Distributed private randomness distillation

Dong Yang, 1, 2∗ Karol Horodecki, 3† and Andreas Winter 4‡

1 Laboratory for Quantum Information, China Jiliang University, 310018 Hangzhou, China
2 Department of Informatics, University of Bergen, 5020 Bergen, Norway
3 Institute of Informatics, National Quantum Information Centre, Department of Physics, Mathematics and Informatics, University of Gdańsk, 80-308 Gdańsk, Poland
4 ICREA & Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

We develop the resource theory of private randomness extraction in the distributed and device-dependent scenario. We begin by introducing the notion of independent random bits, which are bipartite states containing ideal private randomness for each party, and motivate the natural set of free operations. As the main tool of our analysis, we introduce Virtual Quantum State Merging, which is essentially the flip side of Quantum State Merging, without communication. We focus on the bipartite case and find the rate regions achievable in different settings. Perhaps surprisingly, it turns out that local noise can boost randomness extraction. As a consequence of our analysis, we resolve a long-standing problem by giving an operational interpretation for the reverse coherent information in terms of the number of private random bits obtained by sending quantum states from one honest party (server) to another one (client) via the eavesdropped quantum channel.

Introduction. – Randomness is no doubt an important notion, having various applications in modern science and technology. Usually only pseudo-randomness is produced in the classical world, for instance by certain complex algorithms in a computer, where the pseudo-random value is determined by a hidden variable so that it is already implicitly known beforehand. Conceptually, the most straightforward way to ensure a uniformly random bit sequence is to generate it by measuring a quantum state, e.g. the σZ eigenstate |0⟩ in the σX eigenbasis |±⟩. In this way the measurement outcome is completely unpredictable in advance, thus private against any eavesdropper. Here the privacy of randomness comes from the fact that a pure state naturally excludes any correlation with other systems, including Eves. The problem of randomness extraction from a general mixed state has been considered in [1] in the spirit of the decoupling approach [2–4], where a single party Alice shares a mixed state ρAE with an eavesdropper Eve and the goal of Alice is to generate randomness private against Eve. An implicit point in this setting is that we have to assume that whoever, calling Bob, holds the purifying system of ρAE, is a trusted but otherwise completely passive party. The reason is that Alice herself cannot figure out the correlation with Eve’s system without Bob’s assistance. So in the spirit of being cautious in cryptography, we have to assume Eve holds all the purification system, i.e. ρAE is pure, unless Alice knows Bob holds part of it.

We make Bob active, where Alice and Bob trust each other and collaborate together to extract independent randomness private against Eve. This is a very new scenario - distributed private randomness extraction. It is almost immediate to see that this scenario is dual to Slepian-Wolf problem on distributed data compression [17] in classical information theory. Surprisingly a natural dual setting to Slepian-Wolf does not exist in classical framework but does in quantum one! In this Letter, we study the distributed and device-dependent scenario for randomness extraction (for alternative approach, so called device-independent scenario, see [5] and references therein). We begin with defining the notion of independent random bits (ibits), in a picture dual to the standard one, as bipartite states that contain ideal private randomness, and justifying the set of allowed operations which do not increase randomness. Then we introduce our main tool, the Virtual Quantum State Merging (VQSM) protocol, which originates from the Quantum State Merging (QSM) protocol and represents the other face of QSM, less noticed in the literature, to study two-sided and one-sided randomness extraction. In the two-sided setting, we obtain the achievable rate regions in various scenarios, including: either free or no communication, and either free or no local noise. Surprisingly, we show that local noise, usually regarded as a useless resource, can extend the rate region. In the one-sided setting, we determine the optimal rates of randomness extraction in two extremal classes, pure entangled states and separable states, and provide a computable upper bound for general states. It follows also that there is no bound randomness, even in the no communication case. Finally we resolve a long-standing problem by giving an operational interpretation for the reverse coherent information (up to a constant term log d_{out}), as the number of ibits obtained by sending quantum state from one honest party to another one via an eavesdropped quantum channel. In the following, we state and discuss the results carefully, while all proofs are found in the appendix [6].

Ibits and RCLOCC. – In the standard approach, a state having one ideal random bit with respect to Eve is of the form ρ_{AE} = \frac{1}{2}(|0⟩⟨0| + |1⟩⟨1|)_{A} \otimes σ_E where σ_E is an arbitrary state of Eve uncorrelated with system A. In the dual picture, where the purifying system is included.
lemma 1 A bipartite quantum state having two independent
random bits private against Eve is of the form
\[ \alpha_{AB} = \frac{1}{2} \sum_{i,j,k,l=0}^{1} |i\rangle_A \otimes |k\rangle_B \otimes U_{i,j,k,l} \sigma_{A'B'} U_{i,j,k,l}^\dagger. \]

It is important to note that getting ibits is equivalent to obtaining randomness secure in terms of the trace distance from a product state - which is proved to be
composable. We ask what kind of operations are allowed for
free in order to transform an input state into the ibit form. Notice that a pure state \(|0\rangle_A\) is a special form of ibit where \(B\) and \(A'B'\) are trivial (dimension 1), so we cannot allow pure states for free. It is safe to assume a closed system paradigm, much like the framework for distillation of the thermodynamical work represented by a pure state. Also it is quite natural to assume that free operations should allow for local unitary transformation and some form of communication. A good candidate is the set of operations designed for quantifying the localizable purity in a quantum state, which are the so-called Closed Local Operations and Classical Communication (CLOCC) [10, 15]. We restrict this class further by disallowing partial trace, which is not needed in our setting, as partial trace may be regarded as the action of putting a system aside. In total, the set of free op-
erations consists of the following two, and their compo-
sitions: (i) local unitary transformations on subsystems;
and (ii) sending a subsystem through the dephasing chan-
nel in a fixed basis from one party to the other, where the
dephasing channel simulates the classical communication
in the closed system paradigm, and the channel
environment goes to Eve. W.l.o.g. the basis of dephasing
is chosen to be the computational one and these opera-
tions are named Restricted CLOCC (RCLOCC).

It seems that our framework for randomness extraction
would be reduced to that of purity since we have a similar
set of operation, and purity can be used to generate ran-
domness. This would indeed be the case if we considered
randomness extraction under global operations, but it is
very different in the distributed setting on which we fo-
cus in this work. Also we note that it is possible to study
randomness extraction by assuming other restrictions on
free operations, e.g. incoherent operations [16].

Having defined the class of operations, we are able to
ask the key question: for a bipartite quantum state \(\rho_{AB}\)
whose purification is in Eve’s hand, how much private
randomness can Alice and Bob obtain against Eve under
RCLOCC? We mainly consider the asymptotic i.i.d. set-
ing which means we care about the rates. It turns out,
however, that another seemingly useless resource –local
noise, can act as a booster in the process. Here local
noise means a maximally mixed state on Alice’s or Bob’s
side, whose purifying part is under Eve's control, or in
other words, Alice or Bob share a maximally entangled
state with Eve. It is clear that from local noise alone, Al-
ice cannot produce randomness unknown to Eve by von
Neumann measurement on her half, because the outcome
is perfectly correlated with Eve. However local noise may
help when combined with other states. An illuminating
example is given by entanglement swapping: Alice shares
one singlet with Bob and another one with Eve, i.e. the
tripartite state is \(|\Phi\rangle_{A_1B} \otimes |\Phi\rangle_{A_2E}\). Here the state \(\frac{1}{2} |\Psi\rangle_{A_2}

is understood as the local noise on Alice’s side. Observe
that in entanglement swapping, the outcome of the Bell
measurement by Alice is completely random against each
of Bob and Eve separately (no communication between
them). Thus we get that in the case of no communication
between Alice and Bob, Alice can obtain 2 bits of ran-
domness against Eve. However without the local noise
\(\frac{1}{2} |\Psi\rangle_{A_2}\), Alice can only get 1 randomness bit unknown to
Eve. So we have several different settings dependent on
whether randomness is located on two sides or one side,
whether local noise is available or not, or whether com-
munication is allowed or not. Before we state our findings
on the rate regions and the protocols to achieve them, we
introduce the main tool employed.

Virtual Quantum State Merging. – Entanglement swap-
ing shows that local noise can play an important role in
the distributed scenario. If local noise is not freely avail-
able, the parties have always trivial options of generating
it, either say Alice putting aside one copy of state (giv-
ing it to Eve), effectively leaving Bob with a pure entan-
gled state with Eve, or –in case they can communicate–
Alice sending one of her systems to Bob through the
dephasing channel. However, there is a more efficient
way, which results in some private randomness for Al-
ice and at the same time gives Bob local noise. To gain
intuition, let us look at the protocol of Quantum State
Merging (QSM). For our goal, B is the reference system
and QSM is performed from A to E, i.e. the task is to
transform a tripartite state \(|\psi\rangle_{ABE}\) into \(|\phi\rangle_{A_1A_2BE'E_2}\)
by LOCC such that \(\phi_{E'E_2} \approx \psi_{ABE}\), where \(\phi_{E,E'} = T_{A_1A_2E_2} \phi_{A_1A_2BE'E_2}\). In the asymptotic i.i.d. case,
if \(S(A|E)_{\psi} = S(\psi_{AE}) - S(\psi_E) > 0\), the protocol re-
quires an additional rate of \(S(A|E)_{\psi}\) ibits shared be-
tween systems A and E. However, when \(S(A|E)_{\psi} < 0\),
not only does it not need entanglement, but also creates
a rate of \(-S(A|E)_{\psi}\) ibits on \(A_2B_2\). In the protocol, Al-
ice transforms A into \(A_1A_2\) by a unitary, then measures
subsystem \(A_1\) in the computational basis and announces
the outcome to Eve, who further, according to the out-
comes, transforms \(E\) into \(E'E_1E_2\) by proper unitaries.
The amount of classical communication is the size of $A_1$, its rate is the mutual information between Alice and Bob, $I(A : B)_\psi = S(\psi_A) + S(\psi_B) - S(\psi_{AB})$.

