}^{\otimes 2n}) \| T_n^c(\sigma_{AB}^{\otimes 2n}) \otimes T_n^c(\sigma_{AB}^{\otimes 2n})). \tag{98}$$

In a next step we restrict the limit to even n (thereby not changing the limiting value) and make the replacement $n \mapsto n/2$ to obtain

$$\lim_{n\to\infty} \frac{1}{n} D(T_{n/2}^c(\rho_{AB}^{\otimes n})^{\otimes 2} \| T_{n/2}^c(\sigma_{AB}^{\otimes n})^{\otimes 2}). \tag{99}$$

Finally, we redefine $T_{n/2} \mapsto T_n$ and obtain the claim. □

Lemma 14 *For any tri-partite state ρ,*

$$2E_R^\infty(\rho_{B:AE}) \geq D_{A^2\to B^2}^\infty(\rho_{AB}^{\otimes 2}) + 2E_R^\infty(\rho_{B:E}). \tag{100}$$

Proof For a state $\sigma \in \mathrm{SEP}(B:AE)$,

$$
\begin{aligned}
nD(\rho_{ABE}^{\otimes 2}\|\sigma_{ABE}^{\otimes 2}) &= D(\rho^{\otimes 2n}\|\sigma^{\otimes 2n}) && (101)\\
&\geq D(T_n \otimes \mathrm{id}_E(\rho^{\otimes n})^{\otimes 2}\|T_n \otimes \mathrm{id}_E(\sigma^{\otimes n})^{\otimes 2}) && (102)\\
&= D(T_n^c(\rho^{\otimes n})^{\otimes 2}\|T_n^c(\sigma^{\otimes n})^{\otimes 2}) + \sum_{ij} p_i p_j D(\rho_i \otimes \rho_j\|\sigma_i \otimes \sigma_j) && (103)\\
&\geq D(T_n^c(\rho^{\otimes n})^{\otimes 2}\|T_n^c(\sigma^{\otimes n})^{\otimes 2}) + D(T_n^q \otimes \mathrm{id}_E(\rho^{\otimes n}) \otimes T_n^q \otimes \mathrm{id}_E(\rho^{\otimes n})\|\tilde{\sigma} \otimes \tilde{\sigma}) && (104)\\
&\geq D(T_n^c(\rho^{\otimes n})^{\otimes 2}\|T_n^c(\sigma^{\otimes n})^{\otimes 2}) && (105)\\
&\quad + \min_{\tilde{\sigma} \in \mathrm{SEP}(B:E)} D(T_n^q \otimes \mathrm{id}_E(\rho^{\otimes n}) \otimes T_n^q \otimes \mathrm{id}_E(\rho^{\otimes n})\|\tilde{\sigma} \otimes \tilde{\sigma}). && (106)
\end{aligned}
$$

The first inequality follows from the monotonicity of the relative entropy under CPTP maps, the following equality is a direct calculation, where the ensemble $\{p_i, \rho_i\}$ ($\{q_i, \sigma_i\}$) is the output of the instrument $T_n \otimes \mathrm{id}_E$ when applied to $\rho_{ABE}^{\otimes n}$ and $\sigma_{ABE}^{\otimes n}$, respectively. The subsequent inequality is due to convexity of the relative entropy, where we defined the state $\tilde{\sigma} := T_n^q \otimes \mathrm{id}_E(\sigma^{\otimes n})$. Since $T^q \otimes \mathrm{id}_E \in \mathrm{LOCC}(B:AE)$ and $\sigma \in \mathrm{SEP}(B:AE)$, we find $\tilde{\sigma} \in \mathrm{SEP}(B:E)$. This explains the last inequality. Using Lemma 13, the asymptotic continuity of the relative entropy of entanglement [41] and taking the limit $n \to \infty$ proves

$$
D(\rho_{ABE}^{\otimes 2}\|\sigma_{ABE}^{\otimes 2}) \geq D(M(\rho_{AB}^{\otimes 2})\|M(\sigma_{AB}^{\otimes 2})) + \lim_{n\to\infty} \frac{1}{n} \min_{\tilde{\sigma} \in \mathrm{SEP}(B:E)} D(\rho_{BE}^{\otimes n} \otimes \rho_{BE}^{\otimes n}\|\tilde{\sigma}_{BE} \otimes \tilde{\sigma}_{BE}). \quad (107)
$$

We now maximise this statement over measurements, then minimise over σ. This proves

$$
2E_R(\rho_{B:AE}) \geq \inf_{\sigma} \max_{M} D(M(\rho_{AB}^{\otimes 2})\|M(\sigma^{\otimes 2})) + 2E_R^{\infty}(\rho_{B:E}). \quad (108)
$$

The RHS is lower bounded by $D_{A^2\to B^2}(\rho_{AB} \otimes \rho_{AB}) + 2E_R^{\infty}(\rho_{B:E})$. Regularizing this result we obtain the claimed bound. □

Lemma 15

$$
D_{A^2\to B^2}^{\infty}(\rho_{AB} \otimes \rho_{AB}) \leq 4E_{sq}(\rho_{AB}). \quad (109)
$$

Proof From Lemma 14 we have

$$
2E_R^{\infty}(\rho_{B:AE}) - 2E_R^{\infty}(\rho_{B:E}) \geq D_{A^2\to B^2}^{\infty}(\rho_{AB}^{\otimes 2}). \quad (110)
$$

By [48, Lemma 1] the LHS is upper bounded by $2I(A:B|E)_\rho$. Minimizing over all extensions of ρ_{ABE} for a fixed ρ_{AB} proves the claim. □

Since squashed entanglement is lockable [39] but $D^{\infty}_{\mathcal{M}}$ is not, we can improve the squashed entanglement bound. For this we define the *reduced squashed entanglement* [23, 49]:

$$E_{sq\downarrow}(\rho_{AB}) := \inf_{A=A_1A_2, B=B_1B_2} E_{sq}(\rho_{A_1:B_1}) + H(A_2) + H(B_2), \tag{111}$$

where the infimum goes over all splits of A into two subsystems A_1A_2 (likewise for B). Note that, trivially, reduced squashed entanglement is smaller than squashed entanglement. By construction this measure is not lockable and subadditive. Since for every split of the A system (and similarly for the B system)

$$D^{\infty}_{A^2\to B^2}(\rho_{AB}\otimes\rho_{AB}) \leq D^{\infty}_{A_1^2\to B^2}(\rho_{A_1B}\otimes\rho_{A_1B}) + 2I(A_1 : A_2B) \leq 4E_{sq}(\rho_{A_1B}) + 4H(A_1), \tag{112}$$

we obtain the following improved bound.

Lemma 16 *For $\rho_{C_AA} = \rho_{C_BB}$,*

$$D^{\infty}_{C\to AB}(\rho_{AC_A}\otimes\rho_{C_BB}) \leq 4E^{\infty}_{sq\downarrow}(\rho_{AC_A}) \leq 4E_{sq\downarrow}(\rho_{AC_A}). \tag{113}$$

Combining Lemma 16 with Theorem 9 and Lemma 10 we get the following bound.

Corollary 17 *The following inequality holds for all PPT states $\rho_{C_AA} = \rho_{C_BB}$:*

$$K_{A\leftarrow C\to B}(\rho\otimes\rho) \leq 4E^{\infty}_{sq\downarrow}(\rho^{\Gamma}) \leq 4E_{sq\downarrow}(\rho^{\Gamma}). \tag{114}$$

We leave it as an open question whether the relative entropy of entanglement can be much smaller than reduced squashed entanglement, or, in other words, whether the E^{∞}_{R}-bound in Theorem 2 gives a bound stronger than the first bound in Corollary 17.

Interestingly, reduced squashed entanglement can also be used as an upper bound for the traditional distillable entanglement. For this we choose $\tilde{\rho}$ to be the trivial state and apply Lemma 11 and [45, Theorem 2] in order to obtain:

$$D^{\infty}_{A\to B}(\rho_{AB}) \leq D^{\infty}_{A_1\to B}(\rho_{A_1B}) + I(A_1 : A_2B) \leq 2E_{sq}(\rho_{A_1B}) + 2H(A_1). \tag{115}$$

When combined with the bound $D^{\infty}_{A\to B} \geq E_D$ (from [45]) and regularisation we find:

Corollary 18

$$E_D(\rho_{AB}) \leq 2E^{\infty}_{sq\downarrow}(\rho_{AB}). \tag{116}$$

We conjecture that the constant 2 can be replaced by a 1. Note that we could have used $\frac{1}{2}I(A_1 : A_2B)$ in place of $H(A_1)$ in the definition of reduced squashed entanglement. Then, however, the constant could not have been smaller as a simple examples shows. Whereas it is known that $K_D \leq E_{sq}$, we leave explicitly open the question whether $K_D \leq E_{sq\downarrow}$ (even up to a multiplicative constant), a bound that would have paralleled the bound of the classical secret key rate by the reduced intrinsic information: $S(X : Y\|Z) \leq I(X : Y \downarrow\downarrow Z)$ [49].

4. Extensions For Almost PPT States

It is the purpose of this section to extend the squashed entanglement and relative entropy bounds to NPT states that are close to being PPT. We start with some technical lemmas.

The following bounds follow easily from the monotonicity of the relative entropy.

Lemma 19 *For positive $p, \tilde{p}$ we have*

$$D^{\infty}_{LOCC}(\rho^{\otimes p} \otimes \tilde{\rho}^{\otimes \tilde{p}}) \leq pE^{\infty}_R(\rho) + \tilde{p}E^{\infty}_R(\tilde{\rho}), \tag{117}$$

where we defined

$$D^{\infty}_{LOCC}(\rho^{\otimes p} \otimes \tilde{\rho}^{\otimes \tilde{p}}) := \lim_{n\to\infty} \frac{1}{n} D_{LOCC}(\rho^{\lfloor np \rfloor} \otimes \tilde{\rho}^{\lfloor n\tilde{p} \rfloor}). \tag{118}$$

Lemma 20 *Let $\mathcal{M} = LOCC(AB : C)$ or $LOCC(C \to AB)$. Consider a state ρ such that $\rho_+ = p\rho + (1-p)\rho'$ is PPT for some state ρ' (and likewise for $\tilde{\rho}$ and $\tilde{\rho}_+$). Then,*

$$D^{\infty}_{\mathcal{M}}(\rho \otimes \tilde{\rho}) \leq D^{\infty}_{\mathcal{M}}(\rho_+^{1/p} \otimes \tilde{\rho}_+^{1/\tilde{p}}) + h(p)/p + h(\tilde{p})/\tilde{p}. \tag{119}$$

