Catalytic Decoupling of Quantum Information

Christian Majenz,[1,*] Mario Berta,[2] Frédéric Dupuis,[3] Renato Renner,[4] and Matthias Christandl[1]
[1]Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø
[2]Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA
[3]Faculty of Informatics, Masaryk University, Brno, Czech Republic
[4]Institute for Theoretical Physics, ETH Zurich, 8093 Zürich, Switzerland

(Received 24 May 2016; published 23 February 2017)

The decoupling technique is a fundamental tool in quantum information theory with applications ranging from thermodynamics to many-body physics and black hole radiation whereby a quantum system is decoupled from another one by discarding an appropriately chosen part of it. Here, we introduce catalytic decoupling, i.e., decoupling with the help of an independent system. Thereby, we remove a restriction on the standard decoupling notion and present a tight characterization in terms of the max-mutual information. The novel notion unifies various tasks and leads to a resource theory of decoupling.

DOI: 10.1103/PhysRevLett.118.080503

Introduction.—Erasing correlations between quantum systems via local operations, decoupling, is a task that was first studied in the context of quantum information theory [1] (see [2] for an introductory tutorial). It serves as a building block for a variety of tasks in quantum information and quantum cryptography. In particular, decoupling has been crucial for understanding how to distribute quantum information between different parties [3–7] and for understanding how to send quantum information over noisy quantum channels [8–11], as well as randomness extraction [12]. The concept is, however, also very useful in physics (as, e.g., outlined in [13]). Applications range from quantum thermodynamics [14–16] to the studies of black hole radiation [17–19] and solid state physics [20].

Standard decoupling.—The basic idea behind decoupling is the following: if a mixed bipartite quantum state ϱ_{AE} is only weakly correlated, then it should suffice to erase a small part of A to approximately decouple A from E, i.e., to get an approximate product state [see Fig. 1(a)]. More precisely, we say that a bipartite quantum state ϱ_{AE} is ε-decoupled by the partial trace map $\mathcal{T}_{A\to A_1}(\cdot)=\mathrm{Tr}_{A_2}[\cdot]$, with $A=A_1A_2$ if there exists a unitary operation U_A such that

$$\min_{\omega_{A_1}\otimes\omega_E} P[\mathcal{T}_{A\to A_1}(U_A\varrho_{AE}U_A^\dagger),\omega_{A_1}\otimes\omega_E]\le\varepsilon, \tag{1}$$

where the minimum is over all product quantum states $\omega_{A_1}\otimes\omega_E$, and $P(\beta,\gamma)\coloneqq(1-\|\sqrt{\beta}\sqrt{\gamma}\|_1^2)^{1/2}$ denotes the purified distance [21]. The A_1 system is called the decoupled system and the A_2 system, the remainder system—when trying to decouple A from E, we succeed on A_1, and A_2 is the remainder we fail to decouple. The fundamental question that we want to discuss is how large we have to choose the remainder system A_2 in order to achieve ε-decoupling. We denote the minimal remainder system size, i.e., the logarithm of the minimal remainder system dimension, for ε-decoupling A from E in a state ϱ_{AE} by $R^\varepsilon(A;E)_\varrho$. For a formal definition of $R^\varepsilon(A;E)_\varrho$, see Supplemental Material, Definition 18 [22].

Converse.—We first show quite naturally that $R^\varepsilon(A;E)_\varrho$ has to be at least the size of the smooth max-mutual information $I^\varepsilon_{\max}(E{:}A)_\varrho$ present in the initial state ϱ_{AE}. This measure is defined as [11]

$$I^\varepsilon_{\max}(E;A)_\varrho\coloneqq\min_{\bar\varrho} I_{\max}(E;A)_{\bar\varrho},\quad\text{with} \tag{2}$$

$$I_{\max}(E;A)_{\bar\varrho}\coloneqq\log\min\{\mathrm{Tr}\sigma_A|\sigma_A\otimes\bar\varrho_E\ge\bar\varrho_{AE}\}, \tag{3}$$

where the minimum in (2) is over all bipartite quantum states, with $P(\varrho_{AE},\bar\varrho_{AE})\le\varepsilon$ [32], and the minimum in (3) is over all $\sigma_A\ge0$. We note that the definition of the smooth max-mutual information is *a priori* not symmetric in $A{:}E$. However, we have [33]

$$I^\varepsilon_{\max}(E;A)_\varrho\approx I^\varepsilon_{\max}(A;E)_\varrho, \tag{4}$$

where $\approx$ stands for equality up to terms $\mathcal{O}(\log(1/\varepsilon))$. For the converse, we exploit that the smooth max-mutual information is invariant under local unitary operations