Bob’s state remains invariant, so the measurement outcome on $A_1$ is independent of Bob. Meanwhile, Bob’s system is purely entangled with Eve’s, conditional on the measurement outcomes. A key observation is that in our setting, Bob need not care about whether Eve performs the rotations or not. It is sufficient for him to know that the purifying system of his state is at Eve’s side. Furthermore, the outcome of the measurement of $A_1$ is almost decoupled from Eve. This weak correlation can be deleted by the technique of privacy amplification (PA) [14], that works for cq-states; it amounts to tracing out a little bit more from $A_1$, not however affecting the rate of randomness. In this way, Alice can extract randomness independent of Bob and private against Eve at the rate of $I(A : B)_\psi$, while at the same time Bob’s system is (virtually) only entangled with Eve, which acts as noise that can help Bob extract randomness later. The composition of QSM and PA we call Virtual Quantum State Merging. The version we shall need is however much simpler than this, and encapsulated in a doubly decoupling theorem, Theorem 13 in the appendix, based on which we can prove the following theorem.

**Theorem 2** Given $\epsilon, \delta > 0$ and sufficiently large numbers $n, m$, for a state $|\psi\rangle_{ABE}$ satisfying $S(A|B) > 0 > S(B|A)$, there exist local unitaries $U : A^n \to K_A A'$ and $V : B^{n+m} \to K_B B'$ and fixed bases $\{|i\rangle_{K_A}\}$ of $K_A$ and $\{|j\rangle_{K_B}\}$ of $K_B$, $U \otimes V |\psi\rangle^{(n+m)}_{ABE} \equiv \sum_{i,j=1}^{(|K_A|,|K_B|)} \sqrt{p_{ij}} |i\rangle_{K_A} |j\rangle_{K_B} |\psi_{ij}\rangle_{A'B'E^nE^m}$, such that after measurements on $K_A$, $K_B$ in the fixed bases, $\| \sum_{i,j} p_{ij} |i\rangle_{K_A} |j\rangle_{K_B} \otimes |\psi^{ij}_{E^nE^m}\rangle - \frac{I_A}{|K_A|} \otimes \frac{I_B}{|K_B|} \otimes |\psi^{(n+m)}_{E^nE^m}\rangle \|_1 \leq \epsilon$,

when $\log |K_A| = nl(A : B) - n\delta$ and $\log |K_B| = \min\{mI(B : A), mI(B : E) + 2nS(B)\} - m\delta$.

*Two-sided randomness distillation.* Having developed the tool, we are to state the result. It is clear that given a state $\rho_{AB}$, we can obtain the randomness at the rate $R_G(\rho_{AB}) := \log |AB| - S(\rho_{AB})$ if global unitary operations on $AB$ are allowed [8] Lemma 14. Now we ask the same question if the parties are distributed so only local unitary operations are allowed: What is the rate region $(R_A, R_B)$, representing that Alice produces randomness at rate $R_A$ and Bob at rate $R_B$, and their randomness is independent and secret against Eve, who has the purifying system? This is the same as asking for approximate iibs by Lemma 4. We have four different settings in which we allow free or no local noise, and free or no communication. The answer is our first main result, Theorem 3 by composing Theorem 2.

**Theorem 3** The achievable rate regions of $\rho_{AB}$ are:

1) for no communication and no noise, $R_A \leq \log |A| - S(A|B)_+, R_B \leq \log |B| - S(B|A)_+$, and $R_A + R_B \leq R_G$,

2) for free noise but no communication, $R_A \leq \log |A| - S(A|B)\_B$, $R_B \leq \log |B| - S(B|A)_+$, and $R_A + R_B \leq R_G$;

3) for free noise and free communication, $R_A \leq R_G$, $R_B \leq R_G$, and $R_A + R_B \leq R_G$;

4) for free communication but no noise, $R_A \leq \log |A| - \max\{S(B), S(AB)\}$, $R_B \leq \log |B| - \max\{S(A), S(AB)\}$, and $R_A + R_B \leq R_G$.

Further the rate regions in settings 1), 2), 3) are tight. Note how the solutions to 1) and 2) appears to be dual to Slepian-Wolf theorem on distributed data compression [17]. In 3) the entire global purity can be realised on either side as randomness (but not necessarily as purity). We have only achievability in 4) and leave its tightness open.

*One-sided randomness distillation.* We try to understand the achievable rates a bit better for setting 4). The task is that Bob helps Alice to extract randomness against Eve as much as possible. Free communication is allowed but local noise is not, otherwise the setting is understood already. First we derive a formula for the optimal rate. Unfortunately it involves regularization.

**Theorem 4** The randomness that can be extracted from $\rho_{AB}$ on Alice’s side is $R_A = \log |AB| -
inf $\frac{1}{n} \max \{S(E^{(n)}), S(B^{(n)})\}$, where the infimum is taken over all $n$, and $E^{(n)}$ and $B^{(n)}$ are the output systems under RCLOCC acting on $\rho_{AB}$.

At first sight, this theorem is not very useful as it involves possibly complex RCLOCC processing and suffers from the notorious regularization problem. But a further observation that $S(E^{(n)}) \geq nS(E)$ because of the non-decreasing entropy of Eve in every use of the dephasing channel gives the exact optimal rate for a large class of states. Namely, for $\rho_{AB}$ satisfying $S(B) \leq S(E)$, $R_A = \log |AB| - S(AB)$, so all the randomness can be localized on Alice’s side. This is not so strange as it already appears in the rate region of two-sided setting. A large class of bipartite states satisfy this property, for instance the positive partial transpose (PPT) states which includes separable states. Also note that two-way communication is needed to localize the global purity on Alice’s side, but only one-way communication is needed to localize randomness. Such an example is the $\text{cq}$ state $\rho = \frac{1}{2} |00\rangle \langle 00| + |11\rangle \langle 11|$. The nontrivial $R_A(\Phi_{AB}) = \frac{3}{2}$ is Corollary 6 and the other two are straightforward. Before that, we derive an upper bound for a general state that only depends on the initial state.

**Theorem 5** The optimal randomness that can be extracted from $\rho_{AB}$ on Alice’s side is upper bounded as $R_A \leq \log |AB| - \max \{\frac{1}{2} E_\infty(\rho_{AB}) + S(AB), S(AB)\}$.

Here $E_r(\rho) = \inf_{\sigma \in \text{SEP}} \text{Tr} \rho (\log \rho - \log \sigma)$ and $E_\infty(\rho) = \lim_{n \to \infty} \frac{1}{n} E_r(\rho^{\otimes n})$ are the relative entropy of entanglement [13] and its regularisation. From $E_\infty(\rho_{AB}) \geq \max \{S(A)-S(AB), S(B)-S(AB)\}$ [19], we get $E_\infty(\rho_{AB}) + S(AB) \geq \max \{S(A), S(B)\}$. So a computable upper bound is $\log |AB| - \frac{1}{2} \max \{S(A), S(B)\}$. The corollary below shows the tightness for pure states.

**Corollary 6** The optimal randomness on Alice’s side that can be extracted from a singlet is $\frac{3}{4}$.

Corollary 6 can be generalised to a general pure state by noticing that local purity can be freely transmitted. Hence, for $|\phi\rangle_{AB}$, $R_A = \log |AB| - \frac{1}{2} S(A)_\phi$. Further, recall that for $|\phi\rangle_{AB}$, the distillable key $K_D = S(A)_\phi$ [11]. When $|A| = |B| = d$, we have an appealing formula exhibiting the exact balance between localisable and shareable privacy $R_A + \frac{1}{2} K_D = 2 \log d$. For mixed states, from the relation $K_D \leq E_\infty(\rho)$, we get $R_A + \frac{1}{2} K_D \leq 2 \log d$, which can be treated as complementarity between private randomness and key in analogy to complementarity between purity and entanglement [20]. It captures the fact that $R_A$ attains $2 \log d$ on a pure product state while $K_D$ attains maximal value on a maximally entangled state.

**Private randomness capacity.** It is well-known that channel capacity formulas usually involve regularization [21], i.e. optimisation over growing numbers of channel uses due to non-additivity of the relevant quantities [22][25], making a head-on numerical approach impossible. However this does not prevent the regularized non-additive quantities from having good interpretations: the optimal rate to transmit information faithfully through a channel. Ironically, there are additive quantities that however lack an interpretation [24], one of these being the reverse coherent information, whose additivity is well-known [27], but whose interpretation was missing for a long time. The quantity was first introduced in [27] as the negative cb-entropy of a channel. In [28], it is called reverse coherent information and shown to be a lower bound for the entanglement distribution capacity of a channel assisted by classical feedback communication. Here we provide its exact operational interpretation.