Proof We start by applying Lemma 11 to the state $\rho_f^{\otimes n}$ with $\rho_f = p\rho_{A_1C_A} \otimes |1\rangle\langle 1|_{A_2} + (1-p)\rho'_{A_1C_A} \otimes |0\rangle\langle 0|_{A_2}$ and similarly for $\tilde{\rho}$. We have for all $\epsilon > 0$ and sufficiently large n

$$D_{\mathcal{M}}(\rho_f^{\otimes n} \otimes \tilde{\rho}_f^{\otimes \tilde{n}}) \leq D_{\mathcal{M}}(\rho_+^{\otimes n} \otimes \tilde{\rho}_+^{\otimes \tilde{n}}) + nh(p) + \tilde{n}h(\tilde{p}) + \epsilon(n + \tilde{n}), \tag{120}$$

where we used the fact that $I(A_2 : A_1C_A) \leq h(p)$. We now bound the LHS from below. Note that by Lemma 12 a measurement of the flags results in

$$D_{\mathcal{M}}(\rho_f^{\otimes n} \otimes \tilde{\rho}_f^{\otimes \tilde{n}}) \geq \sum_k \binom{n}{k} p^k(1-p)^{n-k} \sum_{\tilde{k}} \binom{\tilde{n}}{\tilde{k}} \tilde{p}^{\tilde{k}}(1-\tilde{p})^{\tilde{n}-\tilde{k}} D_{\mathcal{M}}(\rho^{\otimes k} \otimes \tilde{\rho}^{\otimes \tilde{k}}). \tag{121}$$

By Lemma 12 we can locally apply partial traces resulting in

$$\sum_{k>n(p-\epsilon)} \binom{n}{k} p^k(1-p)^{n-k} \sum_{\tilde{k}} \binom{\tilde{n}}{\tilde{k}} \tilde{p}^{\tilde{k}}(1-\tilde{p})^{\tilde{n}-\tilde{k}} D_{\mathcal{M}}(\rho^{\otimes \lfloor n(p-\epsilon) \rfloor} \otimes \tilde{\rho}^{\otimes \lfloor \tilde{n}(\tilde{p}-\epsilon) \rfloor}). \tag{122}$$

By the Chernoff bound $k > \lfloor n(p-\epsilon)\rfloor$ with probability $1 - e^{-2n\epsilon^2}$. Hence we find

$$D_{\mathcal{M}}(\rho_f^{\otimes n} \otimes \tilde{\rho}_f^{\otimes \tilde{n}}) \geq (1 - e^{-2n\epsilon^2})(1 - e^{-2\tilde{n}\epsilon^2}) D_{\mathcal{M}}(\rho^{\otimes \lfloor n(p-\epsilon)\rfloor} \otimes \tilde{\rho}^{\otimes \lfloor \tilde{n}(\tilde{p}-\epsilon)\rfloor}). \tag{123}$$

Combining this bound with (120) and taking the limit $\lfloor n(p-\epsilon)\rfloor = \lfloor \tilde{n}(\tilde{p}-\epsilon)\rfloor = m \to \infty$ results in

$$D_{\mathcal{M}}^{\infty}(\rho \otimes \tilde{\rho}) \leq D_{\mathcal{M}}^{\infty}(\rho_+^{1/(p-\epsilon)} \otimes \tilde{\rho}_+^{1/(\tilde{p}-\epsilon)}) + h(p) + h(\tilde{p}) + \epsilon, \tag{124}$$

which proves the statement, since ϵ was arbitrary. □

Combining this statement with Theorem 9 and Lemma 19 we find the following result.

Corollary 21 *Let $\rho \in NPT$ be such that $\rho_+ = p\rho + (1-p)\rho' \in PPT$ for some state ρ' (and likewise for $\tilde{\rho}$). Then*

$$K_{A\leftrightarrow C\leftrightarrow B}(\rho \otimes \tilde{\rho}) \leq \frac{1}{p}(E_R^{\infty}(\rho_+^{\Gamma}) + h(p)) + \frac{1}{\tilde{p}}(E_R^{\infty}(\tilde{\rho}_+^{\Gamma}) + h(\tilde{p})). \tag{125}$$

Similarly we can derive a squashed entanglement bound for NPT states.

Corollary 22 *Let $\rho \in NPT$ be such that $\rho_+ = p\rho + (1-p)\rho' \in PPT$ for some state ρ'. Then*

$$\begin{aligned} K_{A\leftarrow C\rightarrow B}(\rho \otimes \rho) &\leq \frac{2}{p}(2E_{sq\downarrow}^{\infty}(\rho_+^{\Gamma}) + h(p)) && (126)\\ &\leq \frac{2}{p}(2E_{sq\downarrow}(\rho_+^{\Gamma}) + h(p)) && (127)\\ &\leq \frac{2}{p}(2E_{sq}(\rho_+^{\Gamma}) + h(p)). && (128) \end{aligned}$$

5. *Example: Exact p-bit close to being PPT*

In previous examples, we had to admix some noise to the p-bits in order to make them PPT and thereby amenable to the bounds. We can now use our bounds for NPT states that are close to being PPT in order directly obtain bounds for exact p-bits.

Theorem 23 *There is a family of private bits $\gamma_{d_s} \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^{d_s} \otimes \mathbb{C}^{d_s}$ with $K_D(\gamma_{d_s}) = 1$, such that $K_S(\gamma_{d_s} \otimes \gamma_{d_s}) \leq \frac{4q}{1-q}\log 2d_s + \frac{2}{1-q}(\eta(q) + h(q))$ with h binary Shannon entropy, $\eta(q) = -q\log q$ and $q = \frac{1}{\sqrt{d_s}+1}$.*

Proof Take γ_{d_s} to be equal to a private bit defined by X from eq. (19). Its matrix reads

$$\gamma_{d_s} = \frac{1}{2}\begin{bmatrix} \sqrt{XX^\dagger} & 0 & 0 & X \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ X^\dagger & 0 & 0 & \sqrt{X^\dagger X} \end{bmatrix}. \tag{129}$$

By [20, Theorem 4], $K_D(\gamma_{d_s}) \leq 1$. Since γ_{d_s} is a private bit we find $K_D(\gamma_{d_s}) = 1$. Theorem 5 shows that γ_{d_s} becomes PPT if mixed with probability p with the separable state $\sigma' = \frac{1}{2}[|0\rangle\langle 0| \otimes |1\rangle\langle 1| \otimes \sqrt{Y^\dagger Y} + |1\rangle\langle 1| \otimes |0\rangle\langle 0| \otimes \sqrt{YY^\dagger}]$. Moreover, resulting PPT state after partial transposition is highly indistinguishable from the separable state σ_{d_s} presented in (23), as it satisfies (24). Since probability with which σ' is admixed $q = \frac{1}{\sqrt{d_s}+1}$ is small, we can apply Corollary 21 (with $p = 1-q$), and Lemma 4 in order to obtain the desired bound. □

D. Entanglement Measures Idea

1. Entanglement Distillation and Cost Bound

We will now present an upper bound on the quantum key repeater rate that depends on the distillable entanglement of the input state.

Theorem 24 *For input states ρ_{AC_A} and $\tilde{\rho}_{C_B B}$ it holds*

$$K_{A\leftarrow C\leftrightarrow B}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) \leq \frac{1}{2}E_D(\tilde{\rho}_{C_B B}) + \frac{1}{2}E_C(\rho_{AC_A}), \tag{130}$$

$$K_{A\leftarrow C\to B}(\rho_{AC_A} \otimes \tilde{\rho}_{C_B B}) \leq \frac{1}{2}E_D^{C_A\to A}(\rho_{AC_A}) + \frac{1}{2}E_C(\tilde{\rho}_{C_B B}) \tag{131}$$

$$\leq \frac{1}{2}E_D(\rho_{AC_A}) + \frac{1}{2}E_C(\tilde{\rho}_{C_B B}). \tag{132}$$

In case of PPT states, we may also transpose the states on the C system.

Our result implies that if one of the input states is bound entangled or has small distillable entanglement, the other state has to 'compensate' this lack of distillability by its entanglement cost. Before proving Theorem 24, we consider the *classical squashed entanglement* [34], denoted by $E_{sq,c}$, a variant of the squashed entanglement where the extensions are restricted to being classical, i.e. $\rho_{ABE} = \sum_i p_i \rho_{AB}^{(i)} \otimes |i\rangle\langle i|_E$. If we further restrict ourselves to $\rho_{ABE} = \sum_i p_i |\Psi^{(i)}\rangle\langle\Psi^{(i)}|_{AB} \otimes |i\rangle\langle i|_E$, i.e. pure states $\rho_i = |\Psi^{(i)}\rangle\langle\Psi^{(i)}|$, we get the *entanglement of formation* as shown in [34]. Clearly, $E_{sq} \leq E_{sq,c} \leq E_F$, and all inequalities can be strict, for example for the antisymmetric state [24, 37]. Furthermore, in [24, 37, 50] it was shown that $K_D \leq E_{sq}$. The proof of Theorem 24 is based on the following Lemmas. First, a small technical observation:

Lemma 25 *For any bipartite state ρ_{AB}, $E_D^{B\to A}(\rho_{AB}) \geq 2E_{sq,c}(\rho_{AB}) - H(B)_\rho$.*

Proof Using the definition of the classical squashed entanglement and the hashing inequality [15], we have $2E_{sq,c}(\rho_{AB}) \leq I(A:B)_\rho = H(B)_\rho - H(B|A)_\rho \leq H(B)_\rho + E_D^{B\to A}(\rho_{AB})$. □

Lemma 25 gives us the following upper bound on the classical squashed entanglement of τ:

Lemma 26 *For $(A \leftarrow C \leftrightarrow B)$-LOCC protocols resulting in $\tau_{A'B'}$ there holds*

$$E_{sq,c}(\tau_{A'B'}) \leq \frac{1}{2}E_D(\tilde{\rho}_{C_B B}) + \frac{1}{2}E_F(\rho_{AC_A}). \tag{133}$$

Proof Any $(A \leftarrow C \leftrightarrow B)$-LOCC protocol can be divided into two steps. First Charlie and Bob perform an LOCC operation on their subsystems, which yields an ensemble $\{p_i, \sigma^{(i)}_{AC'B'}\}$, classically communicate i to Alice and discard C'. After the first step Alice and Bob have the ensemble $\{p_i, \sigma^{(i)}_{AB'} \otimes |i\rangle\langle i|_a\}$. In a second step, Alice performs a local operation that can depend on i, resulting in state $\tau_{A'B'}$.

