SUPPLEMENTAL MATERIAL:

Fundamental work cost of quantum processes

Philippe Faist[1,2,*] and Renato Renner[1]

[1] *Institute for Theoretical Physics, ETH Zurich, 8093 Switzerland*

[2] *Institute for Quantum Information and Matter, Caltech, Pasadena CA, 91125 USA*

(Dated: Jan. 7, 2018)

The supplemental material are structured as follows. Section I offers some preliminary definitions and notation conventions. In Section II we prove the properties of our framework outlined in the main text, namely that any trace-nonincreasing, Γ-sub-preserving map can be dilated to a trace-preserving, Γ-preserving map, as well as the equivalence of a class of battery models. Section III is dedicated to the definition and properties of the coherent relative entropy. Section IV discusses the robustness of battery states to small perturbations. Finally, Section V provides a selection of miscellaneous technical tools which are used in the rest of the paper.

I. TECHNICAL PRELIMINARIES

Let us first fix some notation. The state space of a quantum system S is a Hilbert space $\mathscr{H}_S$ (in this work, we deal exclusively with finite-dimensional spaces), the dimension of which we denote by $|S|$. A quantum state ρ_S of S is a positive semidefinite operator of unit trace acting on $\mathscr{H}_S$. A subnormalized quantum state ρ_S is defined as satisfying $\operatorname{tr}\rho_S \leqslant 1$. In this work, quantum states are normalized to unit trace unless otherwise stated. We use the notation $A \geqslant 0$ to indicate that an operator A is positive semidefinite, and $A \geqslant B$ to indicate that $A - B \geqslant 0$. For any positive semidefinite operator A_S acting on $\mathscr{H}_S$ corresponding to a system S, we denote by $\Pi_S^{A_S}$ the projector onto the support of A_S. Furthermore, all projectors considered in this work are Hermitian. For each system S with Hilbert space $\mathscr{H}_S$, we fix a basis which we denote by $\{|k\rangle_S\}$. Between any two systems A and B of same dimension (which we denote by $\mathscr{H}_A \simeq \mathscr{H}_B$ or $A \simeq B$), we may define a reference (not normalized) entangled ket $|\Phi\rangle_{A:B} := \sum_k |k\rangle_A \otimes |k\rangle_B$, as well as the partial transpose operation $t_{A\to B}(\cdot) = \operatorname{tr}_A[\Phi_{A:B}(\cdot)] = \sum_{kk'} \langle k|\cdot|k'\rangle_A\, |k'\rangle\langle k|_B$ with $\Phi_{A:B} = |\Phi\rangle\langle\Phi|_{A:B}$. Furthermore, for any operator $\Xi_A \geqslant 0$, a ket $|\Xi\rangle_{A:B}$ is a purification of Ξ_A if and only if there exists a ket $|\Phi^\Xi\rangle_{A:B}$ of the form $|\Phi^\Xi\rangle_{A:B} = \sum_j |\chi_j\rangle_A |\chi_j\rangle_B$ with orthonormal sets $\{|\chi_j\rangle_A\}, \{|\chi_j\rangle_B\}$ such that $|\Xi\rangle_{A:B} = \Xi_A^{1/2}|\Phi^\Xi\rangle_{A:B} = \Xi_B^{1/2}|\Phi^\Xi\rangle_{A:B}$ with $\Xi_A = \operatorname{tr}_B|\Xi\rangle\langle\Xi|_{A:B}$ and $\Xi_B = \operatorname{tr}_A|\Xi\rangle\langle\Xi|_{A:B}$ (Schmidt decomposition); the ket $|\Xi\rangle_{A:B}$ is normalized if and only if $\operatorname{tr}\Xi_A = 1$.

Throughout this paper, 'log' denotes the logarithm in base 2.

* phfaist@caltech.edu

A. Logical process and process matrix

We denote by a *logical process* a full description of a logical mapping of input states to output states:

Logical process. *A* logical process $\mathscr{E}_{X\to X'}$ *is a completely positive, trace-preserving map, mapping Hermitian operators on* $\mathscr{H}_X$ *to Hermitian operators on* $\mathscr{H}_{X'}$.

A logical process along with an input state may be characterized by their *process matrix*, defined as the Choi-Jamiołkowski map of the completely positive map, weighted by the input state.

Process matrix. *Let* $\mathscr{E}_{X\to X'}$ *be a logical process, and let* σ_X *be a quantum state. Let* R_X *be a system described by a Hilbert space* $\mathscr{H}_{R_X} \simeq \mathscr{H}_X$*, and let* $|\sigma\rangle_{XR_X} = \sigma_X^{1/2}|\Phi\rangle_{X:R_X}$ *be a purification of* σ_X*. Then the* process matrix *corresponding to* $\mathscr{E}_{X\to X'}$ *and* σ_X *is defined as* $\rho_{X'R_X} = \mathscr{E}_{X\to X'}(|\sigma\rangle\langle\sigma|_{XR_X})$*, where the identity process is understood on* R_X*. The process matrix is itself a normalized quantum state. The (unnormalized)* Choi matrix *of* $\mathscr{E}_{X\to X'}$ *is* $E_{X'R_X} = \mathscr{E}_{X\to X'}(\Phi_{X:R_X})$*, and satisfies* $\operatorname{tr}_{X'}(E_{X'R_X}) = \mathbb{1}_{R_X}$.

The reduced states σ_X and σ_{R_X} of $|\sigma\rangle_{X:R_X}$ on R_X and X, respectively, are related by a partial transpose operation: $\sigma_{R_X} = \operatorname{tr}_X(\sigma_{XR_X}) = t_{X\to R_X}(\sigma_X)$. Furthermore, we have the properties $\rho_{X'R_X} = \sigma_{R_X}^{1/2} E_{X'R_X}\, \sigma_{R_X}^{1/2}$ and $\rho_{R_X} = \operatorname{tr}_{X'}(\rho_{X'R_X}) = \sigma_{R_X}$.

The process matrix in return fully determines the channel $\mathscr{E}_{X\to X'}$ on the support of σ_X, allowing for a full characterization of the input state as well as the logical process on the support of the input.

B. Distance measures on states

For two quantum states ρ, σ, the trace distance is given by $D(\rho,\sigma) = \frac{1}{2}\|\rho-\sigma\|_1$, and their fidelity is defined as $F(\rho,\sigma) = \operatorname{tr}\left[(\rho^{1/2}\sigma\rho^{1/2})^{1/2}\right]$. From the fidelity one can define the *purified distance*[1] as $P(\rho,\sigma) = \sqrt{1-F^2(\rho,\sigma)}$ [3–5].

It will also prove convenient to work with subnormalized quantum states. Following Refs. [3–5], for any two subnormalized states ρ,σ, we define the *(generalized) trace distance* $D(\rho,\sigma) = \frac{1}{2}\|\rho-\sigma\|_1 + \frac{1}{2}|\operatorname{tr}\rho - \operatorname{tr}\sigma|$,

[1] The purified distance is also called *Bures distance* (up to a factor of 2) [1] and coincides to second order with the quantum *angle* [2].

the *(generalized) fidelity* $F(\rho,\sigma) = \mathrm{tr}\big[(\rho^{1/2}\sigma\rho^{1/2})^{1/2}\big] + \sqrt{(1-\mathrm{tr}\,\rho)(1-\mathrm{tr}\,\sigma)}$ and the *(generalized) purified distance* $P(\rho,\sigma) = \sqrt{1-F^2(\rho,\sigma)}$. For any two subnormalized states ρ,σ, we have the useful relation $D(\rho,\sigma) \leqslant P(\rho,\sigma) \leqslant \sqrt{2D(\rho,\sigma)}$.

C. Semidefinite programming

Semidefinite programming is a useful toolbox which brings a rich structure to a certain class of optimization problems. We follow the notation of Refs. [6, 7], where proofs to the statements given here may also be found.

Let A and B be Hermitian matrices, let $\Phi(\cdot)$ be a Hermiticity-preserving superoperator, and let $X \geqslant 0$ be the optimization variable, which is a Hermitian matrix constrained to the cone of positive semidefinite matrices. The prototypical semidefinite program is an optimization problem of the following form:[2]

$$\text{minimize}: \quad \mathrm{tr}(AX) \tag{1a}$$
$$\text{subject to}: \quad \Phi(X) \geqslant B\,. \tag{1b}$$

To any such problem corresponds another, related problem in terms of a different variable $Y \geqslant 0$:

$$\text{maximize}: \quad \mathrm{tr}(BY) \tag{2a}$$
$$\text{subject to}: \quad \Phi^\dagger(Y) \leqslant A\,. \tag{2b}$$

The first problem is called the *primal problem*, and the second, *dual problem*. Either problem is deemed *feasible* if there exists a valid choice of the optimization variable satisfying the corresponding constraint. If there exists a $X \geqslant 0$ such that $\Phi(X) - B$ is positive definite, the primal problem is said to be *strictly feasible*; the dual is *strictly feasible* if there is a $Y \geqslant 0$ such that $A - \Phi^\dagger(Y)$ is positive definite. For these two problems, we define their optimal attained values

$$\alpha = \inf\big\{\mathrm{tr}(AX) : \Phi(X) \geqslant B, X \geqslant 0\big\}\ ; \tag{3a}$$
$$\beta = \sup\big\{\mathrm{tr}(BY) : \Phi^\dagger(Y) \leqslant A, Y \geqslant 0\big\}\ , \tag{3b}$$

with the convention that $\alpha = -\infty$ if the primal problem is not feasible and $\beta = +\infty$ if the dual problem is not feasible.

For any semidefinite program, we have $\alpha \geqslant \beta$, a property called *weak duality*. This convenient relation allows us to immediately bound the optimal attained value of one of the two problems by picking any valid candidate in the other.

For some pairs of problems, we may have $\alpha = \beta$. In those cases we speak of *strong duality*. This is often the case in practice. A useful result here is Slater's theorem, providing sufficient conditions for strong duality [6, Theorem 2.2].

Theorem 1 (Slater's conditions for strong duality)**.** *Consider any semidefinite program written in the form* (1)*, and let its dual problem be given by* (2)*. Then:*

(i) if the primal problem is feasible and the dual is strictly feasible, then strong duality holds and there exists a valid choice X for the primal problem with $\mathrm{tr}(AX) = \alpha$*;*

(ii) if the dual problem is feasible and the primal is strictly feasible, then strong duality holds and there exists a valid choice Y for the dual problem with $\mathrm{tr}(BY) = \beta$.

We note that strong duality in itself doesn't necessarily imply the existence of an optimal choice of variables attaining the infimum or supremum. The existence of optimal primal or dual choices may be explicitly stated by Slater's conditions, or may be deduced by an auxiliary argument such as if the constraints force the optimization region to be compact.

[2] Several equivalent prototypical forms for semidefinite programs exist in the literature.

II. PROPERTIES OF OUR FRAMEWORK

A. Dilation of Γ-sub-preserving maps to Γ-preserving maps

For two systems X, Y, and corresponding operators $\Gamma_X, \Gamma_Y \geqslant 0$, We say that a completely positive map $\Phi_{X\to Y}$ is Γ-*sub-preserving* if it satisfies $\Phi(\Gamma_X) \leqslant \Gamma_Y$. Similarly, $\Phi_{X\to Y}$ is Γ-*preserving* if it satisfies $\Phi(\Gamma_X) = \Gamma_Y$.

From a technical point of view, trace-preserving Γ-preserving maps don't handle nicely systems of varying sizes or with different Γ operators. For example, if X and Y are systems with $\mathrm{tr}\,\Gamma_X \neq \mathrm{tr}\,\Gamma_Y$, there may clearly be no Γ-preserving map from X to Y which is also trace preserving. It turns out that, by focusing on trace-nonincreasing Γ-sub-preserving maps instead, we may circumvent the issue in a physically justified way: A trace-nonincreasing Γ-sub-preserving map can always be seen as a restriction of a Γ-preserving map on a larger system. Furthermore, the ancillas we have to include in this dilation are prepared in, or finish up in, eigenstates of the respective Γ operators.

Proposition 2 (Dilation of Γ-sub-preserving maps)**.** *Let K and L be quantum systems with corresponding Γ_K and Γ_L. Let $\tilde{\Phi}_{K\to L}$ be a trace-nonincreasing, Γ-sub-preserving map. Choose two arbitrary eigenvectors $|\mathrm{k}\rangle_K$ and $|\mathrm{l}\rangle_L$ of Γ_K and Γ_L, respectively. Then there exists a qubit system $\mathscr{H}_Q$ with corresponding Γ_Q diagonal in a basis composed of two orthogonal states $\{|\mathrm{i}\rangle_Q, |\mathrm{f}\rangle_Q\}$, such that there exists a trace-preserving, Γ-preserving map $\Phi_{KLQ\to KLQ}$ satisfying*

$$\tilde{\Phi}_{K\to L}(\cdot) = \langle \mathrm{k\,f}|\ \Phi_{KLQ\to KLQ}\big((\cdot)\otimes|\mathrm{l\,i}\rangle\langle\mathrm{l\,i}|_{LQ}\big)\ |\mathrm{k\,f}\rangle_{KQ}\,. \tag{4}$$

Here, the joint Γ operator on K, L, Q is $\Gamma_{KLQ} = \Gamma_K \otimes \Gamma_L \otimes \Gamma_Q$. Furthermore, the corresponding eigenvalues satisfy

$$\langle \mathrm{l}|\,\Gamma_L\,|\mathrm{l}\rangle_L \langle \mathrm{i}|\,\Gamma_Q\,|\mathrm{i}\rangle_Q = \langle \mathrm{k}|\,\Gamma_K\,|\mathrm{k}\rangle_K \langle \mathrm{f}|\,\Gamma_Q\,|\mathrm{f}\rangle_Q\,. \tag{5}$$

(Proof on page 4.)

This means that for any trace-nonincreasing, Γ-sub-preserving map $\tilde{\Phi}_{K\to L}$, we may find a larger system and a trace-preserving, Γ-preserving map Φ_{KLQ} such that $\tilde{\Phi}_{K\to L}$ is seen as the restriction of Φ_{KLQ} to the case where the input is fixed to $|\mathrm{l\,i}\rangle_{LQ}$ on LQ, where we only consider the subspace of the output in the support of $|\mathrm{k\,f}\rangle_{KQ}$ on KQ.

If the operators $\Gamma_K, \Gamma_L, \Gamma_Q$ come from Hamiltonians H_K, H_L, H_Q as $\Gamma_i = e^{-\beta H_i}$ for a fixed inverse temperature β, then the ancillas are prepared and left in pure energy eigenstates, specifically $|1\mathrm{i}\rangle_{LQ}$ for the input and $|\mathrm{kf}\rangle_{KQ}$ for the output. Furthermore condition (5) ensures that the total energy of the ancillas remains the same:

$$\langle 1|H_L|1\rangle_L + \langle \mathrm{i}|H_Q|\mathrm{i}\rangle_Q = \langle \mathrm{k}|H_K|\mathrm{k}\rangle_K + \langle \mathrm{f}|H_Q|\mathrm{f}\rangle_Q \, . \quad (6)$$

Note that the apparent post-selection in (4) is simply a statement about the output of Φ. This is made clear by the following corollary. For instance, if $\tilde{\Phi}$ is trace-preserving on a certain subspace, then as long as the input state is in that subspace, no post-selection occurs in effect because the output state on KQ is already exactly $|\mathrm{kf}\rangle_{KQ}$, i.e., if we were to project the output onto that state the projection would succeed with certainty. More generally, we show that performing the dilated mapping with the correct input states on the ancillary systems and without any post-selection at all, yields a process matrix which is just as close to the ideal process matrix as the one which would have been achieved with the original trace-decreasing map.