Consider the task of generating private randomness by communicating through a quantum channel. A client Bob who wants to produce randomness private against Eve, has the measurement device but cannot prepare a quantum state himself. He has access to, but does not trust quantum particles around himself, which is potentially entangled with Eve’s system. However there is a quantum channel to Bob from a trusted server Alice who has the power to prepare any quantum state, although the channel itself is eavesdropped by Eve. Before stating the result, let us argue that the model is well-motivated and that it captures the server-client structure possibly realised in future quantum networks: the server having huge devices and able to prepare and manipulate quantum states, a client only able to do limited operations, e.g. units and measurement. The server provides a service to the client via a quantum channel. Second, this new cryptographic model is completely in line with the standard model of transmitting information (bit, qubit, pbit), and thus can be regarded as a new character of a quantum channel. Now the natural capacity question is: “What is the maximal rate of private randomness that can be extracted at Bob’s side when Alice sends quantum states through the channel?” Amazingly, we can answer it completely.

**Theorem 7** The private randomness capacity of a channel $\mathcal{N} : A \rightarrow B$ is given by $R(\mathcal{N}) = \log |B| + \max_{\phi \in \text{SEP}} S(\phi_A) - S(\text{id}_A \otimes \mathcal{N} (\phi_{AA}))$.

The second term is just the reverse coherent information of the channel. So for the first time we provide an operational interpretation for it in a natural Shannon theoretic model. As a matter of fact, the randomness generated at the client is not only private to the eavesdropper but also to the server. The single-letter formula
is concave with respect to the input state, thus efficiently computable. We can draw an interesting comparison between private randomness capacity and purity capacity of a channel: Pure(\mathcal{N}) = \log |\mathcal{B}| - \inf_\phi \frac{1}{2} S_{\min}(\mathcal{N}^{\otimes n}), where S_{\min}(\mathcal{N}) = \min_\phi S(\mathcal{N}(\phi)) is the minimum output entropy. From \[22\], we know that S_{\min}(\mathcal{N}) is not additive thus regularisation is required.

**Summary and outlook.** We have initiated the study of randomness extraction in distributed scenario and provide the tool VQSM to deal with this situation. We found the exact achievable rate regions for various settings of distillation protocols and bipartite states, and gave the long-sought operational interpretation of the reverse coherent information of a quantum channel.

This work opens up a new area, suggesting a wide range of generalisations and extension of the considered scenarios, including the multipartite case, the one-shot scenario based on one-shot QSM \[29\], as well as other cryptographic variants such as the honest-but-curious scenario, in which Alice and Bob need to keep their private randomness in addition secret from each other. The structural analogy between private states and the independent states needs also further exploration. These directions will be developed elsewhere \[30\]. Two notable topics are under study. One is distributed randomness extraction under the framework of coherence theory \[31–33\] where the allowed operations are incoherent, and the other is the strong converse problem for the private randomness capacity. Indeed there are rare cases \[34, 35\] where the allowed operations are coherent and the strong converse theorem is proved only for such cases \[36, 37\]. We leave several interesting questions deserving further investigation, highlighting two: 1) Is the rate region of the two-sided randomness distillation tight in the case of free communication but no noise? We conjecture so, but the proof eludes us so far. 2) Can the formula in Theorem 3 be reduced to a single-letter one? In other words, is the single-letter quantity weakly additive?

**Acknowledgments.** The authors thank Omer Sakarya for confirming the direct relation between secret sharing schemes and the structure of independent states, and Mark Wilde for pointing out the reference \[27\].

DY was supported by the NSFC (grant nos 11375165, 11875244), and by the NFR Project ES546777. KH was supported by the grant Sonata Bis 5 (grant number 2015/18/E/ST2/00327) from the National Science Center. AW was supported by the European Research Council (Advanced Grant IRQUAT), the European Commission (STREP project RAQUEL), the Spanish MINECO, projects FIS2013-40627-P and FIS2016-80681-P, as well as by the Generalitat de Catalunya, CIRIT projects 2014-SGR-966 and 2017-SGR-1127.

---

\[1\] M. Berta, O. Fawzi, and S. Wehner, *IEEE Trans. Inf. Theory* **60**, 1168 (2014).

\[2\] M. Horodecki, J. Oppenheim, and A. Winter, *Nature* **436**, 673 (2005); *Comm. Math. Phys.* **269**, 107 (2007).

\[3\] F. Dupuis, PhD thesis, arXiv[quant-ph]:1004.1641.

\[4\] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, *Comm. Math. Phys.* **328**, 251 (2014).

\[5\] M.N. Bera, A. Acín, M. Kus, M. Mitchell, M. Lewenstein, *Rep. Prog. Phys.* **80**, 124001 (2017).

\[6\] See Supplemental Material at URL for details, which includes additional references \[23, 30\].

\[7\] A. Uhlmann, *Rep. Math. Phys.* **9**, 273 (1976).

\[8\] C. A. Fuchs and J. van de Graaf, *IEEE Trans. Inf. Theory* **45**, 1216-1227 (1999); arXiv:quant-ph/9712042 (1997).

\[9\] M.-Y. Ye, Y.-K. Bai, and Z. D. Wang, *Phys. Rev. A* **78**, 030302 (2008).

\[10\] M. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), U. Sen, and B. Synak-Radtke, *Phys. Rev. A* **71**, 062307 (2005).

\[11\] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, *Phys. Rev. Lett.* **94**, 160502 (2005); *IEEE Trans. Inf. Theory* **55**, 1898 (2009).

\[12\] M. Christandl, PhD thesis, arXiv:quant-ph/0604183.

\[13\] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, arXiv:quant-ph/0409078.

\[14\] R. Renner and R. Koenig, in: Proc. TCC 2005, LNCS 3378, Springer Verlag (2005).

\[15\] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **89**, 180402 (2002).

\[16\] M. Hayashi and H. Zhu, *Phys. Rev. A* **97**, 012302 (2018).

\[17\] D. Slepian and J. K. Wolf, *IEEE Trans. Inf. Theory* **19**, 471 (1973).

\[18\] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, *Phys. Rev. Lett.* **78**, 2275 (1997).

\[19\] I. Devetak and A. Winter, *Proc. R. Soc. Lond. A* **461**, 207 (2005).

\[20\] J. Oppenheim, K. Horodecki, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. A* **68**, 022303 (2003); arXiv:quant-ph/0207025.

\[21\] G. Smith, arXiv:quant-ph/0107255.

\[22\] M. B. Hastings, *Nature Physics* **5**, 255-257 (2009).

\[23\] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, *Phys. Rev. A* **57**(2), 830-839 (1998).

\[24\] K. Li, A. Winter, X. Zou, and G. Guo, *Phys. Rev. Lett.* **103**, 120501 (2009).

\[25\] G. Smith and J. A. Smolin, *Phys. Rev. Lett.* **103**, 120503 (2009).

\[26\] A. W. Cross, K. Li, and G. Smith, *Phys. Rev. Lett.* **118**, 040501 (2017).

\[27\] I. Devetak, M. Junge, C. King, M. B. Ruskai, and Mark Wilde for pointing out the reference \[27\].

\[28\] M. Berta, Diploma thesis, arXiv[quant-ph]:0912.4495.

\[29\] K. Horodecki, O. Sakarya, M. Winczewski, A. Winter, and D. Yang, in preparation (2018).

\[30\] J. A. Æberg, arXiv:quant-ph/0612140.
SUPPLEMENTAL MATERIAL

A. Miscellaneous facts and lemmas

In this appendix, we collect some standard facts about various functionals we use and prove some lemmas that we need in the proofs of main results.

Recall that the fidelity between two mixed states is defined as

\[ F(\rho, \sigma) := \|\sqrt{\rho} \sigma \sqrt{\rho}\|_1 = \text{Tr} \sqrt{\rho \sigma} \sqrt{\rho}, \]

and the trace distance is \( \frac{1}{2} \|\rho - \sigma\|_1 \).

Lemma 8 (Uhlmann [7]) The fidelity is alternatively characterized by the relation

\[ F(\rho_A, \sigma_A) = \max_{U_B} \left| \langle \psi | U_A \rho_A U_A^\dagger \otimes \sigma_B | \psi \rangle \right|, \]

where |\psi\rangle_{AB} and |\psi\rangle_{AB}^\prime are purifications of \( \rho_A \) and \( \sigma_A \) respectively, and \( U_B \) is unitary.

Lemma 9 (Fuchs and van de Graaf [8]) For two states \( \rho \) and \( \sigma \), fidelity and trace distance are related by

\[ 1 - F(\rho, \sigma) \leq \frac{1}{2} \|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)^2}. \]

Lemma 10 (Dupuis et al. [4]) Given a state \( \rho_{AB} \), let \( \epsilon > 0 \) and \( T_{A\rightarrow B} \) be a CPTP map with Choi-Jamiołkowski representation \( \tau_{RB} = (\text{id}_R \otimes T) \Phi_{RA} \). Then we have,

\[ \int_{U_A} \|T(U_A \rho_{AB} U_A^\dagger) - \tau_B \otimes \rho_E\|_1 dU \leq 2^{-\frac{1}{2}H_{\min}(A|E) + H_{\min}(B|R)} + 12\epsilon, \]

where the integral is over the Haar measure on unitary group \( U(A) \).

Lemma 11 If states in an ensemble \( \{p_i, \rho_i\} \) are close to a fixed state \( \sigma \) on average, then most of the ensemble states are close to the fixed state. To be precise, if

\[ \sum_i p_i \|\rho_i - \sigma\|_1 \leq \epsilon \]

then,

\[ \sum_{i: \|\rho_i - \sigma\|_1 \leq \sqrt{\epsilon}} p_i \geq 1 - \sqrt{\epsilon} \]

Proof. It is obvious that

\[ \sqrt{\epsilon} \sum_{i: \|\rho_i - \sigma\|_1 \geq \sqrt{\epsilon}} p_i \leq \sum_{i: \|\rho_i - \sigma\|_1 \geq \sqrt{\epsilon}} p_i \|\rho_i - \sigma\|_1 \leq \epsilon, \]

so

\[ \sum_{i: \|\rho_i - \sigma\|_1 \geq \sqrt{\epsilon}} p_i \leq \sqrt{\epsilon}. \]

Lemma 12 (Horodecki et al. [10]) The function defined by \( g(\rho_{AB}) := S(\rho_{AB}) + E_r(\rho_{AB}) \) is non-decreasing under CLOCC.