Let $\{q_j, |\Psi_j\rangle\langle\Psi_j|_{AC_A}\}$ be an ensemble such that $\rho_{AC_A} = \sum_j q_j |\Psi_j\rangle\langle\Psi_j|_{AC_A}$ and $E_F(\rho_{AC_A}) = \sum_j q_j H(A)_{|\Psi_j\rangle\langle\Psi_j|}$. Applying the first step of the protocol to $|\Psi_j\rangle\langle\Psi_j|_{AC_A} \otimes \tilde{\rho}_{C_B B}$ alone result in the ensemble $\{p_i^{(j)}, \sigma^{(i,j)}_{AB'} \otimes |i\rangle\langle i|_a\}$. By linearity we have $\sigma^{(i)}_{AB'} = \sum_j q_j \sigma^{(i,j)}_{AB'}$. By Lemma 25, for every i, j, we have

$$E_D(\sigma^{(i,j)}_{AB'}) \geq 2E_{sq,c}(\sigma^{(i,j)}_{AB'}) - H(A)_{\sigma^{(i,j)}}. \tag{134}$$

By the concavity of the von Neumann entropy and the fact that the A subsystem remains untouched in the first step, we find

$$\sum_{ij} p_i^{(j)} q_j E_D(\sigma^{(i,j)}_{AB'}) \tag{135}$$

$$\geq 2\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma^{(i,j)}_{AB'}) - \sum_{ij} p_i^{(j)} q_j H(A)_{\sigma^{(i,j)}} \tag{136}$$

$$\geq 2\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma^{(i,j)}_{AB'}) - \sum_{j} q_j H(A)_{|\Psi_j\rangle\langle\Psi_j|} \tag{137}$$

$$= 2\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma^{(i,j)}_{AB'}) - E_F(\rho_{AC_A}). \tag{138}$$

As the second step of the protocol is local, using the convexity and LOCC monotonicity of the classical squashed entanglement [51], we obtain $\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma^{(i,j)}_{AB'}) \geq E_{sq,c}(\tau_{A'B'})$. Note that if Alice and Charlie share a lab they will be able to locally create the ensemble $\{q_j, |\Psi_j\rangle\langle\Psi_j|\}$. This combined with the first part of the protocol provides an $(AC : B)$-LOCC protocol, transferring

$\tilde{\rho}_{C_BB}$ into the ensemble $\{p_i^{(j)}q_j, \sigma_{AB'}^{(i,j)}\}$. By the LOCC monotonicity of the distillable entanglement we have $E_D(\tilde{\rho}_{C_BB}) \geq \sum_{ij} p_i^{(j)} q_j E_D(\sigma_{AB'}^{(i,j)})$, finishing the proof. □

Similarly, we can show the following

Lemma 27 *For $(A \leftarrow C \rightarrow B)$-LOCC protocols resulting in $\tau_{A'B'}$ there holds*

$$E_{sq,c}(\tau_{A'B'}) \leq \frac{1}{2}E_D^{C_A\rightarrow A}(\rho_{AC_A}) + \frac{1}{2}E_F(\tilde{\rho}_{C_BB}), \tag{139}$$

where $E_D^{C_A\rightarrow A}$ describes the one way distillable entanglement.

Proof Any $(A \leftarrow C \rightarrow B)$-LOCC protocol can be divided into two steps. First Charlie performs an operation on his subsystem, which yields an ensemble $\{p_i, \sigma_{AC'B}^{(i)}\}$, classically communicates i to Alice and Bob and discards C'. After the first step Alice and Bob have the ensemble $\{p_i, \sigma_{AB}^{(i)} \otimes |i\rangle\langle i|_a \otimes |i\rangle\langle i|_b\}$. In a second step, Alice and Bob perform local operations that can depend on i, resulting in state $\tau_{A'B'}$.

Let $\{q_j, |\Psi_j\rangle\langle\Psi_j|_{C_BB}\}$ be an ensemble such that $\tilde{\rho}_{C_BB} = \sum_j q_j|\Psi_j\rangle\langle\Psi_j|_{C_BB}$ and $E_F(\tilde{\rho}_{C_BB}) = \sum_j q_j H(B)_{|\Psi_j\rangle\langle\Psi_j|}$. Let applying the first step of the protocol to $\rho_{AC_A} \otimes |\Psi_j\rangle\langle\Psi_j|_{C_BB}$ alone result in the ensemble $\{p_i^{(j)}, \sigma_{AB}^{(i,j)} \otimes |i\rangle\langle i|_a \otimes |i\rangle\langle i|_b\}$. By linearity we have $\sigma_{AB}^{(i)} = \sum_j q_j \sigma_{AB}^{(i,j)}$. By Lemma 25, for every i, j, we have

$$E_D^{B\rightarrow A}(\sigma_{AB}^{(i,j)}) \geq 2E_{sq,c}(\sigma_{AB}^{(i,j)}) - H(B)_{\sigma^{(i,j)}}. \tag{140}$$

By the concavity of the von Neumann entropy and the fact that the B subsystem remains untouched in the first step,

$$\sum_{ij} p_i^{(j)} q_j E_D^{B\rightarrow A}(\sigma_{AB}^{(i,j)}) \tag{141}$$

$$\geq 2\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma_{AB}^{(i,j)}) - \sum_{ij} p_i^{(j)} q_j H(B)_{\sigma^{(i,j)}} \tag{142}$$

$$\geq 2\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma_{AB}^{(i,j)}) - \sum_j q_j H(B)_{|\Psi_j\rangle\langle\Psi_j|} \tag{143}$$

$$= 2\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma_{AB}^{(i,j)}) - E_F(\tilde{\rho}_{C_BB}). \tag{144}$$

As the second step of the protocol is local, using the convexity and LOCC monotonicity of the classical squashed entanglement, we obtain $\sum_{ij} p_i^{(j)} q_j E_{sq,c}(\sigma_{AB}^{(i,j)}) \geq E_{sq,c}(\tau_{A'B'})$. Note that if Charlie and Bob share a lab they will be able to locally create the ensemble $\{q_j, |\Psi_j\rangle\langle\Psi_j|\}$. This combined with the first part of the protocol provides an $(A \leftarrow CB)$-LOCC protocol, transferring

ρ_{AC_A} into the ensemble $\{p_i^{(j)} q_j, \sigma_{AB}^{(i,j)}\}$. By the LOCC monotonicity of the one-way distillable entanglement we have $E_D^{C_A \to A}(\rho_{AC_A}) \geq \sum_{ij} p_i^{(j)} q_j E_D^{B\to A}(\sigma_{AB}^{(i,j)})$, finishing the proof. □

Proof of Theorem 24 Let $\mathcal{M}$ be the class of allowed LOCC protocols and let $\epsilon > 0$. Then there exists n and an $\mathcal{M}$-protocol $\Lambda^{\mathcal{M}}$ such that $\mathrm{Tr}_C \Lambda^{\mathcal{M}} \left((\rho \otimes \tilde{\rho})^{\otimes n} \right) \approx_\epsilon \gamma_{\lfloor nr \rfloor}$ and $r \geq K_{\mathcal{M}}(\rho \otimes \tilde{\rho}) - \epsilon$. Hence, using the fact that $E_{sq}(\gamma_m) \geq m$ for any γ_m [37], as well as the LOCC monotonicity and asymptotic continuity of E_{sq}, it holds

$$nK_{\mathcal{M}}(\rho \otimes \tilde{\rho}) \leq nr + n\epsilon \leq E_{sq}(\gamma_{\lfloor nr \rfloor}) + n\epsilon \leq E_{sq}\left(\mathrm{Tr}_C \Lambda^{\mathcal{M}} \left((\rho \otimes \tilde{\rho})^{\otimes n} \right)\right) + \mathrm{const}\,\epsilon \log(\dim_{A'B'}^n) + f(\epsilon) + n\epsilon, \tag{145}$$

where $f(\epsilon) \to 0$ as $\epsilon \to 0$. By Lemma 26 and 27 for respective classes $\mathcal{M}$ and the fact that $E_{sq} \leq E_{sq,c}$, it holds

$$E_{sq}\left(\mathrm{Tr}_C \Lambda^{A\leftarrow C \leftrightarrow B} \left((\rho \otimes \tilde{\rho})^{\otimes n} \right)\right) \leq \frac{1}{2} E_D(\tilde{\rho}^{\otimes n}) + \frac{1}{2} E_F(\rho^{\otimes n}) \tag{146}$$

and

$$E_{sq}\left(\mathrm{Tr}_C \Lambda^{A\leftarrow C \to B} \left((\rho \otimes \tilde{\rho})^{\otimes n} \right)\right) \leq \frac{1}{2} E_D^{C_A \to A}(\rho^{\otimes n}) + \frac{1}{2} E_F(\tilde{\rho}^{\otimes n}). \tag{147}$$

Let us now divide by n and let $\epsilon \to 0$ and $n \to \infty$. Our bounds then follow from the extensitivity of E_D and the fact that the regularised entanglement of formation equals the entanglement cost. If ρ and $\tilde{\rho}$ are PPT, it can be shown analogously to Lemma 1 that $K_{A\leftarrow C \leftrightarrow B}(\rho \otimes \tilde{\rho}) = K_{A\leftarrow C \leftrightarrow B}(\rho^\Gamma \otimes \tilde{\rho}^\Gamma)$ and $K_{A\leftarrow C \to B}(\rho \otimes \tilde{\rho}) = K_{A\leftarrow C \to B}(\rho^\Gamma \otimes \tilde{\rho}^\Gamma)$, hence we can also partially transpose ρ and $\tilde{\rho}$. □

2. *Example: PPT invariant approximate p-bit (based on data hiding states)*

Note that, even though the results in Section III C may be computed for states without the use of the partial transpose, all examples were in fact computed using that idea. Therefore, until now, we have not been able to demonstrate a nontrivial bound for states that are invariant under the partial transpose operation. It is the goal of this section to demonstrate such an example by help of Theorem 24.

In order to do so, we choose a family of states ρ_m and based on this, consider states of the form $\tilde{\rho}_m := \rho_m \otimes \rho_m^\Gamma$. Note that $\tilde{\rho}_m$ is locally equivalent (by bilocal swap) to its partial transposition and therefore our previous bounds based on the partial transpose idea give no nontrivial results. As we show below, however, for our choice of $\tilde{\rho}_m$ we find $E_D(\tilde{\rho}_m) = 0$ and $E_C(\tilde{\rho}_m) \lesssim 1$. Inserting this into Theorem 24, we find

$$K_{A\leftarrow C \leftrightarrow B}(\tilde{\rho}_m \otimes \tilde{\rho}_m) \lesssim \frac{1}{2}, \tag{148}$$

which is significantly smaller than $K_D(\tilde{\rho}_m) \gtrsim 1$ (see below).

In order to construct ρ_m, we consider a family of states on $B\left(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes (\mathbb{C}^{d^k} \otimes \mathbb{C}^{d^k})^{\otimes m}\right)$ given in [14]:

$$\hat{\rho}_{p,d,k,m} = \frac{1}{N_m} \begin{bmatrix} [p(\frac{\tau_1+\tau_2}{2})]^{\otimes m} & 0 & 0 & [p(\frac{\tau_1-\tau_2}{2})]^{\otimes m} \\ 0 & [(\frac{1}{2}-p)\tau_2]^{\otimes m} & 0 & 0 \\ 0 & 0 & [(\frac{1}{2}-p)\tau_2]^{\otimes m} & 0 \\ [p(\frac{\tau_1-\tau_2}{2})]^{\otimes m} & 0 & 0 & [p(\frac{\tau_1+\tau_2}{2})]^{\otimes m} \end{bmatrix}, \tag{149}$$

where $N_m = 2(p^m) + 2(\frac{1}{2}-p)^m$, $\tau_1 = (\frac{\rho_a+\rho_s}{2})^{\otimes k}$ and $\tau_2 = (\rho_s)^{\otimes k}$, while ρ_s and ρ_a are the d-dimensional symmetric and antisymmetric Werner state, respectively.

The state $\hat{\rho}_{p,d,k,m}$ is PPT iff $p \leq \frac{1}{3}$ and $\frac{1-p}{p} \geq (\frac{d}{d-1})^k$ [14]. We satisfy this condition by setting $p = \frac{1}{3}$, $d = m^2$ and $k = m$, as then $(\frac{d}{d-1})^k < 2$ for $m \geq 2$. Then we define

$$\rho_m := \hat{\rho}_{1/3,m^2,m,m}, \tag{150}$$

with $m \geq 2$. Since also $\tilde{\rho}_m$ is PPT, it is bound entangled and we find $E_D(\tilde{\rho}_m) = 0$. The following lemma assures us of the fact that entanglement of formation of $\tilde{\rho}_m$ is bounded by approximately one.