Corollary 3. *Consider the setting of Proposition 2. Then all the following statements hold.*

(a) Let P be any projector on $\mathscr{H}_K$ and assume that $\tilde{\Phi}$ is trace-preserving on the support of P, i.e., for any state τ supported on P, it holds that $\mathrm{tr}(\tilde{\Phi}(\tau)) = 1$. *Then the mapping Φ_{KLQ} given by Proposition 2 satisfies*

$$\Phi_{KLQ\to KLQ}(\tau \otimes |1\mathrm{i}\rangle\langle 1\mathrm{i}|_{LQ}) = \tilde{\Phi}_{K\to L}(\tau) \otimes |\mathrm{kf}\rangle\langle \mathrm{kf}|_{KQ} \, , \quad (7)$$

for any quantum state τ supported on P.

(b) Let σ_{KR} be any pure state between K and a reference system R. Assume that $\tilde{\Phi}$ satisfies $\mathrm{tr}(\tilde{\Phi}(\sigma_{KR})) = 1$. *Then*

$$\Phi_{KLQ\to KLQ}(\sigma_{KR} \otimes |1\mathrm{i}\rangle\langle 1\mathrm{i}|_{LQ}) = \tilde{\Phi}_{K\to L}(\sigma_{KR}) \otimes |\mathrm{kf}\rangle\langle \mathrm{kf}|_{KQ} \, . \quad (8)$$

(c) Let σ_{KR} be any pure state between K and a reference system R, and let ρ_{LR} be any quantum state. Then the mapping $\Phi_{KLQ\to KLQ}$ provided by Proposition 2 satisfies

$$P\Big(\Phi_{KLQ\to KLQ}(\sigma_{KR} \otimes |1\mathrm{i}\rangle\langle 1\mathrm{i}|_Q), \rho_{LR} \otimes |\mathrm{kf}\rangle\langle \mathrm{kf}|_{LQ}\Big) = P\big(\tilde{\Phi}_{K\to L}(\sigma_{KR}), \rho_{LR}\big) \, . \quad (9)$$

(Proof on page 5.)

B. Equivalence of battery models

Consider a logical process $\mathscr{E}_{X\to X'}$ which is not itself a free operation (i.e., $\mathscr{E}_{X\to X'}(\Gamma_X) \nleqslant \Gamma_{X'}$). It turns out that it is possible to implement this process by investing a certain amount of resources by means of an explicit battery system.

One example of such a battery system is the *information battery*. The information battery is a quantum system A of dimension which we denote by $|A|$, and for which $\Gamma_A = \mathbb{1}_A$. We require the battery to initially be prepared in a state $2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}$ and to finish in a state $2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}$ at the end, where both states are simply a state with a flat spectrum of rank 2^{λ_1} or 2^{λ_2}, and where we require that $\lambda_1, \lambda_2 \geqslant 0$ and that $2^{\lambda_1}, 2^{\lambda_2}$ are integers. If λ_1, λ_2 are themselves integers, this corresponds exactly to having λ_1 or λ_2 qubits in a fully mixed state and the remaining qubits in a pure state.

It is known that this model is equivalent to several other battery models known in the literature [8], notably the work bit (or "wit") [8, 9], or a "weight" system [10, 11]. Here, we point out that these models are in fact different instances of a more general description, making their equivalence manifest.

The most general system we have shown to be usable as a battery system is simply any system W with a arbitrary Γ_W operator, which is restricted to be in states of the form $\sigma = (P\Gamma_W P)/\operatorname{tr} P\Gamma_W$, where P is a projector which commutes with Γ_W. The "value" or "uselessness" of this state is given by the quantity $\log\operatorname{tr}(P\Gamma_W)$. The wit, the weight, as well as the information battery are all special cases of this general model.

The following proposition gives a necessary and sufficient condition as to when it is possible to overcome the Γ-sub-preservation restriction by exploiting a particular charge state change of the battery, and shows how the different battery systems are equivalent. This proves Propositions I and II of the main text.

Proposition 4. *Let $\mathscr{T}_{X\to X'}$ be a completely positive, trace-nonincreasing map. Let $y \in \mathbb{R}$. Then, the following are equivalent:*

(i) The map $\mathscr{T}_{X\to X'}$ satisfies

$$\mathscr{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\,\Gamma_{X'} \, ; \quad (10)$$

(ii) For any $\lambda_1, \lambda_2 \geqslant 0$ such that $2^{\lambda_1}, 2^{\lambda_2}$ are integers and $\lambda_1 - \lambda_2 \leqslant y$, there exists a large enough system A with $\Gamma_A = \mathbb{1}_A$ as well as a trace-nonincreasing, Γ-sub-preserving map $\Phi_{XA\to X'A}$ satisfying for all ω_X,

$$\Phi_{XA\to X'A}\Big(\omega_X \otimes \big(2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}\big)\Big) = \mathscr{T}_{X\to X'}(\omega_X) \otimes \big(2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}\big) \, ; \quad (11)$$

(iii) For a two-level system Q with two orthonormal states $|1\rangle_Q, |2\rangle_Q$, and with $\Gamma_Q = g_1|1\rangle\langle 1|_Q + g_2|2\rangle\langle 2|_Q$ chosen such that $g_2/g_1 \geqslant 2^{-y}$, there exists a trace-nonincreasing, Γ-sub-preserving map $\Phi'_{XQ\to X'Q}$ satisfying for all ω_X,

$$\Phi'_{XQ\to X'Q}\Big(\omega_X \otimes |1\rangle\langle 1|_Q\Big) = \mathscr{T}_{X\to X'}(\omega_X) \otimes |2\rangle\langle 2|_Q \, ; \quad (12)$$

(iv) Let $\tilde{Q}$ be any system and choose two orthogonal states $|1\rangle_{\tilde{Q}}, |2\rangle_{\tilde{Q}}$ which are eigenstates of $\Gamma_{\tilde{Q}}$ corresponding to respective eigenvalues g_1, g_2 which satisfy $g_2/g_1 \geqslant 2^{-y}$. Then there exists a trace-nonincreasing, Γ-sub-preserving map $\Phi'_{X\tilde{Q}\to X'\tilde{Q}}$ satisfying for all ω_X,

$$\Phi'_{X\tilde{Q}\to X'\tilde{Q}}\Big(\omega_X \otimes |1\rangle\langle 1|_{\tilde{Q}}\Big) = \mathscr{T}_{X\to X'}(\omega_X) \otimes |2\rangle\langle 2|_{\tilde{Q}} \, ; \quad (13)$$

(v) Let W_1, W_2 be quantum systems with respective corresponding Γ operators $\Gamma_{W_1}, \Gamma_{W_2}$, and let P_{W_1}, P'_{W_2} be projectors satisfying $[P_{W_1}, \Gamma_{W_1}] = 0$ and $[P'_{W_2}, \Gamma_{W_2}] = 0$, such that

$$\frac{\operatorname{tr} P'_{W_2}\Gamma_{W_2}}{\operatorname{tr} P_{W_1}\Gamma_{W_1}} \geqslant 2^{-y} . \tag{14}$$

Then there exists a Γ-sub-preserving, trace-nonincreasing map $\Phi''_{XW_1\to X'W_2}$ such that for all ω_X,

$$\Phi''_{XW_1\to X'W_2}\left(\omega_X \otimes \frac{P_{W_1}\Gamma_{W_1}P_{W_1}}{\operatorname{tr}(P_{W_1}\Gamma_{W_1})}\right) = \mathscr{T}_{X\to X'}(\omega_X) \otimes \frac{P'_{W_2}\Gamma_{W_2}P'_{W_2}}{\operatorname{tr}(P'_{W_2}\Gamma_{W_2})} . \tag{15}$$

(Proof on page 5.)

C. Proofs

Proof of Proposition 2. By definition, $\tilde{\Phi}_{K\to L}$ satisfies both $\tilde{\Phi}_{K\to L}(\Gamma_K) \leqslant \Gamma_L$ and $\tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L) \leqslant \mathbb{1}_K$. Hence, let $F_K, G_L \geqslant 0$ such that

$$\tilde{\Phi}_{K\to L}(\Gamma_K) = \Gamma_L - G_L ; \tag{16a}$$
$$\tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L) = \mathbb{1}_K - F_K . \tag{16b}$$

Let Π^Γ_L be the projector onto the support of Γ_L. We have $\Pi^\Gamma_L \leqslant \mathbb{1}_L$ and thus $\tilde{\Phi}^\dagger_{K\leftarrow L}(\Pi^\Gamma_L) \leqslant \tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L) \leqslant \mathbb{1}_K$. So define $F'_K \geqslant 0$ such that

$$\tilde{\Phi}^\dagger_{K\leftarrow L}(\Pi^\Gamma_L) = \mathbb{1}_K - F'_K . \tag{16c}$$

Let the system Q be as in the claim, with Γ_Q diagonal in the basis $\{|\mathrm{i}\rangle_Q, |\mathrm{f}\rangle_Q\}$. Define now the completely positive map

$$\begin{aligned}
&\Phi_{KLQ\to KLQ}(\cdot) = \qquad (17)\\
&\tilde{\Phi}_{K\to L}\Big(\langle \mathrm{l\,i}|\cdot|\mathrm{l\,i}\rangle_{LQ}\Big) \otimes |\mathrm{k\,f}\rangle\langle\mathrm{k\,f}|_{KQ}\\
&+\Gamma_K^{1/2}\tilde{\Phi}^\dagger_{K\leftarrow L}\Big(\big(\Gamma_L^{-1/2}\langle\mathrm{k\,f}|_{KQ}\big)(\cdot)\big(\Gamma_L^{-1/2}|\mathrm{k\,f}\rangle_{KQ}\big)\Big)\Gamma_K^{1/2} \otimes |\mathrm{l\,i}\rangle\langle\mathrm{l\,i}|_{LQ}\\
&+\Xi_{KL\to KL}\Big(\langle\mathrm{i}|\cdot|\mathrm{i}\rangle_Q\Big)\otimes|\mathrm{i}\rangle\langle\mathrm{i}|_Q\\
&+\Omega_{KL\to KL}\Big(\langle\mathrm{f}|\cdot|\mathrm{f}\rangle_Q\Big)\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_Q , \qquad (18)
\end{aligned}$$

with some completely positive maps $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ yet to be determined.

First, note that the property (4) is obvious for this Φ_{KLQ}, simply because $|\mathrm{i}\rangle_Q$ and $|\mathrm{f}\rangle_Q$ are orthogonal. It remains to exhibit explicit $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ such that Φ_{KLQ} is trace-preserving and Γ-preserving. Define as shorthands

$$\begin{aligned}
g_\mathrm{k} &= \langle\mathrm{k}|\Gamma_K|\mathrm{k}\rangle_K ; & g_\mathrm{l} &= \langle\mathrm{l}|\Gamma_L|\mathrm{l}\rangle_L ;\\
g_\mathrm{i} &= \langle\mathrm{i}|\Gamma_Q|\mathrm{i}\rangle_Q ; & g_\mathrm{f} &= \langle\mathrm{f}|\Gamma_Q|\mathrm{f}\rangle_Q .
\end{aligned} \tag{19}$$

Note that Condition (5) is then equivalent to

$$g_\mathrm{l}\cdot g_\mathrm{i} = g_\mathrm{k}\cdot g_\mathrm{f} , \tag{20}$$

and that this is straightforwardly satisfied for an appropriate choice of Γ_Q (and hence of $g_\mathrm{i}, g_\mathrm{f}$).

At this point, we'll derive conditions that $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ need to satisfy in order for $\Phi_{KLQ\to KLQ}$ to map Γ_{KLQ} onto itself and to be trace-preserving. Calculate

$$\begin{aligned}
&\Phi_{KLQ\to KLQ}(\Gamma_{KLQ})\\
&\quad= g_\mathrm{l}g_\mathrm{i}\,\tilde{\Phi}_{K\to L}(\Gamma_K)\otimes|\mathrm{k\,f}\rangle\langle\mathrm{k\,f}|_{KQ}\\
&\qquad+g_\mathrm{k}g_\mathrm{f}\,\Gamma_K^{1/2}\tilde{\Phi}^\dagger_{K\leftarrow L}(\Pi^\Gamma_L)\,\Gamma_K^{1/2}\otimes|\mathrm{l\,i}\rangle\langle\mathrm{l\,i}|_{LQ}\\
&\qquad+g_\mathrm{i}\Xi_{KL\to KL}(\Gamma_{KL})\otimes|\mathrm{i}\rangle\langle\mathrm{i}|_Q+g_\mathrm{f}\Omega_{KL\to KL}(\Gamma_{KL})\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_Q\\
&\quad=|\mathrm{f}\rangle\langle\mathrm{f}|_Q\otimes\big[g_\mathrm{l}g_\mathrm{i}(\Gamma_L-G_L)\otimes|\mathrm{k}\rangle\langle\mathrm{k}|_K+g_\mathrm{f}\Omega_{KL\to KL}(\Gamma_{KL})\big]\\
&\qquad+|\mathrm{i}\rangle\langle\mathrm{i}|_Q\otimes\Big[g_\mathrm{k}g_\mathrm{f}\Gamma_K^{1/2}(\mathbb{1}_K-F'_K)\Gamma_K^{1/2}\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_{KQ}\\
&\qquad\qquad+g_\mathrm{i}\Xi_{KL\to KL}(\Gamma_{KL})\Big] .
\end{aligned}\tag{21}$$

We see that in order for this last expression to equal $\Gamma_{KLQ} = g_\mathrm{f}|\mathrm{f}\rangle\langle\mathrm{f}|_Q\otimes\Gamma_{KL} + g_\mathrm{i}|\mathrm{i}\rangle\langle\mathrm{i}|_Q\otimes\Gamma_{KL}$, we need that the terms in square brackets above obey

$$g_\mathrm{l}g_\mathrm{i}(\Gamma_L-G_L)\otimes|\mathrm{k}\rangle\langle\mathrm{k}|_K+g_\mathrm{f}\Omega_{KL\to KL}(\Gamma_{KL}) = g_\mathrm{f}\Gamma_{KL} ; \tag{22a}$$
$$g_\mathrm{k}g_\mathrm{f}\Gamma_K^{1/2}(\mathbb{1}_K-F'_K)\Gamma_K^{1/2}\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_{KQ}+g_\mathrm{i}\Xi_{KL\to KL}(\Gamma_{KL}) = g_\mathrm{i}\Gamma_{KL} . \tag{22b}$$

On the other hand, the adjoint map of $\Phi_{KLQ\to KLQ}$ is relatively straightforward to identify as

$$\begin{aligned}
&\Phi^\dagger_{KLQ\leftarrow KLQ}(\cdot) =\\
&\tilde{\Phi}^\dagger_{K\leftarrow L}\Big(\langle\mathrm{k\,f}|\cdot|\mathrm{k\,f}\rangle_{KQ}\Big)\otimes|\mathrm{l\,i}\rangle\langle\mathrm{l\,i}|_{LQ}\\
&+\Gamma_L^{-1/2}\tilde{\Phi}_{K\to L}\Big(\big(\Gamma_K^{1/2}\langle\mathrm{l\,i}|_{LQ}\big)(\cdot)\big(\Gamma_K^{1/2}|\mathrm{l\,i}\rangle_{LQ}\big)\Big)\Gamma_L^{-1/2}\otimes|\mathrm{k\,f}\rangle\langle\mathrm{k\,f}|_{KQ}\\
&+\Xi^\dagger_{KL\leftarrow KL}\Big(\langle\mathrm{i}|\cdot|\mathrm{i}\rangle_Q\Big)\otimes|\mathrm{i}\rangle\langle\mathrm{i}|_Q\\
&+\Omega^\dagger_{KL\leftarrow KL}\Big(\langle\mathrm{f}|\cdot|\mathrm{f}\rangle_Q\Big)\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_Q .
\end{aligned}\tag{23}$$