Here,

\[ E_r(\rho_{AB}) := \min_{\sigma_{AB} \in \text{SEP}} S(\rho_{AB}\|\sigma_{AB}) \]

is the relative entropy of entanglement where the minimum is taken over the separable states (SEP), and \( S(X|Y) = \text{Tr} (X \log X - Y \log Y) \) is the relative entropy.

B. Proofs

In this appendix, we provide the detailed proofs of the results claimed in the main text.

Proof of Lemma 1. In the usual approach, a quantum state having two independent random bits private against Eve is of the form

\[ \rho_{ABE} = \frac{1}{4} \sum_{i,j=0} |i\rangle_A \otimes |j\rangle_B \otimes |\phi_{ij}\rangle_{AB}. \]

Consider a pure state \( |\phi\rangle_{AA'BB'E} \) whose marginal state \( \phi_{ABE} \) turns into \( \rho_{ABE} \) after measuring subsystems A and B in the computational basis. We denote the dephasing map in the computational basis by \( M \), so the dephased state of \( \phi_{ABE} \) is \( M\phi_{ABE} := \sum_{ij} |i\rangle_A \otimes |j\rangle_B \otimes (T_j |\phi_{ij}\rangle_{AB}). \) When written in the computational basis, the whole state has the form

\[ |\phi\rangle_{AA'BB'E} = \frac{1}{2} \sum_{i,j=0} |i\rangle_A \otimes |j\rangle_B \otimes |\phi_{ij}\rangle_{AB}. \]

The dephased state is

\[ \rho_{AA'BB'E} = \frac{1}{2} \sum_{i,j=0} |i\rangle_A \otimes |j\rangle_B \otimes |\phi_{ij}\rangle_A \otimes \phi_{ij}' \langle A' B' E| \]

\[ \rho_{AA'BB'E} = \frac{1}{2} \sum_{i,j=0} |i\rangle_A \otimes |j\rangle_B \otimes |\phi_{ij}\rangle_A \otimes |\phi_{ij}'\rangle_{AB}. \]
Because $\rho_{ABE} = \text{Tr}_{A'B'} \rho_{AA'B'B'E}$ is of the form Eq. 6, $|\phi_{ij}\rangle_{A'B'E}$ satisfy the condition $\text{Tr}_{A'B'} |\phi_{ij}\rangle_{A'B'E} = \rho_E$ for all $i$ and $j$. Then, by the Schmidt decomposition, there exist unitary operators $U_{ij}$ acting on $A'B'$ such that the following holds

$$|\phi_{ij}\rangle_{A'B'E} = U_{ij} |\phi_0\rangle_{A'B'E},$$  

(9)

where $|\phi_0\rangle_{A'B'E}$ is a fixed purification for $\rho_E$. Using this form and tracing the E part in Eq. 7, we get the state $\alpha_{AA'B'B'}$ of the claimed structure. A similar reasoning gives the state $\alpha_{AA'B'}$ when random bit is only on one side.

The robust version of Lemma 1 is also true in the following sense, which we state directly for general amounts of randomness for Alice and Bob: Let, for positive integers $a$ and $b$,

$$\rho_{AB} = \frac{1}{ab} \sum_{i=0}^{a-1} \sum_{j=0}^{b-1} |i\rangle_1 \otimes |j\rangle_2 \otimes \rho_E,$$

(10)

$$\alpha_{AA'B'B'} = \frac{1}{ab} \sum_{i,j}^{a-1} \sum_{k,l}^{b-1} |i\rangle_2 \otimes |j\rangle_2 \otimes \rho_{AB} U_{ij} \sigma_{AB} U_{ij}^\dagger,$$

(11)

the ideal private randomness state and the ideal idt, respectively. They have the common pure state

$$|\phi\rangle_{AA'B'B'} = \frac{1}{\sqrt{ab}} \sum_{i=0}^{a-1} \sum_{j=0}^{b-1} |i\rangle_2 \otimes |j\rangle_2 \otimes U_{ij} \sigma_{AB} U_{ij}^\dagger |\phi_0\rangle_{A'B'E},$$

(12)

with a suitable purification $|\phi_0\rangle$ of $\rho_E$, by which we mean $M_{\phi_{ABE}} = \rho_{ABE}$, and $|\phi\rangle_{AA'B'B'}$ is a purification of $\alpha_{AA'B'B'}$.

If a ccq state $\tilde{\rho}_{ABE}$ is such that

$$\frac{1}{2} \| \rho_{ABE} - \rho_{ABE} \|_1 \leq \epsilon,$$

then for any pure state $|\tilde{\phi}\rangle_{AA'B'B'E}$ satisfying $M_{\tilde{\phi}_{ABE}} = \tilde{\rho}_{ABE}$, $\frac{1}{2} \| \rho_{ABE} - \tilde{\rho}_{ABE} \|_1 \leq \delta = \sqrt{f(\epsilon)(2 - f(\epsilon))}$ for $f(\epsilon) \to 0$ as $\epsilon \to 0$, from an $\alpha$ of the form 11 with suitable $\sigma_{AB}$, and unitaries $U_{ik}$ acting on $A'B'$.

Conversely, if $\tilde{\alpha}_{AA'B'}$ is such that

$$\frac{1}{2} \| \tilde{\alpha}_{AA'B'} - \alpha_{AA'B} \|_1 \leq \epsilon,$$

then for any purification $|\tilde{\phi}\rangle_{AA'B'E}$ of $\tilde{\alpha}_{AA'B'}$, $\frac{1}{2} \| \tilde{\rho}_{ABE} - \rho_{ABE} \|_1 \leq \delta = \sqrt{f(2 - \epsilon)}$, for a $\tilde{\rho}$ of the form 10 with a suitable $\rho_E$.

For the forward direction, suppose the ccq state $\tilde{\rho}_{ABE}$ is such that $\frac{1}{2} \| \rho_{ABE} - \rho_{ABE} \|_1 \leq \epsilon$. We have a pure state $|\tilde{\phi}\rangle_{AA'B'B'E} = \sum_{ij} \sqrt{p_{ij}} |i\rangle_2 \otimes |j\rangle_2 \otimes \tilde{\rho}_{AB} E$ such that $M_{\tilde{\phi}_{ABE}} = \tilde{\rho}_{ABE}$, where $|\tilde{\phi}_{ij}\rangle_{A'B'E}$ are the purifications of $\tilde{\rho}_{ij}$. Denote a fixed purification of $\rho_E$ as $|\phi_0\rangle_{A'B'E}$. From Uhlmann’s theorem 7, we can choose unitaries $U_{ij}$ on $A'B'$ satisfying $|\tilde{\phi}_{ij}|U_{ij}\phi_0 = F(\tilde{\rho}_{ij}, \rho_E)$ and form the state $|\phi\rangle_{AA'B'B'E}$ in eq. (12) which is clearly the purification for some $\alpha_{AA'B'B'}$. We claim that $|\phi\rangle_{AA'B'B'E}$ is close to $|\phi\rangle_{AA'B'B'E}$ that is sufficient to prove $\phi_{AA'B'B'}$ is close to $\alpha_{AA'B'B'}$.

By the triangle inequality of trace norm and its monotonicity under partial trace, we get

$$\sum_{ij} p_{ij} \| \tilde{\rho}_{ij} - \rho_E \|_1 \leq \left( \sum_{ij} p_{ij} \right) \| |\tilde{\phi}_{ij}|U_{ij}\phi_0\|_1 \leq \left( \sum_{ij} p_{ij} \right) \| \tilde{\rho}_{ij} - \rho_E \|_1 \leq 4\epsilon,$$

from Lemma 1 there exists a good index set $K = \{ ij : \| \tilde{\rho}_{ij} - \rho_E \|_1 \leq 4\epsilon \}$ and $\sum_{ij \in K} p_{ij} \geq 1 - 4\epsilon$. From Lemma 2 we have $|\langle \tilde{\phi}_{ij}|U_{ij}\phi_0| \geq 1 - \sqrt{\epsilon}$ and $\sum_{ij \in K} \sqrt{p_{ij}} \| \tilde{\rho}_{ij} - \rho_E \|_1 \geq 1 - \epsilon$. Then we compute the fidelity,

$$|\langle \tilde{\phi}_{ij}|U_{ij}\phi_0| \geq 1 - \sqrt{\epsilon},$$

(13)

$$\sqrt{\rho_E} \leq \sqrt{\rho_E},$$

(14)

$$\geq (1 - \sqrt{\epsilon}) \sum_{ij \in K} \sqrt{p_{ij}} \sqrt{\frac{1}{ab}},$$

(15)

$$= (1 - \sqrt{\epsilon})(\sum_{ij \in K} \rho_{ij} + \sum_{ij \in K} \rho_{ij}) \sqrt{\frac{1}{ab}},$$

(16)

$$\geq (1 - \sqrt{\epsilon})(1 - \epsilon) \frac{1}{2} \sum_{ij \in K} (p_{ij} + \frac{1}{ab}),$$