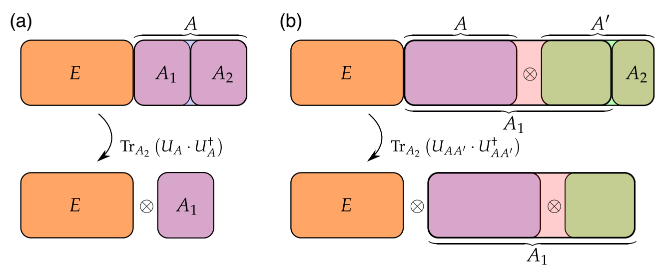

FIG. 1. Schematic representation of (a) standard and (b) catalytic decoupling: tracing out a system A_2 leaves system A_1 decoupled from E. While there is no ancilla for standard decoupling as in (a), catalytic decoupling as in (b) allows us to make use of an additional, already decoupled system A'. The basic question is how large we have to choose the system A_2 such that system A_1 is decoupled from E.

and that it has the so-called nonlocking property (see [34] about information locking). That is, just like the quantum mutual information, it fulfills the inequality [[11], Lemma B.12],

$$I^{\varepsilon}_{\max}(E;A_1A_2)_{\varrho} \leq I^{\varepsilon}_{\max}(E;A_1)_{\varrho} + 2\log|A_2|, \qquad (5)$$

where $|A_2|$ denotes the dimension of A_2. Since the final state is a product state, its smooth max-mutual information $I^{\varepsilon}_{\max}(E;A_1)_{\omega\otimes\omega}$ becomes zero. This means that in order to erase the initial correlations $I^{\varepsilon}_{\max}(E;A)_{\varrho}$, we need at least a remainder system of size [35]

$$R^{\varepsilon}(A;E)_{\varrho} \geq \frac{1}{2} I^{\varepsilon}_{\max}(E;A)_{\varrho}. \qquad (6)$$

Previous works.—Most of the aforementioned decoupling references only give good achievability bounds for states of the form $\varrho_{A^nE^n} = \varrho_{AE}^{\otimes n}$ in the asymptotic limit $n \to \infty$. Whereas this setting is relevant in quantum Shannon theory, it is often a severe restriction for applications in physics. For typical physical situations (e.g., in thermodynamics), there is usually not even a natural decomposition of a large system in n subsystems. A notable exception concerning achievability results is Ref. [13], where the authors show that

$$R^{\varepsilon}(A;E)_{\varrho} \lesssim \frac{1}{2}(H^{\varepsilon'}_{\max}(A)_{\varrho} - H^{\varepsilon'}_{\min}(A|E)_{\varrho}), \quad \text{with} \quad \varepsilon' = \frac{\varepsilon}{5}, \qquad (7)$$

where $\lesssim$ means up to terms $\mathcal{O}(\log(1/\varepsilon))$ here. (We give a proof of this particular statement in the Supplemental Material [22]). Here, $H^{\varepsilon}_{\max}$ and $H^{\varepsilon}_{\min}$ denote the smooth conditional max- and min-entropy, whose exact definitions can be found in the Supplemental Material [22] (or see the textbook [21]). In fact, the results from [13] show that not only decoupling in the sense of (1) is achieved, but moreover, that the decoupled system is also randomized. That is, there exists a quantum state ω_E and a unitary operation U_A such that Eq. (1), with $\omega_{A_1} = 1_{A_1}/|A_1|$, holds and where 1_{A_1} denotes the identity matrix on A_1.

It turns out that there can be an arbitrarily big gap between the converse (6) and the achievability results (7). This is best seen for an example with trivial system E, where the corresponding max-mutual information converse bound becomes zero. In that case, the achievability bound (7) reduces to the difference between the smooth max- and min-entropy and it is known that this can become roughly as big as $\log|A|$ (we provide an explicit example in the Supplemental Material [22]). In order to achieve the converse from (6), we propose in the following a generalized notion of decoupling.