We may thus now derive the conditions on $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ for $\Phi_{KLQ\to KLQ}$ to be trace-preserving. Specifically, we need to ensure that $\Phi^\dagger_{KLQ\leftarrow KLQ}(\mathbb{1}_{KLQ}) = \mathbb{1}_{KLQ}$. A calculation gives us

$$\begin{aligned}
&\Phi^\dagger_{KLQ\leftarrow KLQ}(\mathbb{1}_{KLQ})\\
&\quad=\tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L)\otimes|\mathrm{l\,i}\rangle\langle\mathrm{l\,i}|_{LQ}\\
&\qquad+\Gamma_L^{-1/2}\tilde{\Phi}_{K\to L}(\Gamma_K)\Gamma_L^{-1/2}\otimes|\mathrm{k\,f}\rangle\langle\mathrm{k\,f}|_{KQ}\\
&\qquad+\Xi^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL})\otimes|\mathrm{i}\rangle\langle\mathrm{i}|_Q\\
&\qquad+\Omega^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL})\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_Q .\\
&\quad=|\mathrm{f}\rangle\langle\mathrm{f}|_Q\otimes\Big[\Gamma_L^{-1/2}(\Gamma_L-G_L)\Gamma_L^{-1/2}\otimes|\mathrm{k}\rangle\langle\mathrm{k}|_K\\
&\qquad+\Omega^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL})\Big]\\
&\qquad+|\mathrm{i}\rangle\langle\mathrm{i}|_Q\otimes\Big[(\mathbb{1}_K-F_K)\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L\\
&\qquad+\Xi^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL})\Big] .
\end{aligned}\tag{24}$$

Thus, for $\Phi_{KLQ\to KLQ}$ to be trace-preserving we must have

$$\Gamma_L^{-1/2}(\Gamma_L-G_L)\Gamma_L^{-1/2}\otimes|\mathrm{k}\rangle\langle\mathrm{k}|_K+\Omega^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL}) = \mathbb{1}_{KL} ; \tag{25a}$$
$$(\mathbb{1}_K-F_K)\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L+\Xi^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL}) = \mathbb{1}_{KL} . \tag{25b}$$

Let us now explicitly construct an $\Xi_{KL\to KL}$ which satisfies both (22b) and (25b). These conditions may be written as

$$\Xi_{KL\to KL}(\Gamma_{KL}) = \Gamma_{KL} - g_\mathrm{l}\Gamma_K^{1/2}(\mathbb{1}_K-F'_K)\Gamma_K^{1/2}\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L =: A_{KL} ; \tag{26a}$$
$$\Xi^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL}) = \mathbb{1}_{KL} - (\mathbb{1}_K-F_K)\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L =: B_{KL} \tag{26b}$$

where we have used (20) and defined two new operators A_{KL} and B_{KL}. Observe now that since $g_\mathrm{l}\Gamma_K^{1/2}(\mathbb{1}_K-F'_K)\Gamma_K^{1/2}\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L \leqslant \Gamma_K\otimes(g_\mathrm{l}|\mathrm{l}\rangle\langle\mathrm{l}|_L) \leqslant \Gamma_{KL}$, we have that $A_{KL}\geqslant 0$. Similarly, $(\mathbb{1}_K-F_K)\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L \leqslant \mathbb{1}_{KL}$ and hence $B_{KL}\geqslant 0$.

Let ξ_{KL} be a quantum state defined as follows: If $\operatorname{tr} A_{KL} \neq 0$, then $\xi_{KL} = A_{KL}/\operatorname{tr} A_{KL}$; else $\xi_{KL} = \mathbb{1}_{KL}/|KL|$. Then define

$$\Xi_{KL\to KL}(\cdot) = \operatorname{tr}(B_{KL}(\cdot))\,\xi_{KL}\ . \tag{27}$$

We then have

$$\Xi^{\dagger}_{KL\leftarrow KL}(\mathbb{1}_{KL}) = \operatorname{tr}(\xi_{KL}\mathbb{1}_{KL})\,B_{KL} = B_{KL}\ , \tag{28}$$

thus satisfying condition (26b). On the other hand we have

$$\Xi_{KL\to KL}(\Gamma_{KL}) = \operatorname{tr}[B_{KL}\Gamma_{KL}]\,\xi_{KL}\ , \tag{29}$$

which we need to show equals A_{KL} to satisfy condition (26a). Consider first the case where $\operatorname{tr} A_{KL} = 0$ and hence $A_{KL} = 0$. Then $\Gamma_{KL} = g_1\Gamma_K^{1/2}(\mathbb{1}_K - F'_K)\Gamma_K^{1/2}\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L$, and hence $\Gamma_L = g_1|\mathrm{l}\rangle\langle\mathrm{l}|_L$ and $F'_K = 0$. Since $\tilde{\Phi}^{\dagger}_{K\leftarrow L}(\Pi^{\Gamma}_L) \leqslant \tilde{\Phi}^{\dagger}_{K\leftarrow L}(\mathbb{1}_L)$, we have $F_K \leqslant F'_K$ and thus $F_K = 0$. Then $B_{KL} = \mathbb{1}_K\otimes(\mathbb{1}_L - |\mathrm{l}\rangle\langle\mathrm{l}|_L)$. Thus, B_{KL} has no overlap with $\Gamma_{KL} = \Gamma_K\otimes(g_1|\mathrm{l}\rangle\langle\mathrm{l}|_L)$ and $(29) = 0 = A_{KL}$ as required. Now consider the case where $\operatorname{tr} A_{KL} \neq 0$. We have

$$\begin{aligned}\operatorname{tr} A_{KL} &= \operatorname{tr}\Gamma_{KL} - g_1\operatorname{tr}\left[(\mathbb{1}_K - F'_K)\Gamma_K\right]\\ &= \operatorname{tr}\Gamma_{KL} - g_1\operatorname{tr}\left[\tilde{\Phi}^{\dagger}_{K\leftarrow L}(\Pi^{\Gamma}_L)\,\Gamma_K\right]\\ &= \operatorname{tr}\Gamma_{KL} - g_1\operatorname{tr}\left[\Pi^{\Gamma}_L\tilde{\Phi}_{K\to L}(\Gamma_K)\right]\ .\end{aligned} \tag{30}$$

Now, because $\tilde{\Phi}_{K\to L}(\Gamma_K) \leqslant \Gamma_L$, the operator $\tilde{\Phi}_{K\to L}(\Gamma_K)$ must lie within the support of Γ_L. Thus the projector in the last term of (30) has no effect and can be replaced by an identity operator. We then have

$$\begin{aligned}(30) &= \operatorname{tr}\Gamma_{KL} - g_1\operatorname{tr}\left[\mathbb{1}_L\tilde{\Phi}_{K\to L}(\Gamma_K)\right]\\ &= \operatorname{tr}\Gamma_{KL} - g_1\operatorname{tr}\left[\tilde{\Phi}^{\dagger}_{K\leftarrow L}(\mathbb{1}_L)\,\Gamma_K\right]\\ &= \operatorname{tr}\Gamma_{KL} - g_1\operatorname{tr}\left[(\mathbb{1}_K - F_K)\Gamma_K\right]\\ &= \operatorname{tr}\Gamma_{KL} - \operatorname{tr}\left[(\mathbb{1}_K - F_K)\otimes|\mathrm{l}\rangle\langle\mathrm{l}|_L\Gamma_{KL}\right]\\ &= \operatorname{tr}(B_{KL}\Gamma_{KL})\ .\end{aligned} \tag{31}$$

Since $\operatorname{tr}(B_{KL}\Gamma_{KL}) = \operatorname{tr}(A_{KL})$, we have $(29) = A_{KL}$ as required. We have thus constructed $\Xi_{KL\to KL}$ such that it satisfies conditions (22b) and (25b).

Let's now proceed analogously for $\Omega_{KL\to KL}$. We can rewrite conditions (22a) and (25a) as

$$\Omega_{KL\to KL}(\Gamma_{KL}) = \Gamma_{KL} - g_{\mathrm{k}}|\mathrm{k}\rangle\langle\mathrm{k}|_K\otimes(\Gamma_L - G_L) =: C_{KL}\ ; \tag{32}$$

$$\Omega^{\dagger}_{KL\leftarrow KL}(\mathbb{1}_{KL}) = \mathbb{1}_{KL} - |\mathrm{k}\rangle\langle\mathrm{k}|_K\otimes\Gamma_L^{-1/2}(\Gamma_L - G_L)\Gamma_L^{-1/2} =: D_{KL}\ , \tag{33}$$

defining the operators C_{KL} and D_{KL}. We have $g_{\mathrm{k}}|\mathrm{k}\rangle\langle\mathrm{k}|_K\otimes(\Gamma_L - G_L) \leqslant \Gamma_{KL}$ and thus $C_{KL} \geqslant 0$. Also $\Gamma_L^{-1/2}(\Gamma_L - G_L)\Gamma_L^{-1/2} \leqslant \mathbb{1}_L$ and thus $D_{KL} \geqslant 0$. Proceeding as for $\Xi_{KL\to KL}$, let ω_{KL} be a quantum state defined as $\omega_{KL} = C_{KL}/\operatorname{tr} C_{KL}$ if $\operatorname{tr} C_{KL} \neq 0$ or $\omega_{KL} = \mathbb{1}_{KL}/|KL|$ otherwise. Define

$$\Omega_{KL\to KL}(\cdot) = \operatorname{tr}(D_{KL}(\cdot))\,\omega_{KL}\ . \tag{34}$$

Then

$$\Omega^{\dagger}_{KL\leftarrow KL}(\mathbb{1}_{KL}) = \operatorname{tr}(\omega_{KL}\mathbb{1}_{KL})\,D_{KL} = D_{KL}\ , \tag{35}$$

which satisfies (33). On the other hand, we have

$$\Omega_{KL\to KL}(\Gamma_{KL}) = \operatorname{tr}(D_{KL}\Gamma_{KL})\,\omega_{KL}\ , \tag{36}$$

which we need to show is equal to C_{KL}. First consider the case where $\operatorname{tr} C_{KL} = 0$, i.e. $C_{KL} = 0$. Then $\Gamma_{KL} = g_{\mathrm{k}}|\mathrm{k}\rangle\langle\mathrm{k}|_K\otimes(\Gamma_L - G_L)$, implying that $\Gamma_K = g_{\mathrm{k}}|\mathrm{k}\rangle\langle\mathrm{k}|_K$ and $G_L = 0$. Then $D_{KL} = \mathbb{1}_{KL} - |\mathrm{k}\rangle\langle\mathrm{k}|_K\otimes\Pi^{\Gamma_L}_L = \mathbb{1}_{KL} - \Pi^{\Gamma_{KL}}_{KL}$, and thus D_{KL} has no overlap with Γ_{KL}. It follows that $(36) = 0 = C_{KL}$ as required. Now assume that $\operatorname{tr} C_{KL} \neq 0$. Then

$$\begin{aligned}\operatorname{tr}(D_{KL}\Gamma_{KL}) &= \operatorname{tr}\Gamma_{KL} - g_{\mathrm{k}}\operatorname{tr}\left((\Gamma_L - G_L)\Pi^{\Gamma}_L\right)\\ &= \operatorname{tr}\Gamma_{KL} - g_{\mathrm{k}}\operatorname{tr}(\Gamma_L - G_L) = \operatorname{tr} C_{KL}\ ,\end{aligned} \tag{37}$$

where the projector Π^{Γ}_L has no effect in the second expression since $\Gamma_L - G_L$ is entirely contained within the support of Γ_L. Then again $(36) = C_{KL}$ as required.

We have thus constructed a completely positive, trace preserving map $\Phi_{KLQ\to KLQ}$ which maps Γ_{KLQ} onto itself and which satisfies (4). This concludes the proof. ■

Proof of Corollary 3. The proofs of (a) and (b) exploit the following fact: If a bipartite (normalized) quantum state ζ_{AB} satisfies $\langle\chi|\zeta_{AB}|\chi\rangle_B = \zeta'_A$ for some pure state $|\chi\rangle_B$ and a (normalized) quantum state ζ'_A, then $\zeta_{AB} = \zeta'_A\otimes|x\rangle\langle x|_B$. [Indeed, ζ_{AB} must lie within the support of the projector $\mathbb{1}_A\otimes|\chi\rangle\langle\chi|_B$ since $\operatorname{tr}((\mathbb{1}_{AB} - (\mathbb{1}_A\otimes|\chi\rangle\langle\chi|_B))\,\zeta_{AB}) = \operatorname{tr}(\zeta_{AB}) - \operatorname{tr}(\zeta'_A) = 0$, and hence $\zeta_{AB} = (\mathbb{1}_A\otimes|\chi\rangle\langle\chi|_B)\zeta_{AB}(\mathbb{1}_A\otimes|\chi\rangle\langle\chi|_B) = \zeta'_A\otimes|\chi\rangle\langle\chi|_B$.]

Proof of (a): For any quantum state τ supported on P, we have by assumption $\langle\mathrm{k,f}|\Phi_{KLQ\to KLQ}(\tau_K\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ})|\mathrm{k,f}\rangle = \tilde{\Phi}_{K\to L}(\tau_K)$ with $\operatorname{tr}(\langle\mathrm{k,f}|\Phi_{KLQ\to KLQ}(\tau_K\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ})|\mathrm{k,f}\rangle) = \operatorname{tr}(\tilde{\Phi}_{K\to L}(\tau_K)) = 1$. Using the fact above we conclude that $\Phi_{KLQ\to KLQ}(\tau_K\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ}) = \tilde{\Phi}_{K\to L}(\tau_K)\otimes|\mathrm{k\,f}\rangle\langle\mathrm{k\,f}|_{KQ}$.

Proof of (b): By assumption, $\langle\mathrm{k,f}|\Phi_{KLQ\to KLQ}(\sigma_{KR}\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ})|\mathrm{k,f}\rangle = \tilde{\Phi}_{K\to L}(\sigma_{KR})$ with $\operatorname{tr}(\langle\mathrm{k,f}|\Phi_{KLQ\to KLQ}(\sigma_{KR}\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ})|\mathrm{k,f}\rangle) = \operatorname{tr}(\tilde{\Phi}_{K\to L}(\sigma_{KR})) = 1$. Again a straightforward application of the above fact yields $\Phi_{KLQ\to KLQ}(\sigma_{KR}\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ}) = \tilde{\Phi}_{K\to L}(\sigma_{KR})\otimes|\mathrm{k\,f}\rangle\langle\mathrm{k\,f}|_{KQ}$.