(17)

$$\geq (1 - \sqrt{\epsilon})[(1 - \epsilon) - \frac{1}{2} \sum_{ij \in K} (p_{ij} - \frac{1}{ab})],$$

(18)

$$\geq (1 - \sqrt{\epsilon})(1 - \epsilon) - \frac{1}{2} \sum_{ij \in K} (p_{ij} - \frac{1}{ab} - 2p_{ij}),$$

(19)

$$= (1 - \sqrt{\epsilon})(1 - 2\epsilon - 2\sqrt{\epsilon}) = f(\epsilon),$$

(20)

where Ineq. 15 comes from $|\langle \tilde{\phi}_{ij}|U_{ij}\phi_0| \geq 0$ and $|\langle \tilde{\phi}_{ij}|U_{ij}\phi_0| \geq 1 - \sqrt{\epsilon}$ for $ij \in K$, 17 from $\sqrt{\frac{1}{ab}} \leq \sqrt{x + y}$, 18 from $x - y \leq |x - y|$, and 19 from $\sum_{ij \in K} |p_{ij} - \frac{1}{ab}| \leq \frac{1}{2} \sum_{ij \in K} |p_{ij} - \frac{1}{ab}| \leq \epsilon$ and $\sum_{ij \in K} p_{ij} \geq 1 - \sqrt{4\epsilon}$. Then we get, by monotonicity of the fidelity, that

$$F(\tilde{\phi}_{AA'B'B'}, \alpha_{AA'B'B'}) \geq 1 - f(\epsilon),$$


and so

\[ \frac{1}{2} \| \tilde{\alpha}_{AAB'} - \alpha_{AAB'} \|_1 \leq \sqrt{f(\epsilon)(2 - f(\epsilon))} =: \delta. \]

For the opposite direction, we proceed in a simpler way, namely by Uhlmann’s theorem we have that there exists a unitary \( V_E \) such that

\[ \langle \tilde{\phi} \otimes I \otimes V | \phi \rangle = F(\tilde{\alpha}_{AAB'}, \alpha_{AAB'}) \geq 1 - \epsilon. \]

Then we get, by monotonicity of the fidelity, that

\[ F(M \tilde{\phi}_{ABE}, M(V \phi_{ABE V}^\dagger)) \geq 1 - \epsilon, \]

and so

\[ \frac{1}{2} \| M \tilde{\phi}_{ABE} - M(V \phi_{ABE V}^\dagger) \|_1 \leq \sqrt{\epsilon(2 - \epsilon)} =: \delta. \]

As before, \( M(V \phi_{ABE V}^\dagger) \) is also of the structure of \( \rho_{ABE} \).

\[ \square \]

**Theorem 13** Given \( \epsilon, \delta > 0 \), for all sufficiently large numbers \( n \) of copies of a pure tripartite state \( |\psi\rangle_{ABE} \), there exists a unitary \( U : A^n \rightarrow K'A' \) with a fixed standard basis \( \{|i\rangle\} \) of subsystem \( K \), such that

\[
\left\| \sum_{i=1}^{K^n} \psi(i|K) \otimes \psi_B^n \right\|_1 \leq \epsilon,
\]

\[
\left\| \sum_{i=1}^{K^n} \psi(i|K) \otimes \psi_E^n \right\|_1 \leq \epsilon,
\]

\[
(U_{A^n} \otimes I_{B^nE^n})|\psi\rangle_{ABE}^\otimes = \sum_{i=1}^{K^n} \sqrt{\psi(i|K)} |\psi(i)\rangle_{A'B^nE^n},
\]

when \( \frac{1}{n} \log |K| = \min \{ I(A : B)_\psi, I(A : E)_\psi \} - \delta \).

**Proof.** First we prove a one-shot version of Theorem 13. Namely, for a pure tripartite state \( |\psi\rangle_{ABE} \), there exists a unitary \( U : A \rightarrow KA' \) with a fixed standard basis \( \{|i\rangle\} \) of subsystem \( K \), such that the state of \( K \) after the von Neumann measurement simultaneously decouples from Bob and Eve, as follows:

\[
\left\| \sum_{i=1}^{K^n} \psi(i|K) \otimes \psi_B^n \right\|_1 \leq 2 \left( 2^{\frac{1}{n} \log |K| A\rightarrow|K|} + 2 \epsilon \right),
\]

\[
\left\| \sum_{i=1}^{K^n} \psi(i|K) \otimes \psi_E^n \right\|_1 \leq 2 \left( 2^{\frac{1}{n} \log |K| A\rightarrow|K|} + 2 \epsilon \right),
\]

where \( (U_A \otimes I_{BE})|\psi\rangle_{ABE} = \sum_{i=1}^{K^n} \sqrt{\psi(i|K)} |\psi(i)\rangle_{A'B^nE^n} \) and \( \tau_{KK} = (id_R \otimes T) \Phi_{RA} \) is the Choi-Jamiolkowski state of the CPTP map

\[ T_{A \rightarrow K} (\rho) = \sum_i |i\rangle\langle i| (Tr_{A'} \rho) |i\rangle\langle i|, \]

with \( A = A'K' \), that is tracing \( A' \) first and then performing a von Neumann measurement on \( K \). Notice that \( \tau_{KK} = \frac{I_A}{|K|} \). Now borrowing the idea from \[9\] and Lemma \[10\] it is easy to see that more than one half unitary operators \( U \) satisfy Ineq. \[21\] and more than one half satisfy Ineq. \[22\]. So there must exist a common unitary such that both inequalities are satisfied.

Now, by the asymptotic equipartition property, the \( \epsilon \)-smooth conditional min-entropy reduces to conditional entropy in the asymptotic i.i.d. setting, and \( S(A|E)_{\psi_{AB^n}} = n(S(B) - S(E)) \), \( S(A|B)_{\psi_{AB^n}} = n(S(E) - S(B)) \), as well as \( S(R|K)_\psi = nS(A) - \log |K| \) when we restrict to the typical subspace of \( \psi_{AB} \). Counting the rates, we get that if \( \frac{1}{n} \log |K| = \min \{ I(A : B), I(A : E) \} - \delta \), then the first term would decay exponentially with \( n \) so we can get randomness in \( K \) simultaneously independent of Bob’s state and Eve’s state.

**Remark:** Theorem 13 can be generalized to multipartite states.

**Proof of Theorem 2.** Essentially Theorem 2 says that Theorem 13 is composable in the following way. Suppose \( n \) copies of \( |\psi\rangle_{ABE} \) and \( m \) copies of \( |\psi\rangle_{ABE} \) where \( S(A|E) < 0 < S(B|E) \). From Theorem 13 we know that Alice can get randomness \( nI(A : B)_\psi \) in \( K_A \) from \( |\psi\rangle_{ABE} \). Notice that Bob could virtually regard his state \( \psi_B \) as local noise in the sense that it is entangled with \( A'E' \) even assuming \( A' \) is at Eve’s control which is the worst case (but of course Eve does not really have access to \( A' \) whose state may have information on the randomness). And he can make use of this noise to help extract randomness with the fresh state \( |\psi_{ABE} \rangle \) where the randomness in subsystem \( K_B \) would decouple the system \( A'E' \). So in all randomness in \( K_A \) and \( K_B \) is independent and secure to Eve.

The first step is for Alice to extract randomness from \( |\psi_{ABE} \rangle \). From Theorem 13, we know that there exists \( U : A^n \rightarrow KA' \) such that

\[
U |\psi\rangle_{ABE}^\otimes = \sum_{i=1}^{K^n} \sqrt{\psi(i|K)} |\psi(i)\rangle_{A'B^nE^n},
\]

satisfying

\[
\left\| \sum_{i=1}^{K^n} \psi(i|K) \otimes \psi_B^n \right\|_1 \leq \epsilon,
\]

\[
\left\| \sum_{i=1}^{K^n} \psi(i|K) \otimes \psi_E^n \right\|_1 \leq \epsilon,
\]
when \( \log |K_A| = n[I(A : B) - \delta] \). By the monotonicity of trace norm, we have
\[
\frac{1}{|K_A|} \sum_{i=1}^{|K_A|} p_i |i\rangle |i\rangle_{K_A} \approx \frac{1}{|K_A|} \sum_{i=1}^{|K_A|} |i\rangle |i\rangle_{K_A},
\]
where \( X \approx Y \) means \( \|X - Y\| \leq \epsilon \). Then by the triangle inequality of trace norm, we get
\[
\frac{1}{|K_A|} \sum_{i=1}^{|K_A|} |i\rangle |i\rangle_{K_A} \otimes \psi_B^n \approx \frac{1}{|K_A|} \sum_{i=1}^{|K_A|} |i\rangle |i\rangle_{K_A} \otimes \psi_B^n.
\]
From Lemma \([1]\), we have
\[
\sum_{\{\psi_B^n \approx \psi_B^n\}} \frac{1}{|K_A|} \geq 1 - \sqrt{2\epsilon}.
\] (24)
Similarly we have
\[
\sum_{\{\psi_B^n \approx \psi_B^n\}} \frac{1}{|K_A|} \geq 1 - \sqrt{2\epsilon}.
\] (25)
For both to be true, we have
\[
\sum_{\psi_B^n} \frac{1}{|K_A|} \geq 1 - 2\sqrt{2\epsilon},
\] (26)
where \( \tilde{K}_A := \{ \psi_B^n \approx \psi_B^n \wedge \psi_B^n \approx \psi_B^n \} \). Denote \( R_A = K_A \setminus \tilde{K}_A \), and we can write
\[
\frac{1}{|K_A|} \sum_{i=1}^{|K_A|} |i\rangle |i\rangle_{K_A} \otimes \psi_{A'B^n}^{'}
\]
\[
= \frac{1}{|K_A|} \sum_{i \in \tilde{K}_A} |i\rangle |i\rangle_{K_A} \otimes \psi_{A'B^n}^{'}
\]
\[
+ \frac{1}{|K_A|} \sum_{i \in R_A} |i\rangle |i\rangle_{K_A} \otimes \psi_{A'B^n}^{'}.
\]