Lemma 28 *$\tilde{\rho}_m = \rho_m \otimes \rho_m^\Gamma$ for ρ_m defined in eq. (150) satisfies $E_C(\tilde{\rho}_m) \leq E_F(\tilde{\rho}_m) \leq 1 + \frac{2m^2 \log(2m)}{2^m+1}$. Note that this bound is approximately equal to one for large m.*

Proof Observe first that $E_F(\tilde{\rho}_m) \leq E_F(\rho_m) + E_F(\rho_m^\Gamma)$ due to the subadditivity of E_F. We show now, that $E_F(\rho_m) \leq 1$. Indeed, observe that (for $x = \frac{(1/2-p)^m}{N_m}$)

$$\begin{aligned} \rho_m = (1-2x)\left[\frac{1}{2}|\psi_+\rangle\langle\psi_+| \otimes S_{\text{even}} + \frac{1}{2}|\psi_-\rangle\langle\psi_-| \otimes S_{\text{odd}}\right] + \\ 2x\left[\frac{1}{2}|01\rangle\langle 01| \otimes \tau_2^{\otimes m} + \frac{1}{2}|10\rangle\langle 10| \otimes \tau_2^{\otimes m}\right], \end{aligned} \tag{151}$$

where S_{even} is a uniform mixture (with probability $2^{-(m-1)}$) of all states $\tau_{i_1} \otimes \cdots \otimes \tau_{i_m}$ such that 2 occurs even number of times in string $(i_1, \ldots, i_m)$, and S_{odd} is defined analogously, but with number of 2 being odd, $|\psi_\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$. It is clear from the above formula, that the state ρ_m can be created from 2-qubit maximally entangled state appropriately correlated to the sequences of length m of separable hiding states τ_i, and mixed with probability $2x$ with a separable state $\frac{1}{2}(|01\rangle\langle 01| \otimes \tau_2^{\otimes m} + |10\rangle\langle 10| \otimes \tau_2^{\otimes m})$.

We now bound $E_F(\rho_m^\Gamma)$ from above. Note that

$$\rho_m^\Gamma = \frac{1}{N_m}\begin{bmatrix} [p(\frac{\tau_1+\tau_2}{2})^\Gamma]^{\otimes m} & 0 & 0 & 0 \\ 0 & [(\frac{1}{2}-p)\tau_2^\Gamma]^{\otimes m} & [p(\frac{\tau_1-\tau_2}{2})^\Gamma]^{\otimes m} & 0 \\ 0 & [p(\frac{\tau_1-\tau_2}{2})^\Gamma]^{\otimes m} & [(\frac{1}{2}-p)\tau_2^\Gamma]^{\otimes m} & 0 \\ 0 & 0 & 0 & [p(\frac{\tau_1+\tau_2}{2})^\Gamma]^{\otimes m} \end{bmatrix}, \tag{152}$$

Observe, that $[(\frac{\tau_1+\tau_2}{2})^\Gamma]$ is a separable state, and, therefore, by the convexity of entanglement of formation, $E_F(\rho_m^\Gamma) \leq 2xE_F(\rho'_m)$ where the state ρ'_m is formed by middle block of the above matrix:

$$\rho'_m = \frac{1}{2(\frac{1}{2}-p)^m}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & [(\frac{1}{2}-p)\tau_2^\Gamma]^{\otimes m} & [p(\frac{\tau_1-\tau_2}{2})^\Gamma]^{\otimes m} & 0 \\ 0 & [p(\frac{\tau_1-\tau_2}{2})^\Gamma]^{\otimes m} & [(\frac{1}{2}-p)\tau_2^\Gamma]^{\otimes m} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{153}$$

Since $x \leq \frac{1}{2^m}$, we can safely bound $E_F(\rho'_m)$ by the logarithm of the local dimension of ρ'_m, which equals $2m^{2m^2}$:

$$E_F(\rho_m^\Gamma) \leq 2x \times 2m^2 \log(2m). \tag{154}$$

The assertion follows by inserting $p = 1/3$ and observing that the entanglement cost is upper bounded by the entanglement of formation. □

In the following we show that $K_D(\tilde{\rho}_m) \gtrsim 1$ in the limit of large m. We start by noting that $K_D(\tilde{\rho}_m) \geq K_D(\rho_m)$ and it therefore suffices to lower bound $K_D(\rho_m)$. We first apply a privacy squeezing operation to ρ_m, which gives ρ_m^{ps} [20]. Note, that this operation on ρ_m amounts to the replacement of the blocks of the matrix given in eq. (149) by their respective trace norms. In turn, the ρ_m^{ps} is a 2-qubit state described by the matrix:

$$\begin{bmatrix} a & 0 & 0 & b \\ 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ b & 0 & 0 & a \end{bmatrix}, \tag{155}$$

where $a = \frac{p^m}{N_m}$, $x = \frac{(1/2-p)^m}{N_m}$ and (by eq. 141 of [20]) $b = \frac{(p(1-2^{-m}))^m}{N_m}$. Now, using the fact that the distillable key of ρ_m is lower bounded by the Devetak-Winter quantity of a ccq state of the ρ_m^{ps} (see Corollary 4.26 of [35]), we observe that:

$$K_D(\rho_m) \geq 1 - H(a+b, a-b, x, x), \tag{156}$$

where H is the Shannon entropy. This is what we aimed to prove, as in the limit of large m the above considered distribution approaches $(1,0,0,0)$ for our choice of p. □

Lemmas 26 and 27 are new inequalities for entanglement measures. It might be worth asking, both from a practical and an abstract point of view, whether there are more inequalities of that kind for other entanglement measures. First, let us note that $E(\tau) \leq pE(\rho)+(1-p)E(\tilde{\rho})$ is trivially fulfilled for all LOCC-monotonic measures E and all $0 \leq p \leq 1$. What would be interesting instead, is a relation of the form

$$E(\tau) \leq pE_D(\tilde{\rho}) + (1-p)E(\rho) \quad \text{or} \quad E(\tau) \leq pE_D(\rho) + (1-p)E(\tilde{\rho}), \tag{157}$$

for some measure E and some weight p. If we had a quantum repeater that iterates the swapping operation many times, and bound entangled input states, E would be reduced by a factor $1-p$ with every step. For measures that upper bound the distillable key, such as E_C, E_F, E_{sq}, $E_{sq,c}$, E_R or E_R^∞, this would be a significant limitation to quantum key repeaters with bound entangled input states. The same would hold, if we replaced E_D by E_N or $E_{R,\mathrm{PPT}}$.

We will now show that for $E = E_F$, the entanglement of formation, and $E = E_C$, the entanglement cost, (157) cannot hold for all input states. Assume that Bob and Charlie apply the following LOCC protocol. Charlie performs a generalised Bell state measurement $|\Psi^{\nu\mu}\rangle\langle\Psi^{\nu\mu}|_C$, where $|\Psi^{\nu\mu}\rangle = \frac{1}{\sqrt{d}}\sum_j \omega^{j\nu}|j\rangle \otimes |j+\mu\rangle$ and $\omega = e^{\frac{2\pi i}{d}}$. (Here and in the following the addition is performed modulo d.) Charlie then communicates thr result ν, μ classically to Alice and Bob. Upon receiving message, Bob performs $U^{\nu\mu} = \sum_j \omega^{j\nu}|j\rangle\langle j+\mu|$. Alice and Bob then store μ classically. Charlie's subsystem is then discarded, i.e. given to Eve.

Proposition 29 *For the protocol described above, and any $0 < p \leq 1$, there exist states $\rho, \tilde{\rho}$ such that for $E = E_F$ and $E = E_C$*

$$E(\tau_{AB}) > pE_D(\tilde{\rho}_{C_B B}) + (1-p)E(\rho_{AC_A}) \text{ and } E(\tau_{AB}) > pE_D(\rho_{AC_A}) + (1-p)E(\tilde{\rho}_{C_B B}), \tag{158}$$

where τ is the state resulting from the protocol.

Our counterexamples are of the form $\rho_{AB} = \sum_{i,k=0}^{d-1} a_{ik}|ii\rangle\langle kk|$, which admits a purification $|\Phi\rangle_{ABE} = \frac{1}{\sqrt{d}}\sum_i |ii\rangle \otimes |u_i\rangle$, where $a_{ik} = \frac{1}{d}\langle u_k|u_i\rangle$ and the $|u_i\rangle$ are normalised. Such states are called *maximally correlated*. It is easy to see that $\rho_A = \rho_B = \frac{\mathbb{1}}{d}$. For maximally correlated states the entanglement measures involved simplify and τ can be easily calculated. In particular (see [52] and references therein),

$$E_D(\rho_{AB}) = E_R(\rho_{AB}) = \log d - H(\rho) \tag{159}$$

and

$$E_C(\rho_{AB}) = E_F(\rho_{AB}) = \log d - I_{\text{acc}}\left(\left\{\frac{1}{d}, |u_i\rangle\right\}\right), \tag{160}$$

where $I_{\text{acc}}\left(\left\{\frac{1}{d}, |u_i\rangle\right\}\right) = \sup_{\{A_j\}\text{POVM}} I(i : j)$ is the *accessible information.* Before proceeding with our counterexample for E_F and E_C let us note that (157) with E_R is trivially fulfilled for all maximally correlated states. This can be seen by using the fact that for maximally correlated states $E_R = E_D$ and a simple application of the LOCC monotonicity of E_D.

Lemma 30 *Let ρ_{AC_A} and $\tilde{\rho}_{C_BB}$ be maximally correlated, with purifications $|\Phi^1\rangle_{AC_AE_A} = \frac{1}{\sqrt{d}}\sum_i |ii\rangle_{AC_A} \otimes |u_i\rangle_{E_A}$ and $|\Phi^2\rangle_{C_BBE_B} = \frac{1}{\sqrt{d}}\sum_i |ii\rangle_{C_BB} \otimes |v_i\rangle_{E_B}$, respectively. Then for every $0 < p \leq 1$, (157) with $E = E_F$ or $E = E_C$ implies*

$$\frac{1}{d}\sum_\mu I_{acc}\left(\left\{\frac{1}{d}, |u_i\rangle \otimes |v_{i+\mu}\rangle\right\}\right) \geq pH(\tilde{\rho}) \text{ and} \tag{161}$$

$$\frac{1}{d}\sum_\mu I_{acc}\left(\left\{\frac{1}{d}, |u_i\rangle \otimes |v_{i+\mu}\rangle\right\}\right) \geq pH(\rho). \tag{162}$$

Proof Let $0 < p \leq 1$. Let us first show that maximally correlated states preserve their structure under the protocol assumed in Proposition 29. The protocol results in a state $\tau_{AA'BB'}$ purified by

$$|\tilde{\Phi}\rangle = \sum_{\nu\mu}(\mathbb{1}_{AE_AE_B} \otimes |\Psi^{\nu\mu}\rangle\langle\Psi^{\nu\mu}|_C \otimes U_B^{\nu\mu})|\Phi^1\rangle_{AC_AE_A} \otimes |\Phi^2\rangle_{C_BBE_B} \otimes |\mu\mu\rangle_{ab}|\nu\mu\rangle_{\tilde{E}} \tag{163}$$

$$= \frac{1}{\sqrt{d}}\sum_\mu \underbrace{\frac{1}{\sqrt{d}}\sum_i |ii\rangle_{AB} \otimes |u_i\rangle_{E_A} \otimes |v_{i+\mu}\rangle_{E_B}}_{=:|\tilde{\Phi}^\mu\rangle} \otimes |\mu\mu\rangle_{ab} \otimes \underbrace{\frac{1}{\sqrt{d}}\sum_\nu |\Psi^{\nu\mu}\rangle_C \otimes |\nu\mu\rangle_{\tilde{E}}}_{=:|w_\mu\rangle}. \tag{164}$$