Proof of (c): We know that the mapping Φ_{KLQ} provided by Proposition 2 is such that $\langle\mathrm{k\,f}|\Phi_{KLQ}(\sigma_{KR}\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ})|\mathrm{k\,f}\rangle = \tilde{\Phi}_{K\to L}(\sigma_{KR})$. We exploit the fact that the fidelity does not change if we project one state onto the support of the other state [indeed, we have $F(\sigma,\rho) = \operatorname{tr}\left[(\sigma^{1/2}\rho\sigma^{1/2})^{1/2}\right] = \operatorname{tr}\left[(\sigma^{1/2}\Pi^{\sigma}\rho\,\Pi^{\sigma}\sigma^{1/2})^{1/2}\right] = F(\sigma,\Pi^{\sigma}\rho\Pi^{\sigma})$]. This means in turn that $F\left(\Phi_{KLQ}(\sigma_{KR}\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ}),\rho_{LR}\otimes|\mathrm{k f}\rangle\langle\mathrm{k f}|_{KQ}\right) = F\left((\mathbb{1}_{LR}\otimes|\mathrm{k f}\rangle\langle\mathrm{k f}|_{KQ})\ \Phi_{KLQ}(\sigma_{KR}\otimes|\mathrm{1i}\rangle\langle\mathrm{1i}|_{LQ})\ (\mathbb{1}_{LR}\otimes|\mathrm{k f}\rangle\langle\mathrm{k f}|_{KQ})\right., \left.\rho_{LR}\otimes|\mathrm{k f}\rangle\langle\mathrm{k f}|_{KQ}\right) = F\left(\tilde{\Phi}_{K\to L}(\sigma_{KR})\otimes|\mathrm{k f}\rangle\langle\mathrm{k f}|_{KQ},\rho_{LR}\otimes|\mathrm{k f}\rangle\langle\mathrm{k f}|_{KQ}\right) = F\left(\tilde{\Phi}_{K\to L}(\sigma_{KR}),\rho_{LR}\right)$. This proves the claim since the purified distance is defined in terms of the fidelity. ■

Proof of Proposition 4. The proof consists in showing (i)⇒(v)⇒(iv)⇒(iii)⇒(i) as well as (v)⇒(ii)⇒(i).

(i)⇒(v): By assumption we have $\mathscr{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\Gamma_{X'}$. Let $\Gamma_{W_1},\Gamma_{W_2}$ and P_{W_1},P'_{W_2} satisfy the assumptions in the claim (v), and define the shorthands

$$\sigma^{(1)}_{W_1} = \frac{P_{W_1}\Gamma_{W_1}P_{W_1}}{\operatorname{tr}(P_{W_1}\Gamma_{W_1})}\ ; \qquad \sigma^{(2)}_{W_2} = \frac{P'_{W_2}\Gamma_{W_2}P'_{W_2}}{\operatorname{tr}(P'_{W_2}\Gamma_{W_2})}\ . \tag{38}$$

Define the map

$$\Phi''_{XW_1\to X'W_2}(\cdot) = \mathscr{T}_{X\to X'}\left[\operatorname{tr}_{W_1}\left(P_{W_1}(\cdot)\right)\right]\otimes\sigma^{(2)}_{W_2}\ . \tag{39}$$

This map is completely positive by construction, and is trace nonincreasing because it is a composition of trace nonincreasing maps. We need to show that it is Γ-sub-preserving. We have

$$\begin{aligned}\Phi''_{XW_1\to X'W_2}\left(\Gamma_X\otimes\Gamma_{W_1}\right) &= \left(\operatorname{tr}P_{W_1}\Gamma_{W_1}\right)\cdot\mathscr{T}_{X\to X'}(\Gamma_X)\otimes\sigma^{(2)}_{W_2}\\ &\leqslant 2^{-y}\frac{\operatorname{tr}P_{W_1}\Gamma_{W_1}}{\operatorname{tr}P'_{W_2}\Gamma_{W_2}}\cdot\Gamma_{X'}\otimes\left(P'_{W_2}\Gamma_{W_2}P'_{W_2}\right)\\ &\leqslant \Gamma_{X'}\otimes\Gamma_{W_2}\ ,\end{aligned} \tag{40}$$

using the fact that $P'_{W_2}\Gamma_{W_2}P'_{W_2} \leqslant \Gamma_{W_2}$ since Γ_{W_2} commutes with P'_{W_2}.

(v)⇒(iv): This special case follows directly from (v) with $W_1 = W_2 = \tilde{Q}$, $\Gamma_{W_1} = \Gamma_{W_2} = \Gamma_{\tilde{Q}}$ and by choosing $P_{W_1} = |1\rangle\langle 1|_{\tilde{Q}}$, $P'_{W_2} = |2\rangle\langle 2|_{\tilde{Q}}$. Note that $g_1 = \operatorname{tr}P_{W_1}\Gamma_{W_1}$ and $g_2 = \operatorname{tr}P'_{W_2}\Gamma_{W_2}$ and hence indeed $\left(\operatorname{tr}P'_{W_2}\Gamma_{W_2}\right)/\left(\operatorname{tr}P_{W_1}\Gamma_{W_1}\right) = g_2/g_1 \geqslant 2^{-y}$.

(iv)⇒(iii): This is a trivial special case of (iv).

(iii)⇒(i): Pick $\Gamma_Q, |1\rangle_Q, |2\rangle_Q, g_1, g_2$ such that they satisfy the assumptions of (iii) as well as $g_2/g_1 = 2^{-y}$ and let $\Phi'_{XQ\to X'Q}$ be the corresponding mapping. Observe that for any ω_X

$$\mathscr{T}_{X\to X'}(\omega_X) = \langle 2\,|\,\Phi'_{XQ\to X'Q}\big(\omega_X\otimes|1\rangle\langle 1|_Q\big)\,|\,2\rangle_Q\ . \tag{41}$$

Plugging in $\omega_X = \Gamma_X$, and using the fact that $g_1|1\rangle\langle 1|_Q \leqslant \Gamma_Q$ and that $\Phi'_{XQ\to X'Q}$ is Γ-sub-preserving,

$$\begin{aligned}\mathscr{T}_{X\to X'}(\Gamma_X) &\leqslant \langle 2\,|\,g_1^{-1}\cdot\Phi'_{XQ\to X'Q}\big(\Gamma_X\otimes\Gamma_Q\big)\,|\,2\rangle_Q\ .\\ &\leqslant \langle 2\,|\,g_1^{-1}\cdot\Gamma_{X'}\otimes\Gamma_Q\,|\,2\rangle_Q\ .\\ &= \frac{g_2}{g_1}\cdot\Gamma_{X'} = 2^{-y}\,\Gamma_{X'}\ .\end{aligned} \tag{42}$$

(v)⇒(ii): This is in fact another special case of (v). Let λ_1, λ_2 such that $\lambda_1 - \lambda_2 \leqslant y$ and that $2^{\lambda_1}, 2^{\lambda_2}$ are integers. Let A be any quantum system of dimension at least $\max\{2^{\lambda_1}, 2^{\lambda_2}\}$ and with $\Gamma_A = \mathbb{1}_A$. Now we use our assumption that (v) holds. Choose $W_1 = W_2 = A$, $P_{W_1} = \mathbb{1}_{2^{\lambda_1}}$, $P'_{W_2} = \mathbb{1}_{2^{\lambda_2}}$. Observe that $\mathrm{tr}(P_{W_1}\Gamma_{W_1}) = \mathrm{tr}(P_{W_1}) = 2^{\lambda_1}$ and $\mathrm{tr}(P'_{W_2}\Gamma_{W_2}) = \mathrm{tr}(P'_{W_2}) = 2^{\lambda_2}$, and hence the assumptions of (v) are satisfied. Then we know that there must exist a Γ-sub-preserving, trace-nonincreasing map $\Phi''_{XA\to X'A}$ obeying (15). The latter condition reads by plugging in our choices

$$\Phi''_{XA\to X'A}\big(\omega_X\otimes\big(2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}\big)\big) = \mathscr{T}_{X\to X'}(\omega_X)\otimes\big(2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}\big)\ , \tag{43}$$

for all ω_X. This is exactly the condition that Φ has to fulfill, and hence Φ may be taken equal to the map Φ''. It follows that (ii) is true.

(ii)⇒(i): Consider any $\lambda_1, \lambda_2 \geqslant 0$ with $\lambda_1 - \lambda_2 \leqslant y$. Let $\Phi_{XA\to X'A}$ be the corresponding Γ-sub-preserving map given by the assumption that (ii) holds. Observe that for all ω_X,

$$\mathscr{T}_{X\to X'}(\omega_X) = \mathrm{tr}_A\left\{\mathbb{1}_{2^{\lambda_2}}\ \Phi_{XA\to X'A}\big(\omega_X\otimes\big(2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}\big)\big)\right\}\ . \tag{44}$$

Plugging in $\omega_X = \Gamma_X$, and using the fact that Φ is Γ-sub-preserving,

$$\begin{aligned}\mathscr{T}_{X\to X'}(\Gamma_X) &\leqslant \mathrm{tr}_A\left\{\mathbb{1}_{2^{\lambda_2}}\ \Phi_{XA\to X'A}\big(2^{-\lambda_1}\cdot\Gamma_X\otimes\Gamma_A\big)\right\}\\ &\leqslant 2^{-\lambda_1}\cdot\mathrm{tr}_A\left\{\mathbb{1}_{2^{\lambda_2}}\ \Gamma_{X'}\otimes\Gamma_A\right\}\\ &= 2^{-(\lambda_1-\lambda_2)}\,\Gamma_{X'}\end{aligned} \tag{45}$$

Statement (i) follows by choosing a sequence of (λ_1, λ_2) with $\lambda_1 - \lambda_2 \to y$. ■

III. THE COHERENT RELATIVE ENTROPY

A. Definition and basic properties

Consider two quantum systems X and X', described by respective Γ operators Γ_X and $\Gamma_{X'}$. We would like to perform a logical process from X to X' which is described by the process matrix $\rho_{X'R_X}$, with a reference system $R_X \simeq X$. As we have seen, the process matrix uniquely identifies both an input state σ_X and a trace-nonincreasing, completely positive map $\mathscr{E}_{X\to X'}$ on the support of σ_X.

Because $\rho_{X'R_X}$ only fixes the mapping on the support of σ_X, there may be several trace-nonincreasing, completely positive maps $\mathscr{T}_{X\to X'}$ which implement this given process matrix. The coherent relative entropy is defined as the optimal battery usage achieved by a $\mathscr{T}_{X\to X'}$ with fixed process matrix $\rho_{X'R_X}$, relative to Γ operators $\Gamma_X, \Gamma_{X'}$.

In fact, we allow the implementation to fail with some fixed probability $\varepsilon \geqslant 0$ which can be chosen freely. This allow us to ignore very improbable events. Such a practice is standard in the smooth entropy framework, and it is even necessary in order to make physical statements and recover the correct asymptotic behavior [4, 5, 12]. Hence, we allow the process matrix achieved by the optimization variable $\mathscr{T}_{X\to X'}$ on the given input state to only be ε-close to the requested process matrix $\rho_{X'R_X}$.

By Proposition 4, the optimal number of extracted battery charge y of a fixed $\mathscr{T}_{X\to X'}$ is given by the condition $\mathscr{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\Gamma_{X'}$. We are then directly led to the following definition.

Coherent Relative Entropy. *For a bipartite quantum normalized state $\rho_{X'R_X}$, two positive semidefinite operators Γ_X and $\Gamma_{X'}$ such that $t_{R_X\to X}(\rho_{X'R_X})$ lies in the support of $\Gamma_X\otimes\Gamma_{X'}$, and for $\varepsilon \geqslant 0$, the* coherent relative entropy *is defined as*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \max_{\substack{\mathscr{T}_{X\to X'}(\Gamma_X)\leqslant 2^{-y}\Gamma_{X'}\\ \mathscr{T}^{\dagger}_{X\leftarrow X'}(\mathbb{1}_{X'})\leqslant\mathbb{1}_{R_X}\\ P(\mathscr{T}_{X\to X'}(\sigma_{XR_X}),\rho_{X'R_X})\leqslant\varepsilon}} y\quad , \tag{46}$$

where the optimization ranges over all $y\in\mathbb{R}$ and over all completely positive maps $\mathscr{T}_{X\to X'}$ satisfying the given conditions, and where we use the shorthand $|\sigma\rangle_{XR_X} = \rho_{R_X}^{1/2}\,|\Phi\rangle_{X:R_X}$.

If $\varepsilon = 0$, we may omit the ε superscript altogether:

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \hat{D}^{\varepsilon=0}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ . \tag{47}$$

Clearly, the coherent relative entropy is monotonously increasing in ε, as the optimization set gets larger.

We now introduce the variable $\alpha = 2^{-y}$ and denote by $T_{X'R_X}$ the Choi matrix of $\mathscr{T}_{X\to X'}$, allowing us to write the coherent relative entropy as a semidefinite program.

Proposition 5 (Semidefinite program)**.** *For a bipartite quantum normalized state $\rho_{X'R_X}$, two positive semidefinite operators Γ_X and $\Gamma_{X'}$ such that $t_{R_X\to X}(\rho_{X'R_X})$ lies in the support of $\Gamma_X\otimes\Gamma_{X'}$, and for $\varepsilon\geqslant 0$, the coherent relative entropy may be written as*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = -\log\alpha\ , \tag{48}$$

where α is the optimal solution to the following semidefinite program in terms of the variables $T_{X'R_XE}\geqslant 0, \alpha\geqslant 0$, and dual variables $\mu, \omega_{X'}, X_{R_X}\geqslant 0$, with $|\rho\rangle_{X'R_XE}$ being an arbitrary but fixed purification of $\rho_{X'R_X}$ into an environment system E of dimension at least $|E|\geqslant|X'R_X|$:

Primal problem:

minimize: α

subject to:

$$\mathrm{tr}_{X'}\big[T_{X'R_X}\big]\leqslant\mathbb{1}_{R_X}\quad : X_{R_X} \tag{49a}$$

$$\mathrm{tr}_{R_X}\big[T_{X'R_X}\,\Gamma_{R_X}\big]\leqslant\alpha\,\Gamma_{X'}\quad : \omega_{X'} \tag{49b}$$

$$\mathrm{tr}(\rho_{R_X}^{1/2}\,T_{X'R_XE}\,\rho_{R_X}^{1/2}\,\rho_{X'R_XE})\geqslant 1-\varepsilon^2\quad : \mu \tag{49c}$$

Dual problem:

maximize: $\mu\,(1-\varepsilon^2) - \mathrm{tr}(X_{R_X})$

subject to:

$$\mathrm{tr}\,[\omega_{X'}\Gamma_{X'}]\leqslant 1\quad : \alpha \tag{50a}$$

$$\mu\,\rho_{R_X}^{1/2}\,\rho_{X'R_XE}\,\rho_{R_X}^{1/2}\leqslant\mathbb{1}_E\otimes\big(\omega_{X'}\otimes\Gamma_{R_X}+\mathbb{1}_{X'}\otimes X_{R_X}\big)\quad : T_{X'R_X} \tag{50b}$$

using the shorthand $\Gamma_{R_X} = t_{X\to R_X}(\Gamma_X)$. *(Proof on page 8.)*

In the above, the reference system R_X may be understood as a "mirror system" which allows us to compare how the output and the input of the process are correlated. A classical analogue of R_X would be a memory register which stores a copy of the input. Crucially, in the semidefinite program the "mirror images" Γ_{R_X} and σ_{R_X} of Γ_X and σ_X must be constructed consistently, using the same reference basis on R_X, as encoded in the ket $|\Phi\rangle_{X:R_X}$ and the partial transpose operation $t_{X\to R_X}(\cdot)$. In the semidefinite program, Γ_X needs to be represented on R_X, and general Choi matrices of processes $\mathscr{T}_{X\to X'}$ need to be represented on $X'R_X$, so in general we need $R_X \simeq X$ even if a smaller system could hold a purification of σ_X (for instance, if σ_X is already pure). By contrast, in the definition (46) one could actually choose a more general R_X system: Given σ_X and $\mathscr{E}_{X\to X'}$, one may choose any purification $|\sigma\rangle_{XR_X}$ and correspondingly define $\rho_{X'R_X} = \mathscr{E}_{X\to X'}(\sigma_{XR_X})$.