Catalytic decoupling.—A natural question to ask at this point is if decoupling can be achieved more efficiently in the presence of an already uncorrelated ancilla system (see Fig. 1). Formally, we say that a bipartite quantum state ϱ_{AE} is ε-decoupled catalytically by the ancilla state $\varrho_{A'}$ and the partial trace map $\mathcal{T}_{\bar{A}\to A_1}(\cdot) = \mathrm{Tr}_{A_2}[\cdot]$, with $\bar{A} \equiv AA' \cong A_1A_2$, if there exists a unitary operation $U_{\bar{A}}$ such that

$$\min_{\omega_{A_1}\otimes\omega_E} P[\mathcal{T}_{\bar{A}\to A_1}(U_{\bar{A}}\varrho_{\bar{A}E}U^{\dagger}_{\bar{A}}), \omega_{A_1}\otimes\omega_E] \leq \varepsilon, \qquad (8)$$

$$\text{where } \varrho_{\bar{A}E} = \varrho_{AE}\otimes\varrho_{A'}. \qquad (9)$$

Again, we call the A_1 system the decoupled system and the A_2 system, the remainder system. The term catalytic means that the share of the initially uncorrelated ancilla system A', that becomes part of the decoupled system A_1, stays decoupled (see Fig. 1).

Now, we are interested in the minimal size of the remainder system A_2 in order to achieve ε-decoupling catalytically. We denote the minimal remainder system size for catalytically decoupling A from E in a state ϱ_{AE} by $R^{\varepsilon}_c(A;E)_{\varrho}$. For a formal definition of $R^{\varepsilon}_c(A;E)_{\varrho}$, see Supplemental Material, Definition 19 [22]. Clearly, we have $R^{\varepsilon}_c(A;E)_{\varrho} \leq R^{\varepsilon}(A;E)_{\varrho}$, as we can always choose a trivial ancilla. Moreover, since appending with an ancilla does not change the smooth max-mutual information (see Supplemental Material [22]), the same converse as in (6) still holds.

One may analyze decoupling using a resource-theoretic approach, treating decoupled systems as a resource. A quantum system A coupled to the environment E can yield a decoupled system A_1 of a certain size through standard decoupling. That is, in the resource theory language of [36], we have $\langle\varrho_{AE}\rangle \geq_{\varepsilon} (\log|A| - R^{\varepsilon}(A;E)_{\varrho})[d]$. Here, $x[d]$ denotes x decoupled qubits, and $\geq_{\varepsilon}$ stands for up to error ε (see also [37]), while the set of free operations is given by the unitary operations [38]. Now, our novel paradigm makes use of the possibility that if we already have decoupled qubits, then we are able to decouple a larger system [39]

$$\langle\varrho_{AE}\rangle + n[d] \geq_{\varepsilon} (n + \log|A| - R^{\varepsilon}_c(A;E)_{\varrho})[d], \quad \text{for } n \text{ large enough}. \qquad (10)$$

Note, however, that this inequality is only proved for *specific* initial and final decoupled states used in the presented decoupling protocols.

Tight achievability.—In contrast to standard decoupling as in (1), catalytic decoupling can be achieved with a remainder system size that is essentially equal to the smooth max-mutual information.

Theorem 1: (Catalytic decoupling) For any bipartite quantum state ϱ_{AE} and $0 < \delta \leq \varepsilon \leq 1$, we have:

$$R^{\varepsilon}_c(A;E)_{\varrho} \lesssim \frac{1}{2} I^{\varepsilon-\delta}_{\max}(E;A)_{\varrho}, \qquad (11)$$

where $\lesssim$ stands for smaller or equal up to terms $\mathcal{O}(\log\log|A| + \log(1/\delta))$. We also have the converse

$$R^{\varepsilon}_c(A;E)_{\varrho} \geq \frac{1}{2} I^{\varepsilon}_{\max}(E:A)_{\varrho}. \qquad (12)$$

Note that the converse comes from Eq. (6).