Clearly, $\tau^\mu_{AB} := \text{Tr}_{E_AE_B}|\tilde{\Phi}^\mu\rangle\langle\tilde{\Phi}^\mu|$ is maximally correlated and $\{|w_\mu\rangle\}$ are orthogonal. Therefore Alice and Bobs final state is given by $\tau_{AaBb} = \frac{1}{d}\sum_\mu \tau^\mu_{AB} \otimes |\mu\mu\rangle\langle\mu\mu|_{ab}$. By the convexity and LOCC monotonicity of E_F, it holds that $E_F(\tau) = \frac{1}{d}\sum_\mu E_F(\tau^\mu)$. Since we are dealing with maximally correlated states, the same holds true for E_C. Now, assume that we have (157) with $E = E_F$ or $E = E_C$. Inserting (159) and (160) into (157) gives us

$$\frac{1}{d}\sum_\mu I_{\text{acc}}\left(\left\{\frac{1}{d}, |u_i\rangle \otimes |v_{i+\mu}\rangle\right\}\right) \geq pH(\tilde{\rho}) + (1-p)I_{\text{acc}}\left(\left\{\frac{1}{d}, |u_i\rangle\right\}\right) \tag{165}$$

and the same for ρ and $|v_i\rangle$. Since the accessible information is always non-negative, this implies the Lemma. □

Hence, if we can find an example such that $I_{\text{acc}}(\{\frac{1}{d}, |u_i\rangle \otimes |v_{i+\mu}\rangle\}) < pS(\rho)$ and $I_{\text{acc}}(\{\frac{1}{d}, |u_i\rangle \otimes |v_{i+\mu}\rangle\}) < pS(\tilde{\rho})$ for all μ we will have Proposition 29. For this, we make the following ansatz:

$$|\Phi^1\rangle_{AA'C_AC'_AE_A} = \frac{1}{\sqrt{dn}} \sum_{i=1}^{d} \sum_{j=1}^{n} |ii\rangle_{AC_A} \otimes |jj\rangle_{A'C'_A} \otimes U^j|i\rangle_{E_A}, \tag{166}$$

$$|\Phi^2\rangle_{C_BC'_BBB'E_B} = \frac{1}{\sqrt{dn}} \sum_{i=1}^{d} \sum_{j=1}^{n} |ii\rangle_{C_BB} \otimes |jj\rangle_{C'_BB'} \otimes V^j|i\rangle_{E_B}, \tag{167}$$

where U^j, V^j are unitaries. This is a generalisation of the *flower states* introduced in [23] (see [39]). Replacing the index i with (i, j), hence also d with dn, it is easy to see that those are maximally correlated states. Since $\mathrm{Tr}_{AA'C_AC'_A}|\Phi^1\rangle\langle\Phi^1| = \mathrm{Tr}_{C_BC'_BBB'}|\Phi^2\rangle\langle\Phi^2| = \frac{\mathbb{1}}{d}$, we also have $H(\rho) = H(\tilde{\rho}) = \log d$. Consequently, Proposition 29 follows from Lemma 30 and the next proposition.

Proposition 31 *There exists $d_0 \in \mathbb{N}$ such that for all $d \geq d_0$ and $n = d^8$ there are $2n$ unitaries $U^1, \dots, U^n, V^1, \dots, V^n$ such that for all $\alpha = 1, \dots, n$, $\beta = 1, \dots, d$,*

$$I_{acc}\left(\left\{\frac{1}{dn}, U^j|i\rangle_{E_A} \otimes V^{j+\alpha}|i+\beta\rangle_{E_B}\right\}\right) \leq \mathcal{O}(1). \tag{168}$$

Before we can prove Proposition 31 we need several technical lemmas. Let $n, d \in \mathbb{N}$.

Lemma 32 *For random unitaries U^j, V^j, $j = 1, \dots, n$, $\alpha \in \{1, \dots, n\}$, $\beta \in \{1, \dots, d\}$, and $0 < \delta < \frac{1}{2}$, it holds*

$$\Pr\left\{\frac{1}{dn}\sum_{i=1}^{d}\sum_{j=1}^{n} U^j|i\rangle\langle i|U^{j\dagger} \otimes V^{j+\alpha}|i+\beta\rangle\langle i+\beta|V^{j+\alpha\dagger} \notin \left[\frac{1-\delta}{d^2}\mathbb{1}, \frac{1+\delta}{d^2}\mathbb{1}\right]\right\} \leq 2d^2 \exp\left(-\frac{n\delta^2}{d2\ln 2}\right). \tag{169}$$

Proof Let $\alpha \in \{1, \dots, n\}$, $\beta \in \{1, \dots, d\}$ and $0 < \delta < \frac{1}{2}$. Then,

$$\mathbb{E}_{\mathbf{UV}}\frac{1}{dn}\sum_{i=1}^{d}\sum_{j=1}^{n} U^j|i\rangle\langle i|U^{j\dagger} \otimes V^{j+\alpha}|i+\beta\rangle\langle i+\beta|V^{j+\alpha\dagger} \tag{170}$$

$$= \mathbb{E}_U U|0\rangle\langle 0|U^\dagger \otimes \mathbb{E}_U U|0\rangle\langle 0|U^\dagger = \frac{\mathbb{1}}{d^2}, \tag{171}$$

so [53, Thm. 19] can be applied, yielding the desired property. □

Lemma 33 *For all $\alpha \in \{1, \dots, n\}$, $\beta \in \{1, \dots, d\}$ and $0 < \delta < \frac{1}{2}$, if $n \geq 6d$ and*

$$\frac{1}{dn}\sum_{i=1}^{d}\sum_{j=1}^{n} U^j|i\rangle\langle i|U^{j\dagger} \otimes V^{j+\alpha}|i+\beta\rangle\langle i+\beta|V^{j+\alpha\dagger} \in \left[\frac{1-\delta}{d^2}\mathbb{1}, \frac{1+\delta}{d^2}\mathbb{1}\right], \tag{172}$$

then

$$I_{acc}\left(\left\{\frac{1}{dn}, U^j|i\rangle_{E_A} \otimes V^{j+\alpha}|i+\beta\rangle_{E_B}\right\}\right) \leq \log dn - \inf_{|\varphi\rangle} \tilde{H}^{\alpha\beta}_{\varphi,\delta}(\boldsymbol{U}, \boldsymbol{V}), \tag{173}$$

where $\boldsymbol{U} = (U^1, \ldots, U^n)$, $\boldsymbol{V} = (V^1, \ldots, V^n)$ *and*

$$\tilde{H}^{\alpha\beta}_{\varphi,\delta}(\boldsymbol{U}, \boldsymbol{V}) = \sum_{i=1}^{d}\sum_{j=1}^{n} \eta\left(\frac{d}{n(1+\delta)}\left|\langle\varphi|_{E_AE_B} U^j|i\rangle_{E_A} \otimes V^{j+\alpha}|i+\beta\rangle_{E_B}\right|^2\right), \tag{174}$$

with $\eta(x) = -x\log x$.

Proof Let $\alpha \in \{1, \ldots, n\}$, $\beta \in \{1, \ldots, d\}$ and $0 < \delta < \frac{1}{2}$. Without loss of generality, the optimisation in I_{acc} can be restricted to rank 1 POVMs, hence

$$I_{\text{acc}}\left(\left\{\frac{1}{dn}, U^j|i\rangle_{E_A} \otimes V^{j+\alpha}|i+\beta\rangle_{E_B}\right\}\right) = \sup_{\{\mu_k|\varphi_k\rangle\langle\varphi_k|\}\text{ rank-1 POVM}} I(ij:k) \tag{175}$$

$$= \log dn - \inf_{\{\mu_k|\varphi_k\rangle\langle\varphi_k|\}} \sum_k p(k) H\big(p(ij|k) : i = 1\ldots d, j = 1\ldots n\big) \tag{176}$$

$$\leq \log dn - \inf_{|\varphi_k\rangle \in \mathcal{H}_{E_AE_B}} H\big(p(ij|k) : i = 1\ldots d, j = 1\ldots n\big), \tag{177}$$

where

$$p(ijk) = \frac{\mu_k}{dn}\left|\langle\varphi_k|U^j|i\rangle \otimes V^{j+\alpha}|i+\beta\rangle\right|^2, \tag{178}$$

$$p(k) = \sum_{i=1}^{d}\sum_{j=1}^{n} p(ijk) \text{ and } p(ij|k) = \frac{p(ijk)}{p(k)}. \tag{179}$$

By assumption $p(k) \in \left[\frac{(1-\delta)\mu_k}{d^2}, \frac{(1+\delta)\mu_k}{d^2}\right]$, hence

$$p(ij|k) \geq \frac{d}{n(1+\delta)}\left|\langle\varphi_k|U^j|i\rangle \otimes V^{j+\alpha}|i+\beta\rangle\right|^2 \tag{180}$$

and

$$p(ij|k) \leq \frac{d}{n(1-\delta)}\left|\langle\varphi_k|U^j|i\rangle \otimes V^{j+\alpha}|i+\beta\rangle\right|^2 \leq \frac{1}{e}, \tag{181}$$

for $n \geq 6d$. As $\eta(x)$ is increasing for $x \leq \frac{1}{e}$,

$$H\big(p(ij|k) : i = 1, \ldots, d, j = 1, \ldots, n\big) \geq \sum_{i=1}^{d}\sum_{j=1}^{n} \eta\left(\frac{d}{n(1+\delta)}\left|\langle\varphi|U^j|i\rangle \otimes V^{j+\alpha}|i+\beta\rangle\right|^2\right), \tag{182}$$

finishing the proof. □

Next, we lower bound $\inf_{|\varphi\rangle} \tilde{H}^{\alpha\beta}_{\varphi,\delta}(\mathbf{U}, \mathbf{V})$ using the following *concentration of measure* result:

Theorem 34 *(Theorem 2.4 in [54]) Let* $(\mathcal{X}, g)$ *be a compact connected smooth Riemannian manifold with Ricci curvature* $\geq Ric_{min}(\mathcal{X}) > 0$ *equipped with the normalised Riemannian volume element* $d\mu = \frac{dv}{V}$. *Then for any* λ*-Lipschitz function* F *on* X *and any* $r \geq 0$,

$$\mu\left(\{F \leq \mathbb{E}F - r\}\right) \leq \exp\left(-\frac{Ric_{min}(\mathcal{X})r^2}{2\lambda^2}\right). \tag{183}$$

In order to apply Theorem 34 we need to lower bound the expectation value of $\tilde{H}$.