The dual problem (50) is strictly feasible (choose, e.g., $\omega_{X'} = \mathbb{1}_{X'}/(2\,\mathrm{tr}(\Gamma_{X'}))$, $X_{R_X} = \mathbb{1}_{R_X}$ and $\mu = 1/2$), and $T_{X'R_X} = \rho_{R_X}^{-1/2}\rho_{X'R_X}\rho_{R_X}^{-1/2}$ is a feasible primal candidiate, and hence by Slater's sufficiency conditions (Theorem 1) we have that strong duality holds and there always exists optimal primal candidates. For $\varepsilon > 0$, the primal problem is also strictly feasible (choose $T_{X'R_X} = (1-\varepsilon^2/2)\,\rho_{R_X}^{-1/2}\rho_{X'R_X}\rho_{R_X}^{-1/2} + (\varepsilon^2/4)\mathbb{1}_{X'R_X}/|X'|)$, and there always exists optimal dual candidates as well. However, note that for $\varepsilon = 0$ the primal problem is not always strictly feasible (indeed, constraint (49c) is very strong and fixes the mapping $T_{X'R_X}$ on a subspace; because it must be trace-preserving on that subspace then (49a) cannot be satisfied strictly). This means that there is a possibility that there is no choice of optimal dual variables. However, since strong duality holds, there is always a sequence of choices for dual variables whose attained objective value will converge to the optimal solution of the semidefinite program.

Here are first some basic properties of the coherent relative entropy.

Proposition 6 (Trivial bounds)**.** *For any $0 \leqslant \varepsilon < 1$, we have*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant -\log\mathrm{tr}\,\Gamma_X - \log\|\Gamma_{X'}^{-1}\|_\infty - \log(1-\varepsilon^2)\ ; \quad (51a)$$

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant \log\|\Gamma_X^{-1}\|_\infty + \log\mathrm{tr}\,\Gamma_{X'} - \log(1-\varepsilon^2)\ . \quad (51b)$$

(Proof on page 9.)

In the thermodynamic version of the framework, these bounds can be understood in terms of work extraction. Suppose $\Gamma_X = \Gamma_{X'} = e^{-\beta H_X}$ with a Hamiltonian H_X and an inverse temperature β. Then $\log\|\Gamma_X^{-1}\|_\infty$ (resp. $\log\|\Gamma_{X'}^{-1}\|_\infty$) is β times the maximum energy of R (resp. X'), and similarly, $\mathrm{tr}\,\Gamma_X$ (resp. $\mathrm{tr}\,\Gamma_{X'}$) is the partition function of X (resp. X'). The partition function is directly related to the work cost of erasure (resp. formation) of a thermal state to (resp. from) a pure energy eigenstate of zero energy. In this case, the bounds (51) correspond to the ultimate worst and best cases respectively. The ultimate worst case is that we start off in a thermal state and end up in the highest energy level, whereas the absolute best case would be to start in the highest energy eigenstate and finish in the Gibbs state.

Much like the conditional entropy and relative entropy, the coherent relative entropy is invariant under partial isometries of which $\rho_{X'R}$ and Γ operators lie in the support. In particular, the coherent relative entropy is completely oblivious to dimensions of the Hilbert spaces which are not spanned by Γ_R and $\Gamma_{X'}$.

Proposition 7 (Invariance under isometries)**.** *Let $\tilde{X}$, $\tilde{X}'$ be new systems. Suppose there exist partial isometries $V_{X\to\tilde{X}}$ and $V'_{X'\to\tilde{X}'}$ such that both $t_{R_X\to X}(\rho_{R_X})$ and Γ_X are in the support of $V_{X\to\tilde{X}}$, and both $\rho_{X'}$ and $\Gamma_{X'}$ are in the support of $V'_{X'\to\tilde{X}'}$. Then*

$$\hat{D}^{\varepsilon}_{\tilde{X}\to\tilde{X}'}((V'\otimes V)\,\rho_{X'R_X}\,(V'\otimes V)^\dagger\,\|\,V\Gamma_X V^\dagger, V'\Gamma_{X'}V'^\dagger) = \hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ . \quad (52)$$

(Proof on page 9.)

This proposition allows us to embed states in larger dimensions, as well as to show that it is invariant under simultaneous action of unitaries on the states and the Γ operators.

We may also check the behavior of the coherent relative entropy under re-scaling of the Γ operators (as the latter need not conform to any normalization). Intuitively, in the thermodynamic case where $\Gamma = e^{-\beta H}$ for a Hamiltonian H and an inverse temperature β, the transformation $\Gamma \to a\Gamma$ for a constant factor a yields the Γ operator corresponding to the modified Hamiltonian $H \to H - \beta^{-1}\ln a$, that is, a constant energy shift of all levels. Consequently, we expect that scaling the Γ operators introduces a constant shift in the coherent relative entropy, which would correspond to providing the required energy to compensate for the global change in energy.

Proposition 8 (Scaling the Γ operators)**.** *For any $0 \leqslant \varepsilon < 1$, and for real numbers $a, b > 0$,*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,a\Gamma_X, b\Gamma_{X'}) = \hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) + \log\frac{b}{a}\ . \quad (53)$$

(Proof on page 9.)

The coherent relative entropy furthermore obeys a superadditivity rule, expressing the fact that a joint implementation of two parallel independent processes cannot be worse than two separate implementations of each process.

Proposition 9 (Superadditivity for tensor products)**.** *Let systems X_1, X_1', X_2, X_2' have respective Γ operators Γ_{X_1}, $\Gamma_{X_1'}$, Γ_{X_2}, $\Gamma_{X_2'}$. Let $\rho_{X_1'R_{X_1}}$ and $\zeta_{X_2'R_{X_2}}$ be two quantum states. Then for any $\varepsilon, \varepsilon' \geqslant 0$,*

$$\hat{D}^{\varepsilon''}_{X_1X_2\to X_1'X_2'}(\rho_{X_1'R_{X_1}}\otimes\zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_1}\otimes\Gamma_{X_2}, \Gamma_{X_1'}\otimes\Gamma_{X_2'}) \geqslant \hat{D}^{\varepsilon}_{X_1\to X_1'}(\rho_{X_1'R_{X_1}}\,\|\,\Gamma_{X_1},\Gamma_{X_1'}) + \hat{D}^{\varepsilon'}_{X_2\to X_2'}(\zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_2},\Gamma_{X_2'})\ , \quad (54)$$

where $\varepsilon'' = \sqrt{\varepsilon^2 + \varepsilon'^2}$. *(Proof on page 9.)*

In contrast to measures like the min-entropy and the max-entropy, we do not have equality in general in Proposition 9. One may see this with a simple example analogous to that in Ref. [13]. Consider two qubit systems Q_i with $\Gamma_{Q_i} = g_0|0\rangle\langle 0| + g_1|1\rangle\langle 1|$ (with $i = 1,2$; $g_0 > g_1$). On a single system, performing the logical process $|0\rangle \to |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ has a different cost than the yield of $|+\rangle \to |0\rangle$.[3] However, the transition $|0\rangle \otimes |+\rangle \to |+\rangle \otimes |0\rangle$ can be achieved with a swap operation, which is perfectly Γ-preserving and hence costs no pure qubits.

A further property of the coherent relative entropy can be derived in the case where the Γ operators are restricted by projecting them onto selected eigenkets, while still having the process matrix lying in their support. Then the coherent relative entropy remains unchanged.

Proposition 10 (Restricting the Γ operators)**.** *Let P_X and $P'_{X'}$ be projectors such that $[P_X, \Gamma_X] = 0$ and $[P'_{X'}, \Gamma_{X'}] = 0$. Define $\Gamma'_X = P_X \Gamma_X P_X$ and $\Gamma'_{X'} = P'_{X'} \Gamma_{X'} P'_{X'}$. Let $\rho_{X'R_X}$ be any quantum state with support inside that of $\Gamma'_{X'} \otimes \Gamma'_{R_X}$. Then*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma'_X, \Gamma'_{X'}) = \hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \ . \quad (55)$$

(Proof on page 9.)

Another property relates the coherent relative entropy to that with respect to different Γ operators which represent "at least or at most as much weight on each state," as represented as an operator inequality. Intuitively, this proposition states that if we raise the energy levels at the input and lower the levels at the output, then the process is easier to carry out.

Proposition 11. *Let $\tilde{\Gamma}_X \geqslant 0$ and $\tilde{\Gamma}_{X'} \geqslant 0$ be such that $\tilde{\Gamma}_X \leqslant \Gamma_X$ and $\Gamma_{X'} \leqslant \tilde{\Gamma}_{X'}$. Then for any $\varepsilon \geqslant 0$,*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \tilde{\Gamma}_X, \tilde{\Gamma}_{X'}) \geqslant \hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_R, \Gamma_{X'}) \ . \quad (56)$$

(Proof on page 9.)

We further note that it is possible to rewrite the definition of the coherent relative entropy in a slightly alternative form.

Proposition 12. *The optimization problem defining the coherent relative entropy can be rewritten as*

$$2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \| \Gamma_X, \Gamma_{X'})} = \min_{T_{X'R_X}} \left\| \Gamma_{X'}^{-1/2} \operatorname{tr}_{R_X}\left[T_{X'R_X}\, t_{X\to R_X}(\Gamma_X) \right] \Gamma_{X'}^{-1/2} \right\|_\infty , \quad (57)$$

where the minimization is taken over all positive semidefinite $T_{X'R_X}$ satisfying both conditions (49a) and (49c), and for which the operator $\operatorname{tr}_{R_X}\left(T_{X'R_X}\, t_{X\to R_X}(\Gamma_X)\right)$ lies within the support of $\Gamma_{X'}$. Equivalently,

$$2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \| \Gamma_X, \Gamma_{X'})} = \min_{\mathscr{T}_{X\to X'}} \left\| \Gamma_{X'}^{-1/2} \mathscr{T}_{X\to X'}[\Gamma_X]\, \Gamma_{X'}^{-1/2} \right\|_\infty , \quad (58)$$

where the minimization is taken over all trace nonincreasing, completely positive maps $\mathscr{T}_{X\to X'}$ which satisfy $P(\mathscr{T}_{X\to X'}[\sigma_{XR_X}], \rho_{X'R_X}) \leqslant \varepsilon$ and for which $\mathscr{T}_{X\to X'}(\Gamma_X)$ lies within the support of $\Gamma_{X'}$. *(Proof on page 9.)*

Finally, we present an alternative form of the semidefinite program for the non-smooth coherent relative entropy, i.e., in the case where $\varepsilon = 0$. This version of the semidefinite program will prove useful in some later proofs.

Proposition 13 (Non-smooth specialized semidefinite program)**.** *For a bipartite quantum state $\rho_{X'R_X}$, and two positive semidefinite operators Γ_X and $\Gamma_{X'}$ such that $t_{R_X\to X}(\rho_{X'R_X})$ lies in the support of $\Gamma_X \otimes \Gamma_{X'}$, the non-smooth coherent relative entropy can be written as*

$$\hat{D}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = -\log \alpha \ ; \quad (59)$$

where α is the optimal solution to the following semidefinite program in terms of the variables $T_{X'R_X} \geqslant 0, \alpha \geqslant 0$, and dual variables $Z_{X'R_X} = Z^\dagger_{X'R_X}, \omega_{X'} \geqslant 0, X_{R_X} \geqslant 0$:

Primal problem:

$$\begin{aligned} &\text{minimize:} && \alpha \\ &\text{subject to:} && \operatorname{tr}_{X'}\left[T_{X'R_X}\right] \leqslant \mathbb{1}_{R_X} && : X_{R_X} && \text{(60a)} \\ &&& \operatorname{tr}_{R_X}\left[T_{X'R_X}\, t_{X\to R_X}(\Gamma_X)\right] \leqslant \alpha\, \Gamma_{X'} && : \omega_{X'} && \text{(60b)} \\ &&& \rho_{R_X}^{1/2}\, T_{X'R_X}\, \rho_{R_X}^{1/2} = \rho_{X'R_X} && : Z_{X'R_X} && \text{(60c)} \end{aligned}$$

Dual problem:

$$\begin{aligned} &\text{maximize:} && \operatorname{tr}\left[Z_{X'R_X}\rho_{X'R_X}\right] - \operatorname{tr} X_{R_X} \\ &\text{subject to:} && \operatorname{tr}\left[\omega_{X'}\Gamma_{X'}\right] \leqslant 1 && : \alpha && \text{(61a)} \\ &&& \rho_{R_X}^{1/2}\, Z_{X'R_X}\, \rho_{R_X}^{1/2} \leqslant t_{X\to R_X}(\Gamma_X) \otimes \omega_{X'} + X_{R_X} \otimes \mathbb{1}_{X'} && : T_{X'R_X} && \text{(61b)} \end{aligned}$$

(Proof on page 10.)

Here are the proofs corresponding to this section's propositions.

Proof of Proposition 5. Write $|\sigma\rangle_{XR} = \rho_{R_X}^{1/2}|\Phi\rangle_{X:R_X}$. Let $|\rho\rangle_{X'R_XE}$ be any fixed purification of $\rho_{X'R_X}$ in an environment system E with dimension $|E| \geqslant |X'R_X|$.

First, consider any feasible candidates $T_{X'RE}, \alpha$ for (49). Then, setting $\mathscr{T}_{X\to X'}(\cdot) = \operatorname{tr}_E(T_{X'R_XE}\, t_{X\to R_X}(\cdot))$ and $y = -\log\alpha$ satisfies the requirements of (46), in particular, $F^2(\mathscr{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) \geqslant \operatorname{tr}(\rho_{R_X}^{1/2}\, T_{X'R_XE}\, \rho_{R_X}^{1/2}\, \rho_{X'R_XE}) \geqslant 1 - \varepsilon^2$ by Uhlmann's theorem because $\rho_{R_X}^{1/2}\, T_{X'R_XE}\rho_{R_X}^{1/2}$ is a purification of $\mathscr{T}_{X\to X'}(\sigma_{XR_X})$.

Let $\mathscr{T}_{X\to X'}$ and y be valid candidates in (46). Thanks to Uhlmann's theorem, there exists a pure quantum state $|\tau\rangle_{X'R_XE}$ such that

[3] That the processes $|0\rangle \to |+\rangle$ and $|+\rangle \to |0\rangle$ have different work cost and yield respectively follows from Corollary 30 below. We have $D_{\min,0}(|+\rangle\langle +| \,\|\, \Gamma) = -\log\langle +|\Gamma|+\rangle = -\log[(g_0+g_1)/2]$ and $-D_{\max}(|+\rangle\langle +| \,\|\, \Gamma) = -\log\|\Gamma^{-1/2}|+\rangle\langle +|\Gamma^{-1/2}\|_\infty = -\log\langle +|\Gamma^{-1}|+\rangle = -\log[(g_0^{-1}+g_1^{-1})/2]$ (the argument of the norm is a pure state).