The next step is for Bob to extract randomness by applying unitary \( V : B^{(n+m)} \to K_B B' \) on the state \( |\psi_{A'B'^n}^{'}(A'E^n) \rangle \otimes |\psi_{ABE}^{'}\rangle \), where \( A'E^n \) is virtually at Eve's side and \( |\psi_{A'B'^n}^{'}(A'E^n) \rangle \) is any state satisfying \( \psi_B^n \approx \psi_B^n \)
\( \approx \). Notice that such a unitary \( V \) is only determined by Bob's local states \( \psi_B^n \) and \( |\psi_{ABE}^{'}\rangle \) and is independent of the local state \( \phi_{A'E^n}^{'} \). Thus Bob can apply \( V \) blindly to all states of the form \( |\psi_{A'B'^n}^{'}(A'E^n) \rangle \otimes |\psi_{ABE}^{'}\rangle \). From Theorem \([13]\), we have
\[
V|\psi_{B'^n}(A'E^n)\rangle \otimes |\psi_{ABE}^{'}\rangle = \sum_{j=1}^{K_B} \sqrt{q_j}|j\rangle |K_B \otimes \phi_{B'^n}^{j}\rangle \otimes |\psi_{A'B'^n}^{'}\rangle \otimes |\psi_{ABE}^{'}\rangle.
\]

satisfying
\[
\left| \sum_{j=1}^{K_B} q_j |j\rangle |K_B \otimes \phi_{A'^n}^{j}\rangle \otimes |\psi_{ABE}^{'}\rangle \right| \leq \epsilon,
\]
\[
\left| \sum_{j=1}^{K_B} q_j |j\rangle |K_B \otimes \phi_{A'E^n}^{j}\rangle \otimes |\psi_{ABE}^{'}\rangle \right| \leq \epsilon,
\]
when \( \log |K_B| = \min\{mI(B : A), mI(B : E) + 2nS(B)\} - m\delta \).

Suppose
\[
V|\psi_{B'^n}(A'E^n)\rangle \otimes |\psi_{ABE}^{'}\rangle = \sum_{j=1}^{K_B} \sqrt{q_j}|j\rangle |K_B \otimes |\psi_{A'B'^n}^{'}\rangle \otimes |\psi_{ABE}^{'}\rangle.
\]

For \( i \in \tilde{K}_A \), we have \( \psi_B^n \approx \psi_B^n \approx \phi_{B^n}^{'} \). Using Uhlmann theorem (precisely first using Lemma \([8]\) then Lemma \([5]\)), there exist unitary operators \( W_i \) acting on \( A'E^n \) such that the purification \( W_i|\phi_{B^n}(A'E^n)\rangle W_i^\dagger = \phi_{A'B'^n}^{'}(A'E^n) \approx |\psi_{A'B'^n}^{'}(A'E^n)\rangle \) with \( \phi_B^n = \psi_B^n \) and \( \phi_{A'E^n}^{'} \approx \psi_{A'E^n}^{'} \), where \( \epsilon' = 4\sqrt{2}\epsilon - 2\epsilon \). Then we have
\[
V|\psi_{B'^n}(A'E^n)\rangle \otimes |\psi_{ABE}^{'}\rangle \approx V \otimes W_i|\phi_{B^n}(A'E^n)\rangle \otimes |\psi_{ABE}^{'}\rangle.
\] (29)

From Eq. \([29]\), we get
\[
\sum j q_j |j\rangle |K_B \otimes |\psi_{A'B'^n}^{'}\rangle \otimes |\psi_{ABE}^{'}\rangle,
\]
\[
\approx \epsilon \sum_j q_j |j\rangle |K_B \otimes W_i|\phi_{A'E^n}^{'}\rangle W_i^\dagger \otimes |\psi_{ABE}^{'}\rangle,
\]
\[
\approx \epsilon \sum_j q_j |j\rangle |K_B \otimes |\psi_{A'B'^n}^{'}\rangle \otimes |\psi_{ABE}^{'}\rangle,
\]
where Eq. \([30]\) comes from Eq. \([29]\) and the monotonicity of trace norm under dephasing operation on \( K_B \) and tracing \( A'^n \). (31) from \([28]\), \([32]\) from \( W_i|\phi_{A'E^n}^{'}\rangle W_i^\dagger = \phi_{A'E^n}^{'} \approx |\psi_{A'E^n}^{'}\rangle \), Summing the errors, we have, for \( i \in \tilde{K}_A \)
\[
\sum q_j |j\rangle |K_B \otimes |\psi_{A'B'^n}^{'}\rangle \otimes |\psi_{ABE}^{'}\rangle,
\]
\[
\approx 2\epsilon' + \epsilon \sum_j |j\rangle |K_B \otimes |\psi_{A'B'^n}^{'}\rangle \otimes |\psi_{ABE}^{'}\rangle.
\] (33)
For $i \in \tilde{K}_A$, we also have $\psi_{E^m|n}^\perp \approx \psi_{E^m|n}$, and by monotonicity of trace norm under tracing $A'$,

$$\sum_j q_{ij}|j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i} \leq 2^{\epsilon+\epsilon} \cdot \frac{1}{|K_B|} \sum_j |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i}$$

$$\leq \frac{\sqrt{\epsilon}}{|K_B|} \sum_j |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i} \approx 1$$

$$\sum_j q_{ij}|j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i} \leq \frac{1}{|K_B|} \sum_j |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i}$$

Summing the errors,

$$\sum_j q_{ij}|j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i} \approx \frac{1}{|K_B|} \sum_j |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i}$$

$$\approx 2^{\epsilon'} + \epsilon + \sqrt{2\epsilon}.$$ 

Now combining the two steps and considering the dephasing state, we get

$$\sum_{i \in K_A} p_i|i\rangle|i\rangle_K \otimes \sum_{j \in K_B} q_{ij}|j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i}$$

$$\approx \frac{1}{|K_A|} \sum_{i \in K_A} |i\rangle|i\rangle_K \otimes \sum_{j \in K_B} q_{ij}|j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i},$$

$$= \frac{1}{|K_A|} \sum_{i \in K_A} |i\rangle|i\rangle_K \otimes \sum_{j \in K_B} q_{ij}|j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i},$$

$$+ \frac{1}{|K_A|} \sum_{i \in K_A} |i\rangle|i\rangle_K \otimes \sum_{j \in K_B} q_{ij}|j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i},$$

$$\approx \frac{1}{|K_A|} \sum_{i \in K_A} |i\rangle|i\rangle_K \otimes \sum_{j \in K_B} |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i},$$

$$+ \frac{1}{|K_A|} \sum_{i \in K_A} |i\rangle|i\rangle_K \otimes \sum_{j \in K_B} |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i},$$

$$\approx \frac{1}{|K_A|} \sum_{i \in K_A} |i\rangle|i\rangle_K \otimes \frac{1}{|K_B|} \sum_j |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i},$$

$$\approx \frac{4\sqrt{\epsilon}}{|K_A|} \sum_{i \in K_A} |i\rangle|i\rangle_K \otimes \frac{1}{|K_B|} \sum_j |j\rangle|j\rangle_K \otimes \psi_{E^m|n}^{j|i},$$

where $\frac{4\sqrt{\epsilon}}{|K_A|}$ comes from $\sum_{i \in K_A} \frac{1}{|K_A|} \leq 2\sqrt{2\epsilon}$ and $\|\rho_1 - \rho_2\| \leq 2$ for any two states. We sum up all the errors and get

$$- \frac{1}{|K_A|} \sum_{i = 1}^{K_A} |i\rangle|i\rangle_K \otimes \frac{1}{|K_B|} \sum_{j = 1}^{K_B} |j\rangle|j\rangle_K \otimes \psi_E^{j|i}$$

$$\leq 2\epsilon + 5\sqrt{\epsilon} + 2\sqrt{4\epsilon} - 2\epsilon =: f(\epsilon).$$

The protocol can be composed iteratively in a new round by regarding the indices of $K_A K_B$ as a whole, and noticing that only the existence of the good index set is required in the proof. This means that Alice and Bob need not know how to select the set. They just blindly apply the decoupling unitary locally to all the states, then the dephasing state on the randomness subsystem will approximate the ideal randomness state.

**Proof of Theorem 3.** We first prove for the setting 1) of no noise and no communication. The other three settings can be reduced largely to setting 1), making a couple of simple observations.

The proof for setting 1) proceeds by the analysis of four cases.

**Case 1.** $S(A|E) \leq 0$ and $S(B|E) \leq 0$: The achievable region is

$$R_A \leq \log |A| + S(A|E)$$

$$R_B \leq \log |B| + S(B|E)$$

$$R_A + R_B \leq R_G$$

The extremal points are simply achieved by one party extracting randomness from the local purity part and from the information part by the VQSM protocol and the other just extracting randomness from the local purity part. The rate region is achieved by the time-sharing technique.