Lemma 35 *There exists d_1, such that for $d \geq d_1$, $n = d^8$, $|\varphi\rangle \in \mathcal{H}_{E_A E_B}$, $\alpha \in \{1, \ldots, n\}$, $\beta \in \{1, \ldots, d\}$ and $\delta = \frac{1}{\log dn}$ we have*

$$\mathbb{E}_{\boldsymbol{UV}} \tilde{H}^{\alpha\beta}_{\varphi,\delta}(\boldsymbol{U}, \boldsymbol{V}) \geq \log dn - \mathcal{O}(1), \tag{184}$$

where we are using the Haar measure on $\mathcal{SU}(d)^{2n}$.

For the proof see Section III D 4. We also need the fact that $\mathcal{SU}(d)^{2n}$ is a compact connected smooth Riemannian manifold with positive Ricci curvature (for details see Section III D 4). Next, we need to upper bound the Lipschitz constant of $\tilde{H}$ with respect to the Riemannian metric of $\mathcal{SU}(d)^{2n}$.

Lemma 36 *For every $n > d \geq 8$, $\alpha \in \{1, \ldots, n\}$, $\beta \in \{1, \ldots, d\}$, $0 < \delta < \frac{1}{2}$ and $|\varphi\rangle \in \mathcal{H}_{E_A E_B}$, the Lipschitz constant $\tilde{\lambda}$ of $\tilde{H}^{\alpha\beta}_{\varphi,\delta}$ is upper bounded*

$$\tilde{\lambda} \leq \frac{8d}{\sqrt{n}} \log n. \tag{185}$$

The proof can be found in Section III D 4. Apart from applying Theorem 34 to $\tilde{H}$, we will need the following net result:

Lemma 37 *(Lemma II.4 in [55]) For $0 < x < 1$ there exists a set $\mathcal{M}$ of unit vectors in $\mathcal{H}$ with $|\mathcal{M}| \leq \left(\frac{5}{x}\right)^{2 \dim \mathcal{H}}$ such that for every unit vector $|\varphi\rangle \in \mathcal{H}$ there exists $|\tilde{\varphi}\rangle \in \mathcal{M}$ with $\||\varphi\rangle - |\tilde{\varphi}\rangle\|_2 \leq \frac{x}{2}$. Such an $\mathcal{M}$ is called an "x-net".*

Finally, we will need the Lipschitz constant of $\hat{H}_{\mathbf{UV}} : \mathcal{H}_{E_A E_B} \to \mathbb{R}$, $\hat{H}_{\mathbf{UV}}(|\varphi\rangle) = \tilde{H}^{\alpha\beta}_{\varphi,\delta}(\mathbf{U}, \mathbf{V})$.

Lemma 38 *For every $\boldsymbol{U}, \boldsymbol{V}$, $n > d \geq 8$, $\alpha \in \{1, \ldots, n\}$, $\beta \in \{1, \ldots, d\}$ and $0 < \delta < \frac{1}{2}$ the Lipschitz constant $\hat{\lambda}$ of $\hat{H}_{\boldsymbol{UV}}$ is upper bounded*

$$\hat{\lambda} \leq 4\sqrt{2} d \log n. \tag{186}$$

For the proof see Section III D 4.

Proof of Proposition 31 Let $0 < r < 1$, $0 < \delta < \frac{1}{4}$, $d \geq 8$ and $n = d^8$. By Lemma 37 there exists an $\frac{r}{8\sqrt{2} d \log n}$-net $\mathcal{M}$ of pure states in $\mathcal{H}_{E_A E_B}$ with $|\mathcal{M}| \leq \left(\frac{40\sqrt{2} d \log n}{r}\right)^{2d^2}$. We will first show that there exists a d_0 such that for $d \geq d_0$ there exist $2n$ unitaries $U^1, \ldots, U^n, V^1, \ldots, V^n$ fulfilling

(i) $\tilde{H}^{\alpha\beta}_{\tilde{\varphi},\delta}(\mathbf{UV}) \geq \mathbb{E}_{\mathbf{UV}} \tilde{H}^{\alpha\beta}_{\tilde{\varphi},\delta} - \frac{r}{4}$ $\forall \alpha \in \{1, \ldots, n\}, \beta \in \{1, \ldots, d\}, |\tilde{\varphi}\rangle \in \mathcal{M}$,

(ii) $\frac{1}{dn}\sum_{i=1}^{d}\sum_{j=1}^{n} U^j|i\rangle\langle i|U^{j\dagger}\otimes V^{j+\alpha}|i+\beta\rangle\langle i+\beta|V^{j+\alpha\dagger}\in\left[\frac{1-\delta}{d^2}\mathbb{1},\frac{1+\delta}{d^2}\mathbb{1}\right]$ $\forall\alpha\in\{1,\dots,n\},\beta\in\{1,\dots,d\}$.

Using Theorem 34, Lemma 32 and the union bound, we get

$$\Pr\{\text{not }(i)\text{ or not }(ii)\}\le nd|\mathcal{M}|\exp\left(-\frac{cdr^2}{32\tilde{\lambda}^2}\right)+2nd^3\exp\left(-\frac{n\delta^2}{2d\ln 2}\right) \tag{187}$$

$$\le\frac{1}{2}\exp\left(\left(\ln 4d+\frac{80\sqrt{2}d^3}{r}\right)8\log d-\frac{cr^2d^7}{131072(\log d)^2}\right)+\frac{1}{2}\exp\left(\ln 4+11\ln d-\frac{d^7\delta^2}{2\ln 2}\right), \tag{188}$$

where it has been used that $\mathrm{Ric}_{\min}(d)=cd$ (see Section III D 4). Both exponents can be made negative for large enough d_0 and $d\ge d_0$, implying that $\Pr\{\text{not }(i)\text{ or not }(ii)\}<1$; hence the desired unitaries exist. Now we will show that this implies Proposition 31. By (ii) and Lemma 33,

$$I_{\text{acc}}\left(\left\{\frac{1}{dn},U^j|i\rangle_{E_A}\otimes V^{j+\alpha}|i+\beta\rangle_{E_B}\right\}\right)\le\log dn-\inf_{|\varphi\rangle}\tilde{H}^{\alpha\beta}_{\varphi,\delta}(\mathbf{U},\mathbf{V}). \tag{189}$$

By the definition of the infimum, there exists $|\varphi_0\rangle\in\mathcal{H}_{E_AE_B}$ such that $\tilde{H}^{\alpha\beta}_{\varphi_0,\delta}(\mathbf{U},\mathbf{V})<\inf_{|\varphi\rangle}\tilde{H}^{\alpha\beta}_{\varphi,\delta}(\mathbf{U},\mathbf{V})+\frac{r}{4}$. By Lemma 37, $|\mathcal{M}|$ contains a state $|\tilde{\varphi}_0\rangle$ such that $\||\varphi_0\rangle-|\tilde{\varphi}_0\rangle\|_2\le\frac{r}{16\sqrt{2}d\log n}$. By Lemma 38 then,

$$\left|\tilde{H}^{\alpha\beta}_{\varphi_0,\delta}(\mathbf{U},\mathbf{V})-\tilde{H}^{\alpha\beta}_{\tilde{\varphi}_0,\delta}(\mathbf{U},\mathbf{V})\right|\le\frac{r}{4}. \tag{190}$$

Consequently $\tilde{H}^{\alpha\beta}_{\tilde{\varphi}_0,\delta}(\mathbf{U},\mathbf{V})\le\tilde{H}^{\alpha\beta}_{\varphi_0,\delta}(\mathbf{U},\mathbf{V})+\frac{r}{4}<\inf_{|\varphi\rangle}\tilde{H}^{\alpha\beta}_{\varphi,\delta}(\mathbf{U},\mathbf{V})+\frac{r}{2}$. Setting $d\ge\max\{d_0,d_1\}$ and $\delta=\frac{1}{\log dn}$, we obtain

$$I_{\text{acc}}\left(\left\{\frac{1}{dn},U^j|i\rangle_{E_A}\otimes V^{j+\alpha}|i+\beta\rangle_{E_B}\right\}\right)<\log dn-\tilde{H}^{\alpha\beta}_{\tilde{\varphi}_0,\delta}(\mathbf{U},\mathbf{V})+\frac{r}{2} \tag{191}$$

$$\le\log dn-\mathbb{E}_{\mathbf{U},\mathbf{V}}\tilde{H}^{\alpha\beta}_{\tilde{\varphi}_0,\delta}+\frac{3r}{4} \tag{192}$$

$$\le\mathcal{O}(1), \tag{193}$$

where the second and third inequalities are due to (i) and Lemma 35, respectively. □

4. Technical Lemmas

We will now briefly review some facts about the Riemannian geometry of the special unitary group.

Lemma 39 *$\mathcal{SU}(d)$, thought of as a sub-manifold in $\mathbb{C}^{d\times d}$, and equipped with the Hilbert-Schmidt inner product on its tangent spaces, is a compact connected Riemannian manifold.*

Proof It is known that $\mathcal{SU}(d)$ is a real semi-simple compact connected Lie group [56]. Every real Lie group is a real smooth manifold. Clearly, the Hilbert-Schmidt inner product is a positive definite bilinear form. It is also easy to see that it is smooth. Let $U \in \mathcal{SU}(d)$ and X, Y be some smooth vector fields on $\mathcal{SU}(d)$, i.e. smooth mappings of $\mathcal{SU}(d)$ into its tangent bundle. As it is a composition of smooth maps, the map $U \mapsto \operatorname{Tr}\big(X(U)^\dagger, Y(U)\big)$ is smooth. Hence the Hilbert-Schmidt inner product on the tangent spaces is what is referred to as a "Riemannian metric". A smooth manifold endowed with a Riemannian metric is a Riemannian manifold [57]. □

From [58], we know that there exists $c > 0$ such that

$$\mathrm{Ric}_{\mathrm{min}}(d) := \inf \mathrm{Ric}(x, x) = cd. \tag{194}$$

The infimum is taken over all tangent unit vectors and Ric denotes the Ricci curvature.

Now we can define a Riemannian distance, which is a metric, for $\mathcal{SU}(d)$

$$d_{\mathcal{SU}(d)}(U, U') = \inf_{\gamma:[0,1]\to\mathcal{SU}(d) \text{ s.t. } \gamma(0)=U, \gamma(1)=U'} \int_0^1 \|\gamma'(t)\|_{HS}\, dt. \tag{195}$$

The Cartesian product $\mathcal{SU}(d)^{2n}$ is a Riemannian manifold as well [54]. As for its metric, we have

Lemma 40 *The Riemannian distance of a Cartesian product $\mathcal{M} \times \mathcal{N}$ of Riemannian manifolds is given by the Pythagorean theorem*

$$d_{\mathcal{M}\times\mathcal{N}}((U, V), (\tilde{U}, \tilde{V})) = \sqrt{d_{\mathcal{M}}(U, \tilde{U})^2 + d_{\mathcal{N}}(V, \tilde{V})^2}, \tag{196}$$

for $U, \tilde{U} \in \mathcal{M}$, $V, \tilde{V} \in \mathcal{N}$.