In fact, we not only show that catalytic decoupling in the sense of (8) is achieved, but moreover, that the decoupled system ends up in the marginal of the original state:

$$P(\varrho_{A_1E}, \varrho_{A_1} \otimes \omega_E) \leq \varepsilon \quad \text{for some quantum state } \omega_E. \tag{13}$$

In particular, and in contrast to the standard decoupling results leading to (7), our catalytic decoupling scheme does not randomize the decoupled system, but leaves it invariant (up to the approximation error ε). We can even choose $A_1 = AA_1'$ such that the decoupled system contains the marginal of the input state (A) (plus part of the catalyst, A_1').

In the Supplemental Material [22], we give two conceptually different proofs for Theorem 1. The first proof is based on the standard decoupling techniques from [11,13], combined with the use of embezzling entangled quantum states [40]. For (11), this yields a difference of size at most $\log\log|A| + \mathcal{O}(\log(1/\delta))$ [41]. The second proof is based on the convex splitting technique of Anshu *et. al.* [42]. It allows us to upper bound the difference in (11) with the tighter bound

$$R_c^\varepsilon(A;E)_\varrho - \frac{1}{2} I_{\max}^{\varepsilon-\delta}(E;A)_\varrho \leq \frac{1}{2}\{\log\log I_{\max}^{\varepsilon-\delta}(E;A)_\varrho\}_+ + \mathcal{O}(\log(1/\delta)), \tag{14}$$

where $\{\cdot\}_+ \coloneqq \max\{0,\cdot\}$. Moreover, this argument is constructive and hence, leads to an explicit scheme for decoupling. This improves on the standard decoupling bounds, which are achieved using the probabilistic technique [43] (as, e.g., the previously best known bound (7) from [13]).

Discussion.—The achievability result (11), together with the converse (12), establishes an operational interpretation of the smooth max-information as twice the minimal size of the remainder system to achieve ε-decoupling. We note that the approximation error as well as the smoothing parameter can be made arbitrarily close in (12) and (11) with only a logarithmic penalty. This enables us to make a statement about the case of many independent copies of a state, the so-called IID setting (independent, identically distributed). Following the information-theoretic arguments outlined in [45] (which, in turn, are based on ideas from [46,47]), we find that for states of the form $\varrho_{A^nE^n} = \varrho_{AE}^{\otimes n}$ and large $n \to \infty$,

$$\frac{1}{n} R_c^\varepsilon(A^n;E^n)_{\varrho^{\otimes n}} = \frac{1}{2}\left(I(A:E)_\varrho + \sqrt{\frac{V(A:E)_\varrho}{n}}\Phi^{-1}(\varepsilon)\right) + \mathcal{O}\left(\frac{\log n}{n}\right), \tag{15}$$

with the mutual information $I(A:E)_\varrho = H(A)_\varrho + H(E)_\varrho - H(AE)_\varrho$, featuring the von Neumann entropy $H(A)_\varrho = -\mathrm{Tr}(\varrho_A \log \varrho_A)$, and the mutual information variance $V(A:E)_\varrho$, as well as the cumulative normal distribution function Φ, specified in the Supplemental Material [22]. We note that no such tight (second-order) asymptotic expansion is known for standard decoupling. However, the achievability (7), together with the converse (6), implies that (using the asymptotic equipartition property from [21]),

$$\lim_{n\to\infty} \frac{1}{n} R^\varepsilon(A^n;E^n)_{\varrho^{\otimes n}} = \frac{1}{2} I(A:E)_\varrho \quad \text{for } 0 < \varepsilon < 1. \tag{16}$$

Thus, we can conclude that catalytic decoupling and standard decoupling become equivalent in the first order rate asymptotically: the mutual information quantifies the minimal size of the remainder system.