**Case 2.** $S(A|E) \leq 0$ and $S(B|E) \geq 0$: The achievable region is

$$R_A \leq \log |A| + S(A|E)$$

$$R_B \leq \log |B|$$

$$R_A + R_B \leq R_G$$

One of the extremal points is $(\log |A| + S(A|E), \log |B| - S(B))$ by the same protocol as above. The other extremal point is $(\log |A| - S(E), \log |B|)$, which is achieved as follows: Alice distills local purity from all copies, then performs VQSM on $pn$ copies of the informational part to merging her state to Eve; this produces $R_A = \log |A| - S(A) + pI(A:B)$ of randomness on Alice’s side. Then Bob extracts local randomness on the virtual state $\psi_{ABC}^{\otimes p} \otimes \psi_{AB}^{\otimes pn}$ by state compression and VQSM. If we choose $p$ satisfying the $(1 - p)S(B|E) - pS(B) = 0$, then it will produce $\log |B|$ randomness on Bob.

**Case 3.** $S(A|E) \geq 0$ and $S(B|E) \leq 0$ is analogous.
Case 4. $S(A|E) > 0$ and $S(B|E) > 0$: We suppose Alice and Bob pre-process their states by local compression, separating the purity parts and the remaining information parts that contain the correlation. We only deal with the information part. Neither Alice nor Bob can directly perform VQSM. Alice and Bob will take a “ping-pong” strategy to help each other obtain required singlets between Alice and Eve, and Bob and Eve in order to perform VQSM. It is an activation process and it works only if the activation process could be amplified to more and more copies. So it has to satisfy some constraints. However finally we find that the constraint is always satisfied. So there is no constraint at all!

In detail, suppose that initially Bob puts aside $x$ of his copies (the consumed will be negligible when we count the rate at the end) and will not use them any more. Alice can imagine Bob’s $x$ part is in Eve’s hands, so she would share $xS(A)$ ebits with Eve by her $x$ systems. Then Alice performs VQSM on $n_1$ fresh copies using the $xS(A) = 1$ ebits. From VQSM, we know that one ebit can activate $n_1$ copies satisfying $n_1S(A|E) = 1 = 0$. After VQSM from Alice to Eve, new ebits are created between Bob and Eve, and the number is $n_1S(B)$. Then Bob performs VQSM from Bob to Eve on $n_2$ fresh copies under the assistance of $n_1S(B)$ ebits, where $n_2$ satisfies $n_2S(B|E) = n_1S(B) = 0$. After VQSM from Bob to Eve, Alice would share $n_2S(A)$ ebits with Eve. This is one round. The process can be amplified if only $n_2S(A) > 1$, which reads $S(A)S(B) = S(A|E)S(B|E)$. This is always satisfied, unless we are in a trivial situation. By symmetry, if Bob does VQSM first, still we have the same constraint (i.e., no constraint). Now consider the initial step is performed on $xK$ copies, where $K$ is large such that indeed VQSM is performed. Because at any step of any round, the extraction randomness rate is exactly the same rate of $I(A : B)$, so overall the rate is also $I(A : B)$. In addition with the local purity, the sum of the purity achieves the global purity.

Now we compute the rate: denote $r := \frac{S(A)S(B)}{S(A|E)S(B|E)} > 1$. Suppose the “ping-pong” protocol ends at Alice’s side after $L$ rounds with a large finite number $L$. Then we know that Alice gets the randomness from $N_A$ copies where

$$N_A = n_1 + n_3 + \cdots + n_{L+1}$$

$$= \frac{K}{S(A|E)} (1 + r + r^2 + \cdots + r^{L+1})$$

$$= \frac{K}{S(A|E)} r^{L+2} - 1 \quad \frac{r^{L+1} - 1}{r - 1} ,$$

and Bob from

$$N_B = n_2 + n_4 + \cdots + n_L$$

$$= \frac{KS(B)}{S(A|E)S(B|E)} (1 + r + r^2 + \cdots + r^L)$$

$$= \frac{KS(B)}{S(A|E)S(B|E)} r^{L+1} - 1 \quad \frac{r^L - 1}{r - 1} ,$$

and the total number of copies is $N = N_A + N_B$. The randomness rate depends only on the proportion of copies that VQSM is performed on. So $\frac{N_A}{N} \approx \frac{S(A)}{S(A|E)} = \frac{S(B)}{S(B|E)}$ and $\frac{N_B}{N} = \frac{S(A)I(A : B)}{I(A : B)} S(B|E)$ for large $L$. Thus we get $R_A = \frac{N_A}{N} I(A : B) + \log |A| - S(A) = \log |A|$ and $R_B = \log |B| - S(AB)$, which forms one of the two extremal points of the claimed region. Exchanging the role of Alice and Bob we get the other one ($\log |A| - S(AB), \log |B|$).

Proofs of other settings.

2) Free noise but no communication means that we are effectively applying Theorem 3 to states $\rho_{AB} \otimes \frac{I(A)}{\mathbb{I}} \otimes \frac{I(B)}{\mathbb{I}}$, with large local dimension of $A’$ and $B’$ instead of $\rho_{AB}$. In this way, the two conditional entropies in Theorem 3 can be made positive while the local dimension increases, hence the entire rate region becomes

$$R_A \leq \log |A| - S(AB),$$

$$R_B \leq \log |B| - S(B|A),$$

$$R_A + R_B \leq R_G = \log |AB| - S(AB). \quad (49)$$

3) Free noise and free communication: The rate region is characterised by

$$R_A \leq R_G, \quad R_B \leq R_G, \quad \text{and } R_A + R_B \leq R_G, \quad (50)$$

i.e. the entire global purity can be realised on either side as randomness (but not necessarily as purity)! To see this consider an extreme point of the free-noise region, Eq. [48], $R_A = \log |A| - S(AB) + S(B) = \log |A| - S(A) + I(A : B), R_B = \log |B| - S(B)$. In this case, Bob’s entire randomness, and the part $\log |A| - S(A)$ of Alice are realised as purity obtained by first compressing their respective systems. But purity can be exchanged freely via the dephasing channel, so the rate sum can be concentrated at Alice’s side. Similarly for Bob.

4) For free communication but no noise we have only an achievable region from combining the above Theorem with the sharing of purity observed in the second case: We can certainly achieve all rate pairs with

$$R_A \leq \log |A| - \max \{S(B), S(AB)\},$$

$$R_B \leq \log |B| - \max \{S(A), S(AB)\},$$

$$R_A + R_B \leq R_G = \log |AB| - S(AB). \quad (51)$$

Proof of tightness. We prove the tightness of the rate regions in three of these settings: Theorem 3 (scenario 1), and scenario 2), but leave 3) open. In all settings, $R_A + R_B \leq R_G = \log |AB| - S(AB)$, which clearly
cannot be beaten, even if Alice and Bob can freely cooperate. For setting 2): It clearly cannot be improved, because \( R_G \) is reached for both parties, and nothing can be better. Setting 1): Since there is no communication, Bob can always isolate his purity \( \log |B| - S(B) \) as randomness, regardless of what Alice does. Hence \( R_A \leq R_G - (\log |B| - S(B)) = \log |A| - S(A|B) \). Similarly for Bob’s rate. Theorem 3. Even with free noise, \( R_A \leq \log |A| - S(A|B) \), so this bound is still true without free noise. But also \( R_A \leq \log |A| \) always because Alice has only her system \( A \) to work on. So \( R_A \) is bounded by the minimum of the two; likewise for Bob.

We arrive at the form of the irreducible states by which we mean that all the randomness is in the key part and no extra randomness can be gained from the shield part. So any optimal protocol must reach this kind of state, asymptotically.

Lemma 14 Any irreducible state is of the form

\[
\alpha_{K A'B'} = \frac{1}{|K|} \sum_{j=1}^{|K|} |j\rangle \langle j|_K \otimes \frac{1}{|A'||B'|} U_i U_j^*,
\]

where \( U_i \) are unitaries acting on \( A'B' \).

Proof. Suppose the purification for the irreducible state is \( |\phi\rangle_{AA'B'E} \), where the randomness part is \( A \) and \( A'B' \) is the shield part. From the form of \( \alpha_{AA'B'} \), we know

\[
|\phi\rangle_{AA'B'E} = \frac{1}{|K|} \sum_{i=1}^{|K|} |i\rangle_K \otimes U_i |\phi_0\rangle_{A'B'E}.
\]

Because it is an irreducible state which means the optimal rate of randomness that can be extracted under RCLOCC is \( \log K \), any protocol cannot produce more than \( \log K \) bits. Now consider the BFW protocol \[1\] which is naturally a RCLOCC protocol where Bob does nothing, it creates \( |K| + \log |A'| + S(A|E') \) bits which should be no larger than \( \log K \). Then we get \( \log |A'| + S(A|E') \leq 0 \) that implies \( S(KA|E') = S(B') - S(E) \geq 0 \). So we have \( \log |A'| + S(B') - S(E) \leq 0 \). Notice that \( \alpha_{B'} = 1/|B'| \) (or Bob can do local compression to get purity and send the purity to Alice), and \( S(E) = S(A|B'_i) \leq S(A') + S(B') \leq \log |A'| + S(B') \), the equality holds if and only if \( \phi_{A'B'} = \frac{|A'|}{|A'|} \otimes \frac{|B'|}{|B'|} \). This means that \( |\phi_i\rangle_{A'B'E} = U_i |\Phi_i\rangle_{A'E'} \otimes |\Phi_i\rangle_{B'E} \) up to the local unitaries on \( E_1 E_2 \). Tracing out the \( E \) part we get the desired form.