Proof We know that for tangent vectors x, y, $\|(x, y)\|^2 = \|x\|^2 + \|y\|^2$. We also need the fact that the the length of a curve $L(\gamma) = \int_0^1 \|\gamma'(t)\|\, dt$ is independent of the parametrisation, i.e. for an increasing function $\tau : [0, 1] \to [0, 1]$, it holds $L(\gamma \circ \tau) = L(\gamma)$. Hence it is always possible to find a parametrisation such that $\|\gamma'(t)\|$ is constant, so $L(\gamma) = \|\gamma'(t)\|$. Consequently,

$$\begin{aligned} d_{\mathcal{M}\times\mathcal{N}}((U, V), (\tilde{U}, \tilde{V})) &= \inf_{\gamma\tilde{\gamma}} \int_0^1 \sqrt{\|\gamma'(t)\|^2 + \|\tilde{\gamma}'(t)\|^2} dt &(197)\\ &= \inf_{\gamma\tilde{\gamma}} \sqrt{L(\gamma)^2 + L(\tilde{\gamma})^2} &(198)\\ &= \sqrt{d_{\mathcal{M}}(U, \tilde{U})^2 + d_{\mathcal{N}}(V, \tilde{V})^2}, &(199)\end{aligned}$$

which is what we wanted. □

The minimum Ricci curvature for a Cartesian product of manifolds is just the smallest curvature of the factors. Hence Theorem 34 can be applied to $\tilde{H}$.

Let us now present the proofs that were omitted in the previous section.

Proof of Lemma 35 Let $d \geq 2$, $n = d^8$, $|\varphi\rangle \in \mathcal{H}_{E_A E_B}$, $\alpha \in \{1, \dots, n\}$ and $\beta \in \{1, \dots, d\}$. We need to lower bound $\mathbb{E}\tilde{H}$. For a probability distribution $\{p_i\}$ it holds that $H_2(p) = -\log\left(\sum_i p_i^2\right) \leq \sum_i \eta(p_i) = H(p)$. Here, however, we have $\tilde{p}_{ij} = \frac{d}{n(1+\delta)} \left|\langle\varphi|U^j|i\rangle \otimes V^{j+\alpha}|1+\beta\rangle\right|^2$. Note that $0 \leq \tilde{p}_{ij} \leq \frac{d}{n} \leq \frac{1}{e}$. The $\{\tilde{p}_{ij}\}$ are, in general, no probability distribution. However, Lemma 32 tells us that they are most likely close to one. Namely, for $0 < \delta < \frac{1}{4}$,

$$P\left(\sum_{i=1}^{d}\sum_{j=1}^{n} \tilde{p}_{ij} \notin \left[\frac{1-\delta}{1+\delta}, 1\right]\right) \leq 2d^2 \exp\left(-\frac{n\delta^2}{d\, 2\ln 2}\right). \tag{200}$$

In order to stop H_2 from diverging, let us add a little perturbation that keeps $\tilde{p}_{ij}$ away from 0. Namely, we define

$$\hat{p}_{ij} = (1-\epsilon)\tilde{p}_{ij} + \epsilon\frac{1}{dn}. \tag{201}$$

By concavity and monotonicity of η on $[0, \frac{1}{e}]$,

$$\eta(\hat{p}_{ij}) \leq \eta((1-\epsilon)\tilde{p}_{ij}) + \eta\left(\frac{\epsilon}{nd}\right) \leq \eta(\tilde{p}_{ij}) + \eta\left(\frac{\epsilon}{nd}\right). \tag{202}$$

Hence, choosing $\epsilon = \frac{1}{\log dn}$, we obtain $H(\tilde{p}) \geq H(\hat{p}) - \mathcal{O}(1)$. Next, let us note that if $\sum_{ij} \tilde{p}_{ij} \in \left[\frac{1-\delta}{1+\delta}, 1\right]$, it also holds $\sum_{ij} \hat{p}_{ij} \in \left[\frac{1-\delta}{1+\delta}, 1\right]$. Let us call this event G. If G is true, by Jensen's inequality,

$$H(\hat{p}) \geq \sum_{ij} \hat{p}_{ij} H_2(\hat{p}) - \eta\left(\sum_{ij} \hat{p}_{ij}\right) \geq \frac{1-\delta}{1+\delta} H_2(\hat{p}) - \eta\left(\frac{1-\delta}{1+\delta}\right). \tag{203}$$

Hence,

$$\begin{aligned}
\mathbb{E}_{\mathbf{UV}} H(\tilde{p}) &\geq \mathbb{E}_{\mathbf{UV}} H(\hat{p}) - \mathcal{O}(1) && (204)\\
&\geq \int_G d\mathbf{UV}\; H(\hat{p}) - \mathcal{O}(1) && (205)\\
&\geq \frac{1-\delta}{1+\delta} \int_G d\mathbf{UV}\; H_2(\hat{p}) - \mathcal{O}(1) && (206)\\
&= \frac{1-\delta}{1+\delta} \left(\mathbb{E}_{\mathbf{UV}} H_2(\hat{p}) - \int_{\mathbf{UV} \notin G} d\mathbf{UV}\; H_2(\hat{p})\right) - \mathcal{O}(1) && (207)\\
&\geq \frac{1-\delta}{1+\delta} \left(\mathbb{E}_{\mathbf{UV}} H_2(\hat{p}) - 2d^2 \exp\left(-\frac{n\delta^2}{d\, 2\ln 2}\right) \log\frac{dn}{\epsilon^2}\right) - \mathcal{O}(1), && (208)
\end{aligned}$$

so it is sufficient to lower bound the expectation value of $H_2(\hat{p})$.

$$\mathbb{E}_{\mathbf{UV}} H_2(\hat{p}) \geq -\log \left(\mathbb{E}_{\mathbf{UV}} \sum_{ij} \hat{p}_{ij}^2 \right) \tag{209}$$

$$= -\log \left(nd \, \mathbb{E}_{UV} \hat{p}_{00}^2 \right) \tag{210}$$

$$= -\log \left(nd \left((1-\epsilon)^2 \mathbb{E}_{UV} \tilde{p}_{00}^2 + \frac{2\epsilon(1-\epsilon)}{nd} \mathbb{E}_{UV} \tilde{p}_{00} + \frac{\epsilon^2}{n^2 d^2} \right) \right), \tag{211}$$

where

$$\mathbb{E}_{UV} \tilde{p}_{00} \leq \frac{d}{n} \mathbb{E}_{UV} \operatorname{Tr} \left(|\varphi\rangle\langle\varphi| U|0\rangle\langle 0|U^\dagger \otimes V|0\rangle\langle 0|V^\dagger \right) \tag{212}$$

$$= \frac{d}{n} \operatorname{Tr} \left(|\varphi\rangle\langle\varphi| (\mathbb{E}_U U|0\rangle\langle 0|U^\dagger)^{\otimes 2} \right) \tag{213}$$

$$= \frac{1}{nd} \tag{214}$$

and, using a 2-design,

$$\mathbb{E}_{UV} \tilde{p}_{00}^2 \leq \frac{d^2}{n^2} \mathbb{E}_{UV} \operatorname{Tr} \left(|\varphi\rangle\langle\varphi| U|0\rangle\langle 0|U^\dagger \otimes V|0\rangle\langle 0|V^\dagger \right)^2 \tag{215}$$

$$= \frac{d^2}{n^2} \operatorname{Tr} \left(|\varphi\rangle\langle\varphi|^{\otimes 2} \left((\mathbb{E}_U U|0\rangle\langle 0|U^\dagger)^{\otimes 2} \right)^{\otimes 2} \right) \tag{216}$$

$$= \frac{4}{n^2 (d+1)^2} \operatorname{Tr} \left(|\varphi\rangle\langle\varphi|^{\otimes 2}_{E_A E_B} \Pi^+_{E_A E_A} \otimes \Pi^+_{E_B E_B} \right) \tag{217}$$

$$\leq \frac{4}{n^2 d^2}, \tag{218}$$

where Π^+ denotes the projector onto the symmmetric subspace. Hence,

$$\mathbb{E}_{\mathbf{UV}} H_2(\hat{p}) \geq \log nd - \log \left(4(1-\epsilon)^2 + 2(1-\epsilon)\epsilon + \epsilon^2 \right) \geq \log nd - \log 7. \tag{219}$$

Choosing $\delta = \frac{1}{\log dn}$, for large enough d_1 and $d \geq d_1$ we obtain

$$\mathbb{E}_{\mathbf{UV}} \tilde{H}^{\alpha\beta}_{\varphi,\delta}(\mathbf{U}, \mathbf{V}) = \mathbb{E}_{\mathbf{UV}} H(\tilde{p}) \geq \log dn - \mathcal{O}(1), \tag{220}$$

and we are done. □

Before proving Lemma 36, we need to upper bound the Lipschitz constant of the function $H'_{\beta\delta} : \bigoplus_{j=1}^n \mathcal{H}_{E_A E_B} \to \mathbb{R}$,

$$H'_{\beta\delta}(|\phi_1\rangle, \ldots, |\phi_n\rangle) = \sum_{i=1}^d \sum_{j=1}^n \eta \left(\frac{d}{n(1+\delta)} \operatorname{Tr} \left(|i\rangle\langle i| \otimes |i+\beta\rangle\langle i+\beta| |\phi_j\rangle\langle\phi_j| \right) \right). \tag{221}$$

Note that for $|\phi_j\rangle = U^{j\dagger} \otimes V^{j+\alpha\dagger}|\varphi\rangle$,

$$\tilde{H}^{\alpha\beta}_{\varphi\delta}(\mathbf{UV}) = H'_{\beta\delta}(U^{1\dagger} \otimes V^{1+\alpha\dagger}|\varphi\rangle, \ldots, U^{n\dagger} \otimes V^{n+\alpha\dagger}|\varphi\rangle). \tag{222}$$

Lemma 41 *For all $n > d \geq 8$, $0 < \delta < \frac{1}{2}$, $\beta \in \{1, \ldots, d\}$ the Lipschitz constant λ' of $H'_{\beta\delta}$ is upper bounded*

$$\lambda' \leq \frac{4\sqrt{2}d}{\sqrt{n}} \log n. \tag{223}$$

Proof Let $n > d \geq 8$, $0 < \delta < \frac{1}{2}$ and $\beta \in \{1, \ldots, d\}$. We will make use of the fact that $\lambda'^2 = \sup_{\langle\phi_j|\phi_j\rangle \leq 1 \forall j} \nabla H'_{\beta\delta} \cdot \nabla H'_{\beta\delta}$. Writing $|\phi_j\rangle = \sum_{lm=1}^{d} \phi_{l,m}^{(j)} |lm\rangle$, we get

$$H'_{\beta\delta}(|\phi_1\rangle, \ldots, |\phi_n\rangle) = \sum_{i=1}^{d} \sum_{j=1}^{n} \eta \left(\frac{d}{n(1+\delta)} \left| \phi_{i,i+\beta}^{(j)} \right|^2 \right) = \sum_{i=1}^{d} \sum_{j=1}^{n} \eta \left(c r_{ij}^2 \right), \tag{224}$$

where we have defined $b = \frac{d}{n(1+\delta)}$ and $r_{ij} = \left| \phi_{i,i+\beta}^{(j)} \right|$. By assumption $b < 1$. Computing the gradient we obtain

$$\sup_{\langle\phi_j|\phi_j\rangle \leq 1 \forall j} \nabla H'_{\beta\delta} \cdot \nabla H'_{\beta\delta} = \sup_{\langle\phi_j|\phi_j\rangle \leq 1 \forall j} \frac{4b}{(\ln 2)^2} \sum_{i=1}^{d} \sum_{j=1}^{n} b r_{ij}^2 \left(\ln (b r_{ij}^2) + 1 \right)^2 \tag{225}$$

$$\leq \sup_{\sum_{i=1}^{d} r_{ij}^2 \leq 1 \forall j} \frac{4b}{(\ln 2)^2} \left(\sum_{i=1}^{d} \sum_{j=1}^{n} b r_{ij}^2 \left(\ln b r_{ij}^2 \right)^2 + bn \right) \tag{226}$$

$$= \frac{4bn}{(\ln 2)^2} \left(\sup_{\sum_{i=1}^{d} y_i \leq b,\, y_i \geq 0 \forall i} \sum_{i=1}^{d} y_i (\ln y_i)^2 + b \right) \tag{227}$$