Applications.—We are now going to illustrate the use of catalytic decoupling with various applications. Groisman *et. al.* [48] introduced an operational approach to quantifying the total correlations that are present in a quantum state. In analogy to Landauer's erasure principle [49], they characterize the strength of correlations by the amount of randomness that has to be injected locally to decorrelate the state. This randomizing is done by a random-unitary channel on one of the systems (called local unitary randomizing, A-LUR in [48]):

$$\Lambda(\cdot) = \sum_{i=1}^N p_i U_i(\cdot) U_i^\dagger. \tag{17}$$

We say that the correlations between A and E in a state ϱ_{AE} can be ε-erased by a local mixture of N unitaries on A if Λ_A ε-decouples A from E. That is, if there exists a quantum channel Λ_A of the form (17) such that

$$\min_{\omega_A\otimes\omega_E} P(\Lambda_A(\varrho_{AE}), \omega_A \otimes \omega_E) \leq \varepsilon. \tag{18}$$

We denote the logarithm of the minimal number of unitaries needed for ε erasing the correlations between A and E in a state ϱ_{AE} by $R_U^\varepsilon(A;E)_\varrho$. For a formal definition of $R_U^\varepsilon(A;E)_\varrho$, see Supplemental Material, Definition 20 [22]. Groisman *et. al.* show that for states of the form $\varrho_{A^nE^n} = \varrho_{AE}^{\otimes n}$ for large $n \to \infty$:

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} \frac{1}{n} R_U^\varepsilon(A^n;E^n)_{\varrho^{\otimes n}} = I(A:E)_\varrho. \tag{19}$$

In the following, we will see that the task of *catalytic* local erasure of correlation becomes equivalent to catalytic decoupling [50]. We therefore define $R_{U,c}^\varepsilon(A;E)_\varrho \coloneqq \inf R_U^\varepsilon(AA':E)_{\varrho\otimes\sigma}$, where the infimum is taken over all ancilla systems. This quantity is formally defined in Supplemental Material, Definition 20 [22].

Proposition 2: (Erasure of correlations) For any bipartite quantum state ϱ_{AE}, we have $\frac{1}{2} R_{U,c}^\varepsilon(A;E)_\varrho = R_c^\varepsilon(A;E)_\varrho$. Hence, we get

$$I_{\max}^\varepsilon(E;A)_\varrho \leq R_{U,c}^\varepsilon(A;E)_\varrho \lesssim I_{\max}^{\varepsilon-\delta}(E;A)_\varrho, \tag{20}$$

where $\lesssim$ stands for smaller or equal up to terms $\mathcal{O}(\log\log|A| + \log(1/\delta))$. The same asymptotic expansion as in (15) holds.

This is the generalization of the results in [48] to arbitrary (structureless) states (see also [9,46,47]). It gives an alternative operational characterization of the smooth max-mutual information as the minimal number of unitaries needed for ε erasing the correlations between A and E. The proof of Proposition 2 proceeds as follows. Suppose we have a way of decoupling A from E, with remainder system A_2, and let $|A_2| = 2^k$ for some $k \in \mathbb{N}$. Then, we can think of A_2 as k qubits and erase each of them applying a uniform mixture of the Pauli matrices and the identity. This is a mixture of $4^k = 2^{2k}$ unitaries. Conversely, suppose we have a mixture of $N = 2^{2k}$ unitaries on A that erase the correlations to E. We take the mixed ancilla state $1_{A'_1A'_2}/|A'_1A'_2|$, with $A'_i \cong \mathbb{C}^{2^k}$. Now, we apply the unitaries controlled on an orthonormal basis of maximally entangled states of $A'_1A'_2$. Then, A'_1A are decoupled from E; i.e., we achieved catalytic decoupling with remainder system size $\log|A'_2| = k$ [51].

As a second application, we discuss quantum state merging [1] in whose context decoupling was originally introduced [3,4]. In the communication task of state merging, Alice, Bob and a referee share initially a pure quantum state ψ_{ABR}. Now, Alice has to send her system A to Bob, using as little communication as possible. Any catalytic decoupling theorem naturally leads to a quantum state merging protocol. Since the catalytic decoupling theorem is the abstraction of the results on quantum state merging in [11,42], inserting the bounds from Theorem 1, we recover the following optimal result for the communication cost $q^\varepsilon(A\rangle B)_\varrho$ of merging A to B (up to error $\varepsilon > 0$, see Supplemental Material, Definition 31 [22] for a formal definition).