$F^2(\mathscr{T}_{X\to X'}(\sigma_{XR_X}),\rho_{X'R_X}) = \mathrm{tr}(\tau_{X'R_XE}\,\rho_{X'R_XE})$. Let $V_{X\to X'E}$ be a Stinespring dilation of $\mathscr{T}_{X\to X'}$, i.e., let $V_{X\to X'E}$ satisfy $V^\dagger V \leqslant \mathbb{1}_X$ and $\mathscr{T}_{X\to X'}(\cdot) = \mathrm{tr}_E(V_{X\to X'E}\,(\cdot)\,V^\dagger)$. There exists a unitary W_E such that $|\tau\rangle_{X'R_XE} = W_E\,V_{X\to X'E}|\sigma\rangle_{X:R_X}$, since those two states are both purifications of $\mathscr{T}_{X\to X'}(\sigma_{XR_X})$. Now let $|T\rangle_{X'R_XE} = W_E\,V_{X\to X'E}|\Phi\rangle_{X:R_X}$ and $\alpha = 2^{-y}$. Then, $\mathrm{tr}_{X'E}(T_{X'R_XE}) = \mathrm{tr}_X(V^\dagger\,V\,\Phi_{X:R_X}) \leqslant \mathbb{1}_{R_X}$. Also, $\mathrm{tr}_{R_XE}[T_{X'R_XE}\,\Gamma_{R_X}] = \mathscr{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\Gamma_{X'} = \alpha\Gamma_{X'}$. Finally, $\mathrm{tr}(\rho_{R_X}^{1/2}\,T_{X'R_XE}\,\rho_{R_X}^{1/2}\,\rho_{X'R_XE}) = \mathrm{tr}(W_E\,V_{X\to X'E}\,\sigma_{XR}\,V^\dagger W_E^\dagger\,\rho_{X'R_XE}) = \mathrm{tr}(\tau_{X'R_XE}\,\rho_{X'R_XE}) = F^2(\mathscr{T}_{X\to X'}(\sigma_{XR_X}),\rho_{X'R_X}) \geqslant 1-\varepsilon^2$. ■

Proof of Proposition 6. Let $T_{X'R_XE} = (1-\varepsilon^2)\,\rho_{R_X}^{-1/2}\rho_{X'R_XE}\rho_{R_X}^{-1/2}$ and note that the condition (49c) is fulfilled. On the other hand, $\mathrm{tr}_{X'E}\,T_{X'RE} = (1-\varepsilon^2)\,\Pi_{R_X}^{\rho_{R_X}} \leqslant \mathbb{1}_{R_X}$ fulfilling (49a). Now observe that

$$\mathrm{tr}\big(T_{X'R_X}\,\Gamma_{R_X}\big) = (1-\varepsilon^2)\,\mathrm{tr}\big(\Pi_{R_X}^{\rho_{R_X}}\,\Gamma_{R_X}\big) \leqslant (1-\varepsilon^2)\,\mathrm{tr}(\Gamma_{R_X})\ , \tag{62}$$

and hence $[(1-\varepsilon^2)\,\mathrm{tr}(\Gamma_{R_X})]^{-1}\,\mathrm{tr}_{R_X}(T_{X'R_X}\,\Gamma_{R_X})$ is a subnormalized quantum state, which moreover lives within the support of $\Gamma_{X'}$ by assumption. Hence,

$$[(1-\varepsilon^2)\,\mathrm{tr}(\Gamma_{R_X})]^{-1}\,\mathrm{tr}_{R_X}(T_{X'R_X}\,\Gamma_{R_X}) \leqslant \Pi_{X'}^{\Gamma_{X'}} \leqslant \|\Gamma_{X'}^{-1}\|_\infty\,\Gamma_{X'}\ , \tag{63}$$

noting that $\|\Gamma_{X'}^{-1}\|_\infty^{-1}$ is the minimal nonzero eigenvalue of $\Gamma_{X'}$. Thus, taking $\alpha = (1-\varepsilon^2)\,\mathrm{tr}(\Gamma_{R_X})\,\|\Gamma_{X'}^{-1}\|_\infty$ satisfies (49b) yielding feasible primal candidates, which proves (51a).

Now consider the dual problem. Choosing $\omega_{X'} = (\mathrm{tr}\,\Gamma_{X'})^{-1}\,\mathbb{1}_{X'}$ immediately satisfies (50a). Using $\rho_{X'R_XE} \leqslant \mathbb{1}_{X'R_XE}$ and $\rho_{R_X} \leqslant \Pi_{R_X}^{\Gamma_{R_X}}$, we have

$$\begin{aligned}\mu\,\rho_{R_X}^{1/2}\rho_{X'R_XE}\rho_{R_X}^{1/2} &\leqslant \mu\,\Pi_{R_X}^{\Gamma_{R_X}} \otimes \mathbb{1}_{X'E}\\ &= \mu\,(\mathrm{tr}\,\Gamma_{X'})\,\mathbb{1}_E \otimes \omega_{X'} \otimes \Pi_R^{\Gamma_{R_X}}\\ &\leqslant \mu\,(\mathrm{tr}\,\Gamma_{X'})\,\|\Gamma_{R_X}^{-1}\|_\infty\,\mathbb{1}_E \otimes \omega_{X'} \otimes \Gamma_{R_X}\ ,\end{aligned} \tag{64}$$

so we choose $\mu = (\mathrm{tr}\,\Gamma_{X'})^{-1}\,\|\Gamma_{R_X}^{-1}\|_\infty^{-1}$ and $X_{R_X} = 0$ in order to fulfill (50b), which proves (51b). ■

Proof of Proposition 7. This is clearly the case, because the semidefinite problem lies entirely within the support of the isometries. Formally, any choice of variables for the original problem can be mapped in the new spaces through these partial isometries, and vice versa, and the attained values remain the same. Hence the optimal value of the problem is also the same. ■

Proof of Proposition 8. Consider the optimal primal candidiates $T_{X'R_XE}$ and α for the problem defining $2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. Then $T_{X'R_XE}$ and $ab^{-1}\alpha$ are feasible primal candidates for the semidefinite program with the scaled Γ operators. Hence

$$2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,a\Gamma_X,b\Gamma_{X'})} \leqslant \frac{a}{b}\alpha = \frac{a}{b}\,2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}\ . \tag{65}$$

The opposite direction follows by applying the same argument to the reverse situation with $\Gamma_X \to a^{-1}\Gamma_X$, $\Gamma_{X'} \to b^{-1}\Gamma_{X'}$. ■

Proof of Proposition 9. Let $T_{X_1'R_{X_1}E_1},\alpha_1$ and $T_{X_2'R_{X_2}E_2},\alpha_2$ be the optimal choice of primal variables for $2^{-\hat{D}^{\varepsilon}_{X_1\to X_1'}(\rho_{X_1'R_{X_1}}\,\|\,\Gamma_{X_1},\Gamma_{X_1'})}$ and $2^{-\hat{D}^{\varepsilon'}_{X_2\to X_2'}(\zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_2},\Gamma_{X_2'})}$, respectively. Now, let $\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}E_1E_2} = T_{X_1'R_{X_1}E_1} \otimes T_{X_2'R_{X_2}E_2}$ and $\bar{\alpha} = \alpha_1\alpha_2$. Then

$$\mathrm{tr}_{R_{X_1}R_{X_2}}\big[\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}}\,\Gamma_{R_{X_1}} \otimes \Gamma_{R_{X_2}}\big] \leqslant \alpha_1\alpha_2\,\Gamma_{X_1'} \otimes \Gamma_{X_2'}\ ; \tag{66}$$

$$\mathrm{tr}_{X_1'X_2'}\big[\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}}\big] \leqslant \mathbb{1}_{R_{X_1}} \otimes \mathbb{1}_{R_{X_2}}\ , \tag{67}$$

and

$$\begin{aligned}\mathrm{tr}\big[(\rho_{R_{X_1}}^{1/2} \otimes \zeta_{R_{X_2}}^{1/2})\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}E_1E_2}(\rho_{R_{X_1}}^{1/2} \otimes \zeta_{R_{X_2}}^{1/2})\,\cdot\\ \rho_{X_1'R_{X_1}E_1} \otimes \zeta_{X_2'R_{X_2}E_2}\big] \geqslant (1-\varepsilon^2)(1-\varepsilon'^2) \geqslant 1-\varepsilon''^2\ ,\end{aligned} \tag{68}$$

and hence this choice of variables is feasible for the tensor product problem. We then have

$$\begin{aligned}2^{-\hat{D}^{\varepsilon''}_{X_1X_2\to X_1'X_2'}(\rho_{X_1'R_{X_1}} \otimes \zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_1}\otimes\Gamma_{X_2},\Gamma_{X_1'}\otimes\Gamma_{X_2'})} &\leqslant \alpha_1\alpha_2\\ &= 2^{-\left[\hat{D}^{\varepsilon}_{X_1\to X_1'}(\rho_{X_1'R_{X_1}}\,\|\,\Gamma_{X_1},\Gamma_{X_1'}) + \hat{D}^{\varepsilon'}_{X_2\to X_2'}(\zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_2},\Gamma_{X_2'})\right]}\ .\end{aligned}$$

■

Proof of Proposition 10. Let $T_{X'R_XE}$ and α be the optimal feasible candidates for the primal semidefinite problem defining $2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. Let $T'_{X'R_XE} = (P'_{X'} \otimes P_{R_X})\,T_{X'RE}\,(P'_{X'} \otimes P_{R_X})$ and $\alpha' = \alpha$, writing $P_{R_X} = t_{X\to R_X}(P_X)$. Then

$$\begin{aligned}\mathrm{tr}_{X'}\,T'_{X'R_X} = P_{R_X}\,\mathrm{tr}_{X'}\big[P'_{X'}T_{X'R_X}\big]\,P_{R_X} &\leqslant P_{R_X}\,\mathrm{tr}_{X'}\big(T_{X'R_X}\big)\,P_{R_X}\\ &\leqslant P_{R_X} \leqslant \mathbb{1}_{R_X}\ ,\end{aligned} \tag{69}$$

satisfying (49a), and

$$\begin{aligned}&\mathrm{tr}[\rho_{R_X}^{1/2}T'_{X'R_XE}\rho_{R_X}^{1/2}\,\rho_{X'R_XE}]\\ &\qquad = \mathrm{tr}[\rho_{R_X}^{1/2}T_{X'R_XE}\rho_{R_X}^{1/2}\,\rho_{X'R_XE}] \geqslant 1-\varepsilon^2\ ,\end{aligned} \tag{70}$$

where the first equality holds because ρ_{R_X} and $\rho_{X'R_XE}$ already lie within the support of P_{R_X} and $P'_{X'} \otimes P_{R_X} \otimes \mathbb{1}_E$, respectively, and hence those projectors have no effect. Hence (49c) is fulfilled. Now we have

$$\begin{aligned}\mathrm{tr}_{R_X}[T'_{X'R_X}\,\Gamma'_{R_X}] &= \mathrm{tr}_{R_X}[(P'_{X'} \otimes P_{R_X})T_{X'R_X}(P'_{X'} \otimes P_{R_X})\,\Gamma_{R_X}]\\ &\leqslant P'_{X'}\,\mathrm{tr}_{R_X}[T_{X'R_X}\Gamma_{R_X}]P'_{X'}\\ &\leqslant P'_{X'}\,(\alpha\Gamma_{X'})\,P'_{X'} = \alpha'\Gamma'_{X'}\ ,\end{aligned} \tag{71}$$

using the fact that $\Gamma'_{R_X} \leqslant \Gamma_{R_X}$ (because $[P_{R_X},\Gamma_{R_X}] = 0$). Hence

$$2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma'_X,\Gamma'_{X'})} \leqslant 2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}\ . \tag{72}$$

Let μ, X_{R_X} and $\omega_{X'}$ be any dual feasible candidates for $2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. Now let $\mu' = \mu$, $X'_{R_X} = P_{R_X}\,X_{R_X}\,P_{R_X}$ and $\omega_{X'} = P'_{X'}\,\omega'_{X'}\,P'_{X'}$. Then $\mathrm{tr}(\omega'_{X'}\Gamma'_{X'}) = \mathrm{tr}(\omega_{X'}\Gamma'_{X'}) \leqslant \mathrm{tr}(\omega_{X'}\Gamma_{X'}) \leqslant 1$ (using the fact that $\Gamma'_{X'} \leqslant \Gamma_{X'}$ since $[\Gamma_{X'},P'_{X'}] = 0$), in accordance with (50a). Also, apply $(P'_{X'} \otimes P_{R_X})(\cdot)(P'_{X'} \otimes P_{R_X})$ onto the dual constraint (50b) to immediately see that μ', $\omega'_{X'}$ and X_{R_X} obey the new constraint with Γ'_{R_X}. Finally, the attained dual value is

$$\mu'\,(1-\varepsilon^2) - \mathrm{tr}(X'_{R_X}) \geqslant \mu\,(1-\varepsilon^2) - \mathrm{tr}(X_{R_X})\ . \tag{73}$$

Hence, we now have

$$2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma'_X,\Gamma'_{X'})} \geqslant 2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}\ , \tag{74}$$

which completes the proof. ■

Proof of Proposition 11. Let $T_{X'R_XE}$ and α be the optimal solution to the semidefinite program for $2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. They are then also feasible candidates for the semidefinite program for $2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\bar{\Gamma}_X,\bar{\Gamma}_{X'})}$, because the only condition that changes is (49b), which is obviously still satisfied. ■

Proof of Proposition 12. Let $T_{X'R_X}$ be any candidate in the primal problem. If $\mathrm{tr}_R\big(T_{X'R_X}\big)$ does not lie within the support of $\Gamma_{X'}$, then condition (49b) is not satisfied and the candidate is not primal feasible; we can hence ignore it in the minimization. Otherwise, by conjugating condition (49b) by $\Gamma_{X'}^{-1/2}$, we see that (49b) is equivalent to

$$\Gamma_{X'}^{-1/2}\,\mathrm{tr}_{R_X}\big[T_{X'R_X}\,t_{X\to R_X}(\Gamma_X)\big]\,\Gamma_{X'}^{-1/2} \leqslant \alpha\,\Pi_{X'}^{\Gamma_{X'}}\ , \tag{75}$$

which in turn is equivalent to

$$\Gamma_{X'}^{-1/2}\,\mathrm{tr}_{R_X}\big[T_{X'R_X}\,t_{X\to R_X}(\Gamma_R)\big]\,\Gamma_{X'}^{-1/2} \leqslant \alpha\,\mathbb{1}\ , \tag{76}$$

because the left hand side of (75) is entirely within the support of its right hand side. Now, the optimal α which corresponds to this fixed $T_{X'R_X}$ is given by $\|\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}\left[T_{X'R_X}\,t_{X\to R_X}(\Gamma_X)\right]\Gamma_{X'}^{-1/2}\|_\infty$. This chain of equivalences may be followed in reverse order, establishing the equivalence of the minimization problems.

The formulation in terms of channels follows immediately from the translation of one formalism to the other. ∎

Proof of Proposition 13. In the case $\varepsilon = 0$, the conditions in (46) reduce to

$$\mathscr{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\Gamma_{X'} \;;$$
$$\mathscr{T}^\dagger_{X\leftarrow X'}(\mathbb{1}_{X'}) \leqslant \mathbb{1}_X \;;$$
$$\mathscr{T}_{X\to X'}(\sigma_{XR_X}) = \rho_{X'R_X} \;,$$

where we write $|\sigma\rangle_{XR_X} = \rho_{R_X}^{1/2}|\Phi\rangle_{X:R_X}$. These conditions, when written in terms of the Choi matrix $T_{X'R_X}$ corresponding to $\mathscr{T}_{X\to X'}$, yield precisely the semidefinite program given in the claim. ∎

B. Some special cases

In this section, we look at some instructive special cases where the coherent relative entropy can be evaluated exactly.