Using Lagrange multipliers, it can be shown that for $d \geq 8$ the maximum is attained at $y_i = \frac{b}{d}$, hence

$$\lambda'^2 \leq \frac{4b^2 n}{(\ln 2)^2} \left(\left(\ln \frac{b}{d} \right)^2 + 1 \right) \leq \frac{32 d^2}{n} (\log n)^2, \tag{228}$$

finishing the proof. □

Proof of Lemma 36 Let $U_1, \ldots, U_n, V_1, \ldots, V_n, U'_1, \ldots, U'_n, V'_1, \ldots, V'_n \in \mathcal{SU}(d)$. Then

$$\left| \tilde{H}_{\varphi\delta}^{\alpha\beta}(\mathbf{U}, \mathbf{V}) - \tilde{H}_{\varphi\delta}^{\alpha\beta}(\mathbf{U}', \mathbf{V}') \right| \leq \lambda' \left\| \bigoplus_{j=1}^{n} \left(U_j^\dagger \otimes V_{j+\alpha}^\dagger - U_j'^\dagger \otimes V_{j+\alpha}'^\dagger \right) |\varphi\rangle \right\|_2 \tag{229}$$

$$= \lambda' \sqrt{\sum_{j=1}^{n} \left\| \left(U_j^\dagger \otimes V_{j+\alpha}^\dagger - U_j'^\dagger \otimes V_{j+\alpha}'^\dagger \right) |\varphi\rangle \right\|_2^2} \tag{230}$$

$$\leq \lambda' \sqrt{\sum_{j=1}^{n} \left\| \left(U_j^\dagger \otimes V_{j+\alpha}^\dagger - U_j'^\dagger \otimes V_{j+\alpha}'^\dagger \right) \right\|_\infty^2} \tag{231}$$

$$\leq \sqrt{2} \lambda' \sqrt{\sum_{j=1}^{n} \left\| U_j - U'_j \right\|_\infty^2 + \sum_{j=1}^{n} \left\| V_j - V'_j \right\|_\infty^2}. \tag{232}$$

Since

$$d_{\mathrm{Riem}}(U,U') = \inf_{\gamma} \int_a^b \left\|\gamma'(t)\right\|_{HS} dt \geq \inf_{\gamma} \left\| \int_a^b \gamma'(t)dt \right\|_{HS} \tag{233}$$

$$= \inf_{\gamma} \|\gamma(a) - \gamma(b)\|_{HS} = \left\|U - U'\right\|_{HS} \geq \left\|U - U'\right\|_{\infty}, \tag{234}$$

we get $\tilde{\lambda} = \sqrt{2}\lambda'$. Applying Lemma 41 finishes the proof. □

Proof of Lemma 38 Let $\mathbf{U}, \mathbf{V} \in \mathcal{SU}(d)^d$, $\alpha \in \{1, \ldots, n\}$, $\beta \in \{1, \ldots, d\}$. Then for all $|\varphi\rangle, |\varphi'\rangle \in \mathcal{H}$,

$$\left|\hat{H}_{\mathbf{UV}}(|\varphi\rangle) - \hat{H}_{\mathbf{UV}}(|\varphi'\rangle)\right| \leq \lambda' \left\| \bigoplus_{j=1}^{n} U^j \otimes V^{j+\alpha} \left(|\varphi\rangle - |\varphi'\rangle\right) \right\|_2 \tag{235}$$

$$= \lambda' \sqrt{\sum_{j=1}^{n} \|U^j \otimes V^{j+\alpha} \left(|\varphi\rangle - |\varphi'\rangle\right)\|_2^2} \tag{236}$$

$$= \lambda' \sqrt{n} \left\||\varphi\rangle - |\varphi'\rangle\right\|_2, \tag{237}$$

where we have used that the Hilbert space norm is unitarily invariant. □

Acknowledgments

Part of this work was done when the authors attended the program “Mathematical Challenges in Quantum Information”, Aug-Dec 2013, at the Isaac Newton Institute for Mathematical Sciences, Cambridge, whose hospitality is gratefully acknowledged. MC thanks Robert Pisarczyk for helpful discussions and the members of the Centre for Quantum Information and Foundations, DAMTP, University of Cambridge for their hospitality. KH thanks Michał and Paweł Horodecki and Jonathan Oppenheim for helpful discussions relating to Lemma 11.

SB was supported by the ERC Advanced Grant “IRQUAT”. MC acknowledges financial support by the Swiss National Science Foundation (grants PP00P2-128455, 20CH21-138799 (CHIST-ERA project CQC)), the Swiss National Center of Competence in Research ’Quantum Science and Technology (QSIT)’, the Swiss State Secretariat for Education and Research supporting COST action MP1006 and the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement no. 337603. KH is supported by ERC AdG grant QOLAPS. AW was supported by the Spanish MINECO, project FIS2008-01236 with the support of FEDER funds, the EC STREP “RAQUEL”, the ERC Advanced Grant “IRQUAT”, and the Philip Leverhulme Trust.

[1] C. A. Fuchs and A. Peres, Phys. Rev. A **53**, 2038 (1996).

[2] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Computer Society Press, New York, Bangalore, India, December 1984, 1984), pp. 175–179.

[3] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[4] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[5] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[6] C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).

[7] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **81**, 5932 (1998).

[8] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).

[9] D. Stucki, N. Walenta, F. Vannel, R. T. Thew, N. Gisin, H. Zbinden, S. Gray, C. R. Towery, and S. Ten, New J. Phys. **11**, 075003 (2009).

[10] N. Sangouard, C. Simon, H. De Riedmatten, and N. Gisin, Rev. Mod. Phys. **83**, 33 (2011).

[11] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).

[12] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).

[13] M. Curty, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. Lett. **92**, 217903 (2004).

[14] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 160502 (2005).

[15] I. Devetak and A. Winter, Proc. R. Soc. Lond. A **461**, 207 (2005).

[16] M. Christandl, R. König, and R. Renner, Phys. Rev. Lett. **102**, 020504 (2009).

[17] R. Renner and R. König, in *Theory of Cryptography*, edited by J. Kilian (Springer Berlin Heidelberg, 2005), vol. 3378 of *Lecture Notes in Computer Science*, pp. 407–425.

[18] K. Horodecki, D. Leung, H.-K. Lo, and J. Oppenheim, Phys. Rev. Lett. **96**, 070501 (2006).

[19] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim, IEEE Trans. Inf. Theory **54**, 2604 (2008).

[20] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, IEEE Trans. Inf. Theory **55**, 1898 (2009).

[21] K. Horodecki, Ł. Pankowski, M. Horodecki, and P. Horodecki, IEEE Trans. Inf. Theory **54**, 2621 (2008).

[22] M. Piani, Phys. Rev. Lett. **103**, 160504 (2009).

[23] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 200501 (2005).

[24] M. Christandl, N. Schuch, and A. Winter, Commun. Math. Phys. **311**, 397 (2012).

[25] G. Gour and B. C. Sanders, Phys. Rev. Lett. **93**, 260501 (2004).

[26] G. Gour, Phys. Rev. A **71**, 012318 (2005).

[27] G. Gour, Phys. Rev. A **72**, 042318 (2005).

[28] S. Lee, J. S. Kim, and B. C. Sanders, Phys. Lett. A **375**, 411 (2011).

[29] M. Christandl, in *Banff International Research Station (BIRS) Report on Workshop 12w5084: Operator structures in quantum information theory* (2012), p. 5, URL `http://www.birs.ca/`.

[30] S. Bäuml, Diploma thesis, Ludwig Maximilians Universität München, Munich, Germany (2010), URL `http://www.maths.bris.ac.uk/~masmgb/Diploma_thesis.pdf`.

[31] A. Hansen, Master thesis, ETH Zurich (2013), URL `http://www.qit.ethz.ch/paperPDFs/Hansen-Masterarbeit.pdf/`.

[32] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).

[33] M. Christandl and A. Winter, J. Math. Phys. **45**, 829 (2004).

[34] R. Tucci (2002), arXiv:quant-ph/0202144.

[35] K. Horodecki, Ph.D. thesis, University of Warsaw (2008), URL `http://www.mimuw.edu.pl/wiadomosci/aktualnosci/doktoraty/pliki/karol_horodecki/doktorat-kh.pdf`.

[36] M. Fannes, Commun. Math. Phys. **31**, 291 (1973).

[37] M. Christandl, N. Schuch, and A. Winter, Phys. Rev. Lett. **104**, 240405 (2009).

[38] K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor, Phys. Rev. Lett. **87**, 217902 (2001).

[39] M. Christandl and A. Winter, IEEE Trans. Inf. Theory **51**, 3159 (2005).

[40] W. Matthews, S. Wehner, and A. Winter, Commun. Math. Phys. **291**, 3 (2009).

[41] M. J. Donald and M. Horodecki, Phys. Lett. A **264**, 257 (1999).

[42] M. Horodecki, P. Horodecki, and J. Oppenheim (2013), private communication.

[43] P. Horodecki and R. Augusiak, Phys. Rev. A **74**, 010302 (2006).

[44] M. Christandl and R. Pisarczyk (2013), private communication.

[45] K. Li and A. Winter, Commun. Math. Phys. **326**, 63 (2014).

[46] B. Groisman, S. Popescu, and A. Winter, Phys. Rev. A **72**, 032317 (2005).

[47] N. Linden, S. Popescu, B. Schumacher, and M. Westmoreland (1999), arXiv:quant-ph/9912039.

[48] F. G. S. L. Brandão, M. Christandl, and J. Yard, Commun. Math. Phys. **306**, 805 (2011).

[49] R. Renner and S. Wolf, in *Advances in Cryptology, EUROCRYPT 2003*, edited by E. Biham (Springer Berlin Heidelberg, 2003), vol. 2656 of *Lecture Notes in Computer Science*, pp. 562–577, ISBN 978-3-540-14039-9.

[50] M. Christandl, Ph.D. thesis, University of Cambridge (2006), arXiv:quant-ph/0604183.

[51] D. Yang, K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, and W. Song, IEEE Trans. Inf. Theory **55**, 3375 (2009).

[52] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[53] R. Ahlswede and A. Winter, IEEE Trans. Inf. Theory **48**, 569 (2002).

[54] M. Ledoux, *The concentration of measure phenomenon*, vol. 89 of *Mathematical Surveys & Monographs* (American Mathematical Society, 2005).

[55] P. Hayden, D. Leung, P. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004).

[56] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, vol. 222 of *Graduate Texts in Mathematics* (Springer Verlag, 2003).

[57] M. P. Do Carmo, *Riemannian Geometry* (Springer Verlag, 1992).

[58] G. Blower, *Random matrices: high dimensional phenomena*, vol. 367 (Cambridge University Press, 2009).