Proposition 3: (Coherent quantum state merging) Let ϱ_{ABR} be a pure tripartite quantum state shared between Alice, Bob and a referee. If Alice and Bob have arbitrary entanglement assistance at hand, then Alice can send her system A to Bob up to error $\varepsilon > 0$ in purified distance using

$$q^\varepsilon(A\rangle B)_\varrho \lesssim \frac{1}{2} I_{\max}^{\varepsilon/3}(R;A)_\varrho \tag{21}$$

qubits of quantum communication, where $\lesssim$ stands for smaller or equal up to terms $\mathcal{O}(\log\log|A| + \log(1/\delta))$.

We note that, in the asymptotic limit, standard decoupling is sufficient to obtain

$$\lim_{n\to\infty} \frac{1}{n} q^\varepsilon(A^n\rangle B^n)_{\varrho^{\otimes n}} = \frac{1}{2} I(R:A)_\varrho, \tag{22}$$

which is optimal [4]. For the general setup, there is an issue known as entanglement spread [52], and for the proof of Proposition 3, we make use of catalytic decoupling and Uhlmann's theorem [53]. In the following, we present a proof sketch but defer the full argument to the Supplemental Material [22]. Setting $\delta = \varepsilon/6$ in Theorem 1 shows that there exists an ancilla state $\varrho_{A'}$ and a unitary $U_{AA'\to A_1A_2}$ such that A_1 is $\varepsilon/2$-decoupled from R and

$$\log|A_2| \lesssim \frac{1}{2} I_{\max}^{\varepsilon/3}(R:A)_\varrho. \tag{23}$$

Now, Alice and Bob take a pure entangled state $\varrho_{A'B'}$ where Alice's part A' is in state $\varrho_{A'}$, the required ancilla. She applies the unitary $U_{AA'\to A_1A_2}$ and sends A_2 to Bob. The decoupling condition and the triangle inequality for the purified distance imply that $P(\varrho_{A_1R}, \varrho_{A_1} \otimes \varrho_R) \leq \varepsilon$, so by Uhlmann's theorem, there exists a unitary $U_{A_2B\to ABB_1}$, acting on Bob's system such that

$$P(U\varrho_{A_1A_2BR}U^\dagger, \varrho_{A_1B_1} \otimes \varrho_{ABR}) \leq \varepsilon, \tag{24}$$

where $\varrho_{A_1B_1}$ is a purification of ϱ_{A_1}, and we omitted the subscript of U. This implies that Bob has systems AB after applying U.

Finally, we show in the Supplemental Material [22] that catalytic decoupling directly implies the achievability bound for quantum state redistribution of Anshu *et. al.* [42] (see [23,54] for alternative bounds).

Extensions.—We have analyzed how well the partial trace map $\mathcal{T}_{A\to A_1}(\cdot) = \mathrm{Tr}_{A_2}[\cdot]$ decouples. However, as originally suggested in [13], we can also study quantum channels $\mathcal{T}_{A\to B}(\cdot)$ that add noise in an arbitrary way in order to achieve decoupling. To further clarify the important difference between standard decoupling and catalytic decoupling, as well as to correct the faulty [[13], Corollary 4.2], we now give a converse for the decoupling behavior of general quantum channels.

Proposition 4: (Correction of Corollary 4.2 from [13]) If for a bipartite quantum state ϱ_{AE} and a quantum channel $\mathcal{T}_{A\to B}$,

$$\int dU_A P\left[\mathcal{T}_{A\to B}(U_A\varrho_{AE}U_A^\dagger), \mathcal{T}_{A\to B}\left(\frac{1_A}{|A|}\right) \otimes \varrho_E\right] \leq \varepsilon, \tag{25}$$

then we have

$$H_{\min}^{\varepsilon'}(A|E)_\varrho + H_{\max}^{\varepsilon}(A'|B)_\tau \gtrsim 0, \quad \text{with} \quad \varepsilon' = 15\sqrt{\varepsilon}, \tag{26}$$

where $\tau_{A'B} = \mathcal{T}_{A\to B}(\phi^+_{A'A})$ is the Choi-Jamioł kowski state.