The first proposition concerns identity mappings. It is a property that one would expect very naturally: If the process matrix corresponds to the identity mapping on the support of the input, and if the Γ operators coincide, then the process should be a free operation and should not require a battery. This property may seem like a triviality, but it is in fact not so obvious to prove: Indeed, because the coherent relative entropy is a function of the process matrix only, the implementation can choose to implement whatever process it likes on the complement of the support of the input state. In other words, this proposition tells us that there is no way to extract work by exploiting the freedom on this complementary subspace when performing the identity map on the support of σ_X.

Proposition 14 (Identity mapping)**.** *Let $\mathrm{id}_{X\to X'}$ be the identity map from a system X to a system $X'\simeq X$. Assume that $\Gamma_{X'} = \mathrm{id}_{X\to X'}(\Gamma_X)$. Let σ_X be any state on X, let $R_X\simeq X$ and $|\sigma\rangle_{XR_X} = \sigma_X^{1/2}|\Phi\rangle_{X:R_X}$, and let $|\rho\rangle_{X'R_X}$ be the process matrix of the identity process applied on σ_X, i.e. $\rho_{X'R_X} = \mathrm{id}_{X\to X'}(\sigma_{XR_X})$. Then*

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = 0\;. \tag{77}$$

Proof of Proposition 14. Let $\Phi_{X'R_X} = \mathrm{id}_{X\to X'}(\Phi_{X:R_X})$ be the unnormalized maximally entangled state on X' and R_X such that $\rho_{X'R_X} = \rho_{R_X}^{1/2}\Phi_{X'R_X}\rho_{R_X}^{1/2}$.

First we show that $\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant 0$. Consider the mapping $\mathscr{T}_{X\to X'} = \mathrm{id}_{X\to X'}$ and $y=0$, i.e., consider the identity mapping as an implementation candidate. This clearly satisfies the requirements of the maximization in (46) for $\varepsilon = 0$, and thus

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant 0\;. \tag{78}$$

We prove the reverse direction by exhibiting dual candidates for the problem given in Proposition 13. The tricky part is that there might not be an optimal choice of dual variables. The best we can do in general is to come up with a sequence of choices for dual candidates whose attained value converges to 1. For any $\mu > 0$, let

$$Z_{X'R_X} = \mu\rho_{R_X}^{-1/2}\Phi_{X'R_X}\rho_{R_X}^{-1/2}\;;\qquad \omega_{X'} = \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Pi_{X'}^{\rho_{X'}}\;. \tag{79}$$

Then $\operatorname{tr}(\omega_{X'}\Gamma_{X'}) = 1$, satisfying the dual constraint (61a). Let's now study (61b):

$$\begin{aligned}&\rho_{R_X}^{1/2}Z_{X'R_X}\rho_{R_X}^{1/2} - \Gamma_{R_X}\otimes\omega_{X'}\\ &\qquad= \mu\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X':R_X}\,\Pi_{R_X}^{\rho_{R_X}} - \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\otimes\Pi_{X'}^{\rho_{X'}}\;.\end{aligned} \tag{80}$$

The operator $\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R_X}\,\Pi_{R_X}^{\rho_{R_X}}$ is a rank-1 positive operator with support within $\Pi_{R_X}^{\rho_{R_X}}\otimes\Pi_{X'}^{\rho_{X'}}$, and its nonzero eigenvalue is given by

$$\operatorname{tr}\left(\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\right) = \operatorname{rank}\rho_{R_X}\;. \tag{81}$$

Let $r = \operatorname{rank}\rho_{R_X}$. We then have $\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}} \leqslant r\Pi_{R_X}^{\rho_{R_X}}\otimes\Pi_{X'}^{\rho_{X'}}$ and we may continue our calculation:

$$(80) \leqslant \left(\mu r\Pi_{R_X}^{\rho_{R_X}} - \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\right)\otimes\Pi_{X'}^{\rho_{X'}}\;. \tag{82}$$

Now, let P_{R_X} be the projector onto the eigenspaces associated to the positive (or null) eigenvalues of the operator $\left(\mu r\Pi_{R_X}^{\rho_{R_X}} - \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\right)$, and let

$$X_{R_X} = P_{R_X}\left(\mu r\Pi_{R_X}^{\rho_{R_X}} - \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\right)P_{R_X}\;. \tag{83}$$

Then

$$(82) \leqslant X_{R_X}\otimes\mathbb{1}_{X'}\;. \tag{84}$$

Hence, for any $\mu > 0$, this choice of dual variables satisfies the dual constraints. The value attained by this choice of variables is given by

$$\operatorname{tr}[Z_{X'R}\rho_{X'R}] - \operatorname{tr}X_{R_X} = \mu\operatorname{tr}\left[\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R}\right] - \operatorname{tr}X_{R_X}\;. \tag{85}$$

As the object $\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}$ is rank-1, we have thanks to (81) that $\operatorname{tr}\left[\left(\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\right)^2\right] = \left(\operatorname{tr}\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\right)^2 = r^2$. Then

$$\begin{aligned}(85) &= \mu r^2 - \operatorname{tr}X_{R_X}\\ &= \mu r^2 - \mu r\operatorname{tr}\left(P_{R_X}\Pi_{R_X}^{\rho_{R_X}}\right) + \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\operatorname{tr}(P_{R_X}\Gamma_{R_X})\\ &\geqslant \mu r^2 - \mu r\operatorname{tr}\left(\Pi_{R_X}^{\rho_{R_X}}\right) + \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\operatorname{tr}(P_{R_X}\Gamma_{R_X})\\ &\geqslant \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\operatorname{tr}(P_{R_X}\Gamma_{R_X})\;,\end{aligned} \tag{86}$$

recalling that $\operatorname{tr}\Pi_{R_X}^{\rho_{R_X}} = \operatorname{rank}\rho_{R_X} = r$.

Next episode: the Lemma awakens. Take $A = \mu r\Pi_{R_X}^{\rho_{R_X}}$ and $B = \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}$; Lemma 43 then asserts that there exists a constant c independent of μ such that

$$\Pi_{R_X}^{\rho_{R_X}} \leqslant P_{R_X} + \frac{c}{\mu}\mathbb{1}\;. \tag{87}$$

Hence,

$$(86) \geqslant \left(\operatorname{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\left(\operatorname{tr}\left[\Pi_{R_X}^{\rho_{R_X}}\Gamma_{R_X}\right] - \frac{c}{\mu}\operatorname{tr}\Gamma_{R_X}\right) = 1 - O(1/\mu)\;. \tag{88}$$

Taking $\mu\to\infty$ yields successive feasible dual candidates with attained objective value converging to 1, hence proving that

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant 0\;.$$ ∎

An essentially trivial proposition immediately follows from the fact that Γ-sub-preserving maps are admissible operations, and hence don't cost anything in our framework:

Proposition 15. *Let σ_X be a quantum state and let $\mathscr{E}_{X\to X'}$ be a Γ-sub-preserving logical process. With the process matrix $\rho_{X'R} = \mathscr{E}_{X\to X'}(\sigma_X^{1/2}\,\Phi_{X:R_X}\,\sigma_X^{1/2})$, we have for any $\varepsilon \geqslant 0$,*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant 0\ . \tag{89}$$

Proof of Proposition 15. The process $\mathscr{E}_{X\to X'}$ itself is a valid optimization candidate in (58), and clearly $\|\Gamma_{X'}^{-1/2}\,\mathscr{E}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\|_\infty \leqslant \|\Pi_{X'}^{\Gamma_{X'}}\|_\infty \leqslant 1$ because $\mathscr{E}_{X\to X'}$ is Γ-sub-preserving. ■

In general, the coherent relative entropy depends on the precise logical process used to map the input and output states. However, there are some classes of states for which the coherent relative entropy depends only on the input and output state.

The following proposition tells us that one may map the $\Gamma_X/\operatorname{tr}\Gamma_X$ state to the $\Gamma_{X'}/\operatorname{tr}\Gamma_{X'}$ state in however way one wants, i.e. regardless of the logical process, and yet in any case the coherent relative entropy is given by the ratio $\operatorname{tr}\Gamma_{X'}/\operatorname{tr}\Gamma_X$. This is a consequence of allowing any Γ-preserving maps to be performed for free, and this ratio comes about from the normalization of the respective input and output states.

Proposition 16. *Let P_X and $P'_{X'}$ be projectors with $[P_X,\Gamma_X]=0$ and $[P'_{X'},\Gamma_{X'}]=0$. Let $\rho_{X'R_X}$ be a bipartite quantum state with reduced states $\rho_{R_X} = t_{X\to R_X}[(P_X\Gamma_X P_X)/\operatorname{tr}(P_X\Gamma_X)]$ and $\rho_{X'} = (P'_{X'}\Gamma_{X'}P'_{X'})/\operatorname{tr}(P'_{X'}\Gamma_{X'})$. Then, for any $\varepsilon\geqslant 0$,*

$$\begin{aligned}&\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\quad = \log\operatorname{tr}(P'_{X'}\Gamma_{X'}) - \log\operatorname{tr}(P_X\Gamma_X) + \log[1/(1-\varepsilon^2)]\ .\end{aligned} \tag{90}$$

Proof of Proposition 16. Let $|\rho\rangle_{X'R_XE}$ be a purification of $\rho_{X'R_X}$ into a (large enough) system E, and consider the semidefinite program given by Proposition 5. We give feasible primal and dual candidates which achieve the same value. First, let $T_{X'R_XE} = (1-\varepsilon^2)\,\rho_{R_X}^{-1/2}\,\rho_{X'R_XE}\,\rho_{R_X}^{-1/2}$. We have $\operatorname{tr}_{X'E}(T_{X'R_XE}) = (1-\varepsilon^2)\,\Pi_{R_X}^{\rho_{R_X}} \leqslant \mathbb{1}_{R_X}$ as required by (49a). Also, since $\rho_{R_X} = P_{R_X}\Gamma_{R_X}P_{R_X}/\operatorname{tr}(P_{R_X}\Gamma_{R_X})$ and $\rho_{X'} = P'_{X'}\Gamma_{X'}P'_{X'}/\operatorname{tr}(P'_{X'}\Gamma_{X'})$, we have $\operatorname{tr}_{R_XE}(T_{X'R_XE}\,\Gamma_{R_X}) = (1-\varepsilon^2)\operatorname{tr}(P_{R_X}\Gamma_{R_X})\operatorname{tr}_{R_X}(\rho_{X'R_X}P_X) = (1-\varepsilon^2)\operatorname{tr}(P_{R_X}\Gamma_{R_X})\,\rho_{X'} \leqslant \alpha\Gamma_{X'}$, where we have defined $\alpha = (1-\varepsilon^2)\operatorname{tr}(P_{R_X}\Gamma_{R_X})/\operatorname{tr}(P'_{X'}\Gamma_{X'})$ and noting that $[P'_{X'},\Gamma_{X'}]=0$, hence satisfying (49b). Finally, we have $\operatorname{tr}\big[\rho_{R_X}^{1/2}\,T_{X'R_XE}\,\rho_{R_X}^{1/2}\,\rho_{X'R_XE}\big] = (1-\varepsilon^2)$ which satisfies (49c). This choice of primal variables is feasible, and attains the value α.

Now we exhibit feasible dual candidates. Let $\mu = \operatorname{tr}(P_{R_X}\Gamma_{R_X})/\operatorname{tr}(P'_{X'}\Gamma_{X'})$, $\omega_{X'} = P'_{X'}/\operatorname{tr}(P'_{X'}\Gamma_{X'})$ and $X_{R_X}=0$, and note that (50a) is automatically satisfied. Then, since $\rho_{X'R_XE} \leqslant \mathbb{1}_E\otimes P'_{X'}\otimes P_{R_X}$, we have

$$\begin{aligned}\mu\,\rho_{R_X}^{1/2}\,\rho_{X'R_XE}\,\rho_{R_X}^{1/2} &\leqslant \frac{\operatorname{tr}P_{R_X}\Gamma_{R_X}}{\operatorname{tr}P'_{X'}\Gamma_{X'}}\,\mathbb{1}_E\otimes P'_{X'}\otimes\rho_{R_X}\\ &\leqslant \mathbb{1}_E\otimes\omega_{X'}\otimes\Gamma_{R_X}\ ,\end{aligned} \tag{91}$$

keeping in mind that $[P_{R_X},\Gamma_{R_X}]=0$, and hence condition (50b) is satisfied. The value attained by this choice of variables is simply $\mu\,(1-\varepsilon^2) - \operatorname{tr}X_{R_X} = \alpha$, hence proving that this is the optimal solution of the semidefinite program. Calculating $-\log\alpha$ completes the proof. ■

We note that for this special type of states we have the nice expression for their relative entropy to Γ.

Proposition 17. *If $\Gamma\geqslant 0$ and P is a projector with $[P,\Gamma]=0$, then*

$$\begin{aligned}D\left(\frac{P\Gamma P}{\operatorname{tr}P\Gamma}\,\middle\|\,\Gamma\right) &= D_{\min,0}\left(\frac{P\Gamma P}{\operatorname{tr}P\Gamma}\,\middle\|\,\Gamma\right) = D_{\max}\left(\frac{P\Gamma P}{\operatorname{tr}P\Gamma}\,\middle\|\,\Gamma\right)\\ &= -\log\operatorname{tr}P\Gamma\ .\end{aligned} \tag{92}$$

Proof of Proposition 17. Write as shorthand $\rho = P\Gamma P/\operatorname{tr}P\Gamma$. Then

$$\begin{aligned}2^{D_{\max}(\rho\,\|\,\Gamma)} &= \|\Gamma^{-1/2}\,\rho\,\Gamma^{-1/2}\|_\infty\\ &= (\operatorname{tr}P\Gamma)^{-1}\,\|\Gamma^{-1/2}\,P\,\Gamma\,P\,\Gamma^{-1/2}\|_\infty\\ &= (\operatorname{tr}P\Gamma)^{-1}\,\|\Gamma^{-1/2}\,\Gamma^{1/2}\,P\,\Gamma^{1/2}\,\Gamma^{-1/2}\|_\infty\\ &= (\operatorname{tr}P\Gamma)^{-1}\ ,\end{aligned} \tag{93}$$

since $[P,\Gamma]=0$. Also, observing that $\Pi^{\rho}=P$,

$$2^{-D_{\min,0}(\rho\,\|\,\Gamma)} = \operatorname{tr}(\Pi^{\rho}\Gamma) = \operatorname{tr}(P\Gamma)\ . \tag{94}$$

The expression $D(\rho\,\|\,\Gamma)$ is thus also equal to $-\log\operatorname{tr}P\Gamma$ since we know that $D_{\min,0}(\rho\,\|\,\Gamma) \leqslant D(\rho\,\|\,\Gamma) \leqslant D_{\max}(\rho\,\|\,\Gamma)$ [14, Lemma 10]. ■

Notably, the states of the form $P\Gamma P/\operatorname{tr}(P\Gamma)$ for $[P,\Gamma]=0$ are precisely those general type of states which we allowed on battery systems in item (v) of Proposition 4.