From Lemma 14 it is clear the last step of an optimal protocol is VQSM when Eve’s system is included. The irreducible form expresses that entropy of the Eve is not less than Bob’s entropy.

Proof of Theorem 4

Denote the quantity at the right-hand side as RHS.

“\( R_A \geq \) RHS”: For any RCLOCC that transforms \( n \) copies of \( \rho_{AB} \) to a state \( \sigma_{AB} \) whose purification is denoted as \( |\phi\rangle_{A'B'E'} \). We notice that in case of free communication and no noise in the two-side randomness scenario, Alice can get randomness at the rate \( \log |A'| - \max\{S(B'), S(E')\} = n \log |AB| - \max\{S(B), S(E')\} \). So we have \( R_A \geq \max\{n \log |AB| - \max\{S(B)/n, S(E')/n\}\} \) = RHS.

\( R_A \leq \) RHS”: Given \( n \) copies of \( \rho_{AB} \), a protocol to extract randomness ends in a state \( \omega_{K A'B'} \) that is close to the state \( \alpha_{K A'B'} = \frac{1}{|K|} \sum_{j=1}^{|K|} |j\rangle \langle j|_K \otimes U_j \sigma_{A'B'} U_j^* \) in the sense \( |\omega_{K A'B'} - \alpha_{K A'B'}| \leq \epsilon(n) \) when \( n \to \infty \). Now suppose the purification state for \( \omega_{K A'B'} \) is \( |\psi\rangle_{A'B'} \). From the continuity of entropy, we get \( |S(\omega_{K A'B'}) - S(\alpha_{K A'B'})| \leq \epsilon(n) |\log |AB|| + h(\epsilon(n)) \) which amounts to \( |S(\omega_{E'}) - S(\alpha_{E'})| \leq \epsilon(n) |\log |AB|| + h(\epsilon(n)) \), and \( |S(\omega_{E'}) - S(\alpha_{E'})| \leq \epsilon(n) |\log |B'|| + h(\epsilon(n)) \). Then we have \( |K|/n \leq \log |AB| - \log |B'|/n + \log \epsilon(n) |\log |AB|| + h(\epsilon(n)) \). More directly we have \( |K|/n \leq \log |AB| - \log |B'|/n + \log \epsilon(n) |\log |AB|| + h(\epsilon(n)) \). Take the supremum over \( n \), and notice that if for a finite \( n \), the supremum is obtained, then for any \( mn \), it is achieved. So we get \( R_A = \sup \log |K|/n \leq \) RHS.

Proof of Theorem 5

It is straightforward that the randomness \( R_A \) that can be extracted on Alice’s side is not larger than the randomness extracted by global operation which is equal to the global purity. So \( R_A \leq \log |AB| - S(AB) \). Suppose for \( n \) copies of \( \rho_{AB} \) whose purification is \( |\phi\rangle_{ABE} \) with \( E \) on Eve’s side, we get an approximate irreducible state \( \alpha_{K A'B'} \) under RCLOCC operations, where \( K \) is the randomness part and \( A'B' \) is the shield part, \( K'A \) is on Alice’s side and \( B' \) on Bob’s, and its purification part is denoted as \( E' \). Then we get

\[
\log |K| \leq n \log |AB| - S(E'),
\]

\[
= n \log |AB| - S(E') - nS(E) - nS(E),
\]

\[
\leq n \log |AB| - [E_r(\rho_{AB}) - E_r(\alpha_{K A'B'} | - nS(E),
\]

\[
\leq n \log |AB| - E_r(\rho_{AB}) - nS(E) + S(B'),
\]

\[
\leq n \log |AB| - E_r(\rho_{AB}) - nS(E) + \log |B'|,
\]

\[
= n \log |AB| - E_r(\rho_{AB}) - nS(E) + \log |K A'| - \log |K A'|,
\]

\[
\leq 2n \log |AB| - E_r(\rho_{AB}) - nS(AB) - \log |K|,
\]

where Ineq. 52 comes from the BFW protocol [1], Eq. 53 from additivity of entropy for product state, Eq. 54 from Lemma 12, Ineq. 55 from the fact that \( E_r(\rho_{AB}) \leq S(Y) \), Ineq. 56 from S(X) ≤ log |X|, and
Ineq. \((58)\) from that the dimension of \(AB\) is invariant. So we obtain \(\log |K| \leq n \log |AB| - \frac{1}{2} [E^{\infty}(\rho_{AB}) + S(\rho_{AB})]\). Since it holds for any randomness extraction, we get \(R_A = \sup \log |K|/n \leq \log |AB| - [E^{\infty}(\rho_{AB}) + S(\rho_{AB})]/2\). □

**Proof of Corollary** \(7\) We get this by constructing an explicit protocol. Consider two singlets \(|\Phi_{AB} \otimes |\Phi_{A'B'}\rangle\). Bob sends \(B_1\) through the dephasing channel and the resulting state is \(\frac{1}{2}(|00\rangle + |11\rangle)_{A_1B_1E}\) when we include Eve. Alice performs a CNOT to get \(|\Phi_{A_1E} \otimes |0\rangle_{B_1}\), where the pure state \(|0\rangle_{B_1}\) can produce 1 ibit. Then Alice performs a Bell measurement on \(A_1A_2\) to get another 2 ibits, in total 3 ibits from 2 copies of the state. This is optimal by Theorem \(5\). □

**Proof of Theorem** \(7\) Suppose for \(n\) uses of the channel, the final tripartite state among Alice, Bob, and Eve is

\[|\psi\rangle_{A^nB^nE^n} = V^{\otimes n}(|\psi\rangle_{A^nA^n}),\]

where \(V : A' \rightarrow BE\) is an isometry purification of the channel \(N\). Then with free noise, from case 2) we know for such a state, Bob can extract randomness with \(n \log |B| - S(B^n|A^n)_{\psi_{A^nB^n}} = n \log |B| + [S(A^n) - S(A^n_{\psi_{A^nB^n}})]\). Further, from Theorem \(13\) we know that the randomness value which is represented by the dephased system \(K_B\) is decoupled from both system \(A\) of the sender, and \(E\) of the eavesdropper, thus is private to each of them individually.

To maximize the quantity amounts to maximize the second term which is none other than the reverse coherent information, and from \(27\), we know that the reverse coherent information of a channel is additive. For the sake of completeness, we write the proof. First using the duality relation between conditional entropy of a pure state \(|\psi\rangle_{ABE}\), we have

\[S(A) - S(AB) = -S(B|A) = S(B|E).\]

Then by the chain rule \(S(XY|Z) = S(X|Z) + S(Y|XZ)\) and the monotonicity of conditional entropy \(S(X|Y) \leq S(X|E(Y))\) under quantum operation \(\mathcal{E}\), we get

\[S(B_1B_2|E_1E_2) = S(B_1|E_1E_2) + S(B_2|B_1E_1E_2), \leq S(B_1|E_1) + S(B_2|E_2),\]

which immediately implies the additivity

\[
\begin{align*}
\max_{|\phi\rangle_{A_1A_2}} [S(\phi_A) - S(N_1 \otimes N_2(\phi_{AA_A}A_2))] \\
&= \max_{|\phi\rangle_{A_1A_2}} [S(\phi_A) - S(N_1(\phi_{AA_A}))] \\
&\quad + \max_{|\phi\rangle_{A_1A_2}} [S(\phi_A) - S(N_2(\phi_{AA_A}))].
\end{align*}
\]

So optimisation over \(n\) channel use is reduced to the optimization over one channel use.

Consider the state of the form

\[p\rho_{BE} \otimes |0\rangle_F + (1 - p)\sigma_{BE} \otimes |1\rangle_F.\]

By the linearity of isometry \(V : A' \rightarrow BE\) on input state and the monotonicity of the conditional entropy \(S(B|E) \geq S(B|EF)\), we conclude that \(S(B|E)\) is concave with respect to the input state,

\[S(B|E)_{p\rho_{BE} + (1 - p)\sigma_{BE}} \geq pS(B|E)_{\rho_{BE}} + (1 - p)S(B|E)_{\sigma_{BE}},\]

which amounts to the reverse coherent information is concave to the input state, thus the optimisation problem is efficiently computable. □

**Lemma 15** For a state \(\rho_{AB}\), if global operations are allowed, the randomness extraction rate is \(R_G(\rho_{AB}) = I_G(\rho_{AB}) = \log |AB| - S(\rho_{AB})\).

**Proof.** It is clear that a pure state \(|0\rangle\) is the simplest ibit. So given a state \(\rho_{AB}\), if global operations are allowed, then we can obtain purity at the rate \(I_G = \log |AB| - S(\rho_{AB})\). From the purity the same amount of randomness can be obtained which implies \(R_G \geq I_G\).

On the other side, any randomness extraction ends up in a state close to \(\alpha\) state, on which we can undo the twisting \(U\) to get a pure state on the key part which means \(I_G \geq R_G\). So we get \(I_G = R_G\). □

**Lemma 16** Non-existence of bound randomness states; given bipartite state \(\rho_{AB}\), if randomness can be extracted by global operations on \(AB\), then nonzero randomness can be extracted by RCLOOC operations.

**Proof.** Given a state \(\rho_{AB}\), if none-zero randomness can be extracted by global operations, then it means that \(R_G = \log |AB| - S(AB) > 0\). Notice that Theorem \(3\) already implies that there exists \((R_A, R_B)\) such that \(R_A + R_B = R_G > 0\) can be achieved. □