In the Supplemental Material [22], we prove Proposition 4 starting from [[13], Theorem 4.1] (from which the faulty [[13], Corollary 4.2] was derived). The crucial difference of Proposition 4 to the erroneous version is the assumption that not only decoupling, but decoupling and randomizing are achieved:

$$\mathcal{T}_{A\to B}(\varrho_A) \otimes \varrho_E \text{ vs } \mathcal{T}_{A\to B}\left(\frac{1_A}{|A|}\right) \otimes \varrho_E. \tag{27}$$

For example, a product state $\varrho_{AE} = \varrho_A \otimes \varrho_E$ with ϱ_A pure has $H_{\min}^{\varepsilon'}(A|E)_\varrho = 0$. It is, however, already perfectly decoupled by the identity map on A, which has $H_{\max}^{\varepsilon}(A|B)_\tau \approx -\log|A|$.

In turn, applying the converse bound (26) to the partial trace map $\mathcal{T}_{A\to A_1}(\cdot) = \mathrm{Tr}_{A_2}[\cdot]$ shows that the standard decoupling bound (7), in terms of a difference of smooth max- and min-entropy, is natural if we ask for decoupling and randomizing. However, if we are not interested in randomizing, but only in decoupling, then our main result about catalytic decoupling (Theorem 1) shows that the smooth max-mutual information is the relevant measure.

Conclusion.—In this work, we introduced the notion of catalytic decoupling. As our main result, we established that the minimal remainder system size for decoupling is given by half the smooth max-mutual information. Moreover, we have shown that catalytic decoupling for general (structureless) states naturally quantifies the resources needed in the erasure of correlation model from [48] and for quantum state merging as in [11]. All of this strengthens the smooth max-mutual information as the one-shot generalization of the quantum mutual information. Finally, given that standard decoupling has already proven useful in various areas of physics (see the Refs. in the introduction), we believe that catalytic decoupling has manifold applications that remain to be explored.

We thank Mark M. Wilde for comments on an earlier version of this paper. MC and CM acknowledge financial support from the European Research Council (ERC Grant Agreement No. 337603), the Danish Council for Independent Research (Sapere Aude), the Villum Centre of Excellence for the Mathematics of Quantum Theory (QMATH), and the Swiss National Science Foundation (Project No. PP00P2-150734). MB acknowledges funding by the SNSF through a fellowship, funding by the Institute for Quantum Information and Matter (IQIM), an NSF Physics Frontiers Center (NFS Grant No. PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028), and funding support from the ARO grant for Research on Quantum Algorithms at the IQIM (W911NF-12-1-0521). FD acknowledges the support of the Czech Science Foundation (GA ČR) Project No. GA16-22211S, and FD and RR acknowledge the support of the EU FP7 under Grant Agreement No. 323970 (RAQUEL).