In fact, we may prove a slightly more general version of Proposition 16 for the case $\varepsilon=0$: it suffices that the reduced state on the input is of the form $\Gamma_X/\operatorname{tr}\Gamma_X$, and then the coherent relative entropy is oblivious to any correlation between input and output, or equivalently, to which process is exactly implemented, and depends only on the reduced states on the input and the output.

Proposition 18. *Let $\rho_{X'R_X}$ such that $\operatorname{tr}_{X'}\rho_{X'R_X} = \Gamma_{R_X}/\operatorname{tr}\Gamma_{R_X}$. Then*

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = -\log\operatorname{tr}\Gamma_X - D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})\ . \tag{95}$$

Proof of Proposition 18. Take any $T_{X'R_X}$ satisfying $\rho_{R_X}^{1/2}T_{X'R_X}\rho_{R_X}^{1/2} = \rho_{X'R_X}$ and $\operatorname{tr}_{X'}T_{X'R_X} \leqslant \mathbb{1}_{R_X}$. Then since $\operatorname{tr}(\Gamma_{R_X})\,\rho_{R_X} = \Gamma_{R_X}$, we have

$$\begin{aligned}\operatorname{tr}_{R_X}(T_{X'R_X}\Gamma_{R_X}) &= \operatorname{tr}(\Gamma_{R_X})\operatorname{tr}_{R_X}(\rho_{R_X}^{1/2}T_{X'R_X}\rho_{R_X}^{1/2})\\ &= \operatorname{tr}(\Gamma_{R_X})\operatorname{tr}_{R_X}(\rho_{X'R_X}) = \operatorname{tr}(\Gamma_{R_X})\,\rho_{X'}\ ,\end{aligned} \tag{96}$$

and thus

$$\begin{aligned}-\log\big\|\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}[T_{X'R_X}\Gamma_{R_X}]\,\Gamma_{X'}^{-1/2}\big\|_\infty &= -\log\big\|\Gamma_{X'}^{-1/2}\rho_{X'}\Gamma_{X'}^{-1/2}\big\|_\infty\\ &= -D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})\ .\end{aligned} \tag{97}$$

This argument holds in particular for the optimal such $T_{X'R_X}$. ■

Remarkably, if $\operatorname{tr}_{R_X}\rho_{X'R_X} = \Gamma_{X'}/\operatorname{tr}\Gamma_{X'}$, the coherent relative entropy may still depend on the exact process, and does not necessarily reduce to a difference of input and output terms as in (95). This can be seen by considering the unitary process $\mathscr{U}$ which swaps two levels $|0\rangle,|1\rangle$, choosing $\Gamma = g_0|0\rangle\langle 0| + g_1|1\rangle\langle 1|$ (with $g_0+g_1=1$ and $g_0>g_1$) for both input and output, and using the input state $\sigma = g_1|0\rangle\langle 0| + g_0|1\rangle\langle 1|$: in this case, σ maps to Γ, but $-\log\|\Gamma^{-1/2}\,\mathscr{U}(\Gamma)\,\Gamma^{-1/2}\|_\infty = -D_{\max}(\sigma\,\|\,\Gamma)$ whereas there are processes which map σ to Γ, such as $\mathscr{T}(\cdot) = \operatorname{tr}(\Pi^{\sigma}(\cdot))\,\Gamma$, which achieve a coherent relative entropy of $D_{\min,0}(\sigma\,\|\,\Gamma)$.

C. Data processing inequality

The data processing inequality is an important property desirable for an information measure. Intuitively, it asserts that processing information cannot make it more "valuable."

In our case, the data processing inequality asserts that post-processing, or applying a map to both the output state and output Γ, may only increase the coherent relative entropy.

Proposition 19 (Data processing inequality)**.** *Let $\rho_{X'R_X}$ be a quantum state and let $\Gamma_X, \Gamma_{X'} \geqslant 0$. Let $\mathscr{F}_{X'\to X''}$ be a trace-preserving, completely positive map. Then, for any $\varepsilon \geqslant 0$,*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \leqslant \hat{D}^{\varepsilon}_{X\to X''}(\mathscr{F}_{X'\to X''}(\rho_{X'R_X}) \,\|\, \Gamma_X, \mathscr{F}_{X'\to X''}(\Gamma_{X'})) \ . \tag{98}$$

Proof of Proposition 19. Let $\mathscr{T}_{X\to X'}, y$ be optimal candidates for the optimization defining $2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \| \Gamma_X, \Gamma_{X'})}$ in (46). We construct an optimization candidate for the coherent relative entropy of the post-processed state. Let $\mathscr{T}'_{X\to X''} = \mathscr{F}_{X'\to X''} \circ \mathscr{T}_{X\to X'}$. Then $\mathscr{T}'^{\dagger}_{X\leftarrow X''}(\mathbb{1}_{X''}) = \mathscr{T}^{\dagger}_{X\leftarrow X'}(\mathscr{F}^{\dagger}_{X'\leftarrow X''}(\mathbb{1}_{X''})) \leqslant \mathbb{1}_{R_X}$ because $\mathscr{F}_{X'\to X''}$ is trace-preserving. Also, $\mathscr{T}'_{X\to X''}(\Gamma_X) \leqslant \alpha\, \mathscr{F}_{X'\to X''}(\Gamma_{X'})$. Finally, writing $|\sigma\rangle_{X R_X} = \rho^{1/2}_{R_X} |\Phi\rangle_{X:R_X}$, we have $P(\mathscr{T}'_{X\to X''}(\sigma_{XR}), \mathscr{F}_{X'\to X''}(\rho_{X'R_X})) \leqslant P(\mathscr{T}_{X\to X'}(\sigma_{XR}), \rho_{X'R_X}) \leqslant \varepsilon$. ■

The case of pre-processing, i.e. when a map is applied to the input before the actual mapping is carried out, is less clear how to formulate. Indeed, the expression $\hat{D}^{\varepsilon}_{R_{\tilde{X}}\to X'}(\mathscr{F}_{R_X\to R_{\tilde{X}}}(\rho_{X'R_X}) \,\|\, \mathscr{F}_{X\to\tilde{X}}(\Gamma_X), \Gamma_{X'})$ would correspond to the not-so-natural setting where one implements a process matrix defined by the state resulting when two logical processes are applied on both the system X of interest and the reference system R_X on a pure state $|\sigma\rangle_{XR_X}$. However, a more general statement about composing processes can be derived in the form of a chain rule, which is the subject of the next section.

D. Chain rule

If two individual processes are concatenated, what can be said of the coherent relative entropy of the combined processes? As one would expect, it turns out that the optimal battery use of implementing directly a composition of logical maps can only be better than the sum of the battery uses of the individual realizations of each map.

Proposition 20 (Chain rule)**.** *Consider three systems X, X', X'' with corresponding $\Gamma_X, \Gamma_{X'}, \Gamma_{X''} \geqslant 0$, and let $R_X \simeq X, R_{X'} \simeq X'$. Let σ_X be a quantum state. Let $\mathscr{E}^{(1)}_{X\to X'}$ and $\mathscr{E}^{(2)}_{X'\to X''}$ be two completely positive, trace-nonincreasing maps such that $\mathrm{tr}[\mathscr{E}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_X))] = 1$. Let $\varepsilon, \varepsilon' \geqslant 0$. Then:*

$$\hat{D}^{\varepsilon}_{X\to X'}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) \,\|\, \Gamma_X, \Gamma_{X'}) + \hat{D}^{\varepsilon'}_{X'\to X''}(\mathscr{E}^{(2)}_{X'\to X''}(\rho'_{X'R'_X}) \,\|\, \Gamma_{X'}, \Gamma_{X''}) \leqslant \hat{D}^{\varepsilon+\varepsilon'}_{X\to X''}(\mathscr{E}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X})) \,\|\, \Gamma_X, \Gamma_{X''}) \ , \tag{99}$$

where $|\sigma\rangle_{XR_X} = \sigma_X^{1/2} |\Phi\rangle_{X:R_X}$ and $|\rho'\rangle_{X'R_{X'}} = (\mathscr{E}^{(1)}_{X\to X'}(\sigma_X))^{1/2} |\Phi\rangle_{X':R_{X'}}$.

Proof of Proposition 20. Let $\mathscr{T}^{(1)}_{X\to X'}$, y_1 be optimal choices in (46) for $\hat{D}^{\varepsilon}_{X\to X'}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) \,\|\, \Gamma_X, \Gamma_{X'})$, and let $\mathscr{T}^{(2)}_{X\to X'}$, y_2 be optimal choices for $\hat{D}^{\varepsilon'}_{X'\to X''}(\mathscr{E}^{(2)}_{X'\to X''}(\rho'_{X'R'_X}) \,\|\, \Gamma_{X'}, \Gamma_{X''})$. Let $V_{X\to X'E}$ be a Stinespring dilation of $\mathscr{E}^{(1)}_{X\to X'}$, such that $\mathscr{E}^{(1)}_{X\to X'}(\cdot) = \mathrm{tr}_E[V_{X\to X'E}(\cdot)V^{\dagger}]$. Now, as two different purifications of $\mathscr{E}^{(1)}_{X\to X'}(\sigma_X) = \rho'_{X'}$, there must exist a partial isometry $W_{R_{X'}\to R_X E}$ such that $V_{X\to X'E}|\sigma\rangle_{XR_X} = W_{R_{X'}\to R_X E}|\rho'\rangle_{X'R_{X'}}$. Define $\mathscr{F}_{R_{X'}\to R_X}(\cdot) = \mathrm{tr}_E(W_{R_{X'}\to R_X E}(\cdot)W^{\dagger})$, and note that $\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \mathscr{F}_{R_{X'}\to R_X}(\rho'_{X'R_{X'}})$. Now, let $\mathscr{T}_{X\to X''} = \mathscr{T}^{(2)}_{X'\to X''} \circ \mathscr{T}^{(1)}_{X\to X'}$, and note that

$$\begin{aligned} &P\big[\mathscr{T}_{X\to X''}(\sigma_{XR_X}), \mathscr{E}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\big] \\ &\quad\leqslant P\big[\mathscr{T}^{(2)}_{X'\to X''}(\mathscr{T}^{(1)}_{X\to X'}(\sigma_{XR_X})), \mathscr{T}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\big] \\ &\qquad + P\big[\mathscr{T}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X})), \mathscr{E}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\big] \\ &\quad\leqslant P\big[\mathscr{T}^{(1)}_{X\to X'}(\sigma_{XR_X}), \mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X})\big] \\ &\qquad + P\big[\mathscr{T}^{(2)}_{X'\to X''}(\rho'_{X'R_{X'}}), \mathscr{E}^{(2)}_{X'\to X''}(\rho'_{X'R_{X'}})\big] \\ &\quad\leqslant \varepsilon + \varepsilon' \ . \end{aligned} \tag{100}$$

where in second inequality we have used twice the fact that the purified distance cannot decrease under application of a completely positive, trace-nonincreasing map, and that $\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \mathscr{F}_{R_{X'}\to R_X}(\rho'_{X'R_{X'}})$. Observe finally that

$$\mathscr{T}_{X\to X''}(\Gamma_X) = \mathscr{T}^{(2)}_{X'\to X''}(\mathscr{T}^{(1)}_{X\to X'}(\Gamma_X)) \leqslant 2^{-y_1}\, \mathscr{T}^{(2)}_{X'\to X''}(\Gamma_{X'}) \leqslant 2^{-y_1-y_2}\, \Gamma_{X''} \ , \tag{101}$$

proving that $\mathscr{T}_{X\to X''}$, $y = y_1 + y_2$ are valid optimization candidates in (46) for $\hat{D}^{\varepsilon+\varepsilon'}_{X\to X''}(\mathscr{E}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X})) \,\|\, \Gamma_X, \Gamma_{X''})$, proving the claim. ■

Corollary 21 (Chain rule in terms of states)**.** *Consider systems A, B, C and $R_A \simeq A$, $R_B \simeq B$. Let $\Gamma_C \geqslant 0$, $\Gamma_{AB} \geqslant 0$ and write $\Gamma_A = \mathrm{tr}_B[\Gamma_{AB}]$. Let $\tau_{CR_AR_B}$ be any tripartite state. Then, for $\varepsilon, \varepsilon' \geqslant 0$,*

$$\hat{D}^{\varepsilon}_{A\to AB}(\rho_{ABR_A} \,\|\, \Gamma_A, \Gamma_{AB}) + \hat{D}^{\varepsilon'}_{AB\to C}(\tau_{CR_AR_B} \,\|\, \Gamma_{AB}, \Gamma_C) \leqslant \hat{D}^{\varepsilon+\varepsilon'}_{A\to C}(\tau_{CR_A} \,\|\, \Gamma_A, \Gamma_C) \ , \tag{102}$$

where $\rho_{ABR_A} = \mathrm{tr}_{R_B}[\tau^{1/2}_{R_AR_B}\, \Phi_{AB:R_AR_B}\, \tau^{1/2}_{R_AR_B}]$.

Proof of Corollary 21. Define systems $X = A$, $X' = AB$ and $X'' = C$. Let

$$\mathscr{E}^{(1)}_{X\to X'}(\cdot) = \mathrm{tr}_{R_A}[\rho^{-1/2}_{R_A}\, \rho_{ABR_A}\, \rho^{-1/2}_{R_A}\, t_{A\to R_A}(\cdot)] \ ; \tag{103a}$$

$$\mathscr{E}^{(2)}_{X'\to X''}(\cdot) = \mathrm{tr}_{R_AR_B}[\tau^{-1/2}_{R_AR_B}\, \tau_{CR_AR_B}\, \tau^{-1/2}_{R_AR_B}\, t_{AB\to R_AR_B}(\cdot)] \ . \tag{103b}$$

These mappings are trace nonincreasing. Let $\sigma_X = t_{R_X\to X}(\tau_{R_X}) = t_{R_X\to X}(\rho_{R_X})$. We see that $\mathscr{E}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_X)) = \mathscr{E}^{(2)}_{X'\to X''}(\rho_{AB}) =$

$\mathscr{E}^{(2)}_{X'\to X''}(t_{R_AR_B\to AB}(\tau_{R_AR_B})) = \tau_C$ which has unit trace as required. Furthermore, let $|\sigma\rangle_{XR_X} = \sigma_X^{1/2}\,|\Phi\rangle_{X:R_X} = \sigma_A^{1/2}\,|\Phi\rangle_{A:R_A}$ and $|\rho'\rangle_{X'R_{X'}} = (\rho_{AB}^{1/2})\,|\Phi\rangle_{AB:R_AR_B} = (\tau_{R_AR_B}^{1/2})\,|\Phi\rangle_{ABR_AR_B}$. Now calculate

$$\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \Pi^{\rho_{R_A}}_{R_A}\ \mathrm{tr}_{\tilde{R}_A}\big[\rho_{AB\tilde{R}_A}\, t_{A\to\tilde{R}_A}(\Phi_{A:R_A})\big]\,\Pi^{\rho_{R_A}}_{R_A} = \rho_{ABR_A}\ , \quad (104)$$

as well as

$$\mathscr{E}^{(2)}_{X'\to X''}(\rho'_{X'R_{X'}}) = \Pi^{\tau_{R_AR_B}}_{R_AR_B}\ \mathrm{tr}_{\tilde{R}_A\tilde{R}_B}\big[\tau_{C\tilde{R}_A\tilde{R}_B}\, t_{AB\to\tilde{