[*]majenz@math.ku.dk

[1] M. Horodecki, J. Oppenheim, and A. Winter, Nature (London) **436**, 673 (2005).
[2] P. Hayden, Tutorial QIP Singapore (2011), http://qip2011.quantumlah.org/images/QIPtutorial1.pdf.
[3] M. Horodecki, J. Oppenheim, and A. Winter, Commun. Math. Phys. **269**, 107 (2006).
[4] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, Proc. R. Soc. A **465**, 2537 (2009).
[5] Z. Luo and I. Devetak, IEEE Trans. Inf. Theory **55**, 1331 (2009).
[6] J. T. Yard and I. Devetak, IEEE Trans. Inf. Theory **55**, 5339 (2009).
[7] I. Devetak and J. Yard, Phys. Rev. Lett. **100**, 230501 (2008).
[8] P. Hayden, M. Horodecki, A. Winter, and J. Yard, Open Syst. Inf. Dyn. **15**, 7 (2008).
[9] F. Dupuis, Ph.D. thesis, Université de Montréal, 2009.
[10] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, IEEE Trans. Inf. Theory **60**, 2926 (2014).
[11] M. Berta, M. Christandl, and R. Renner, Commun. Math. Phys. **306**, 579 (2011).
[12] M. Berta, O. Fawzi, and S. Wehner, IEEE Trans. Inf. Theory **60**, 1168 (2014).
[13] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, Commun. Math. Phys. **328**, 251 (2014).
[14] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, Nature (London) **474**, 61 (2011).
[15] J. Aberg, Nat. Commun. **4**, 1925 (2013).
[16] A. Hutter, Master's thesis, ETH Zurich, 2011.
[17] P. Hayden and J. Preskill, J. High Energy Phys. 07 (2007) 120.
[18] S. L. Braunstein, S. Pirandola, and K. Zyczkowski, Phys. Rev. Lett. **110**, 101301 (2013).
[19] S. L. Braunstein and A. K. Pati, Phys. Rev. Lett. **98**, 080502 (2007).
[20] F. G. Brandão and M. Horodecki, Nat. Phys. **9**, 721 (2013).
[21] M. Tomamichel, *Quantum Information Processing with Finite Resources—Mathematical Foundations* (Springer International Publishing, Heidelberg, 2016).
[22] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.118.080503, which includes Refs. [23–31], for formal definition and full proofs.
[23] N. Datta, M.-H. Hsieh, and J. Oppenheim, arXiv:1409.4352.
[24] M. Tomamichel, R. Colbeck, and R. Renner, IEEE Trans. Inf. Theory **56**, 4674 (2010).
[25] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, England, 2010).
[26] R. Renner, Ph.D. thesis, ETH Zurich, 2005.
[27] R. Renner and S. Wolf, in *Proceedings of the IEEE International Symposium on Information Theory* (IEEE, New York, 2004) p. 233.
[28] R. König, R. Renner, and C. Schaffner, IEEE Trans. Inf. Theory **55**, 4337 (2009).
[29] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner, IEEE Trans. Inf. Theory **59**, 2603 (2013).
[30] H. Umegaki, in *Kodai Mathematical Seminar Reports* (Department of Mathematics, Tokyo Institute of Technology, Tokyo, 1962), pp. 59–85
[31] K. Li, Ann. Stat. **42**, 171 (2014).
[32] More precisely, this minimum is taken over subnormalized states; see Supplemental Material [22].
[33] N. Ciganovic, N. J. Beaudry, and R. Renner, IEEE Trans. Inf. Theory **60**, 1573 (2014).
[34] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal, Phys. Rev. Lett. **92**, 067902 (2004).
[35] Note that $I^{\varepsilon}_{\max}(E;A)$ and $R^{\varepsilon}(A;E)$ are *a priori* not symmetric, even though the smooth max-mutual information is approximately [as discussed in Eq. (4)]. Here, the chosen quantity gives the cleanest bounds.
[36] I. Devetak, A. Harrow, and A. Winter, IEEE Trans. Inf. Theory **54**, 4587 (2008).
[37] N. Datta and M.-H. Hsieh, New J. Phys. **13**, 093042 (2011).
[38] L. del Rio, L. Kraemer, and R. Renner, arXiv:1511.08818.
[39] A more careful discussion of this has to include the fact that a large part of the catalyst can be given back unchanged.
[40] W. van Dam and P. Hayden, Phys. Rev. A **67**, 060302 (2003).
[41] The term $\log\log|A|$ can be improved to be logarithmic in the smooth max-information when accepting a slightly worse leading order term.
[42] A. Anshu, V. K. Devabathini, and R. Jain, arXiv:1410.3031.
[43] A partial derandomization can be achieved using (approximate) unitary 2-designs [44].

[44] O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner, New J. Phys. **15**, 053022 (2013).
[45] M. Tomamichel and M. Hayashi, IEEE Trans. Inf. Theory **59**, 7693 (2013).
[46] M. Hayashi and H. Nagaoka, IEEE Trans. Inf. Theory **49**, 1753 (2003).
[47] H. Nagaoka and M. Hayashi, IEEE Trans. Inf. Theory **53**, 534 (2007).
[48] B. Groisman, S. Popescu, and A. Winter, Phys. Rev. A **72**, 032317 (2005).
[49] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[50] Also see [48] for a slightly different notion of catalyticity in this context.
[51] A similar argument can be made for arbitrary mixtures of unitaries with a nonuniform probability distribution.
[52] A. W. Harrow, in *Proceedings of the XVI International Congress on Mathematical Physics* (World Scientific, Singapore, 2009).
[53] A. Uhlmann, Ann. Phys. (N.Y.) **497**, 524 (1985).
[54] M. Berta, M. Christandl, and D. Touchette, IEEE Trans. Inf. Theory **62**, 1425 (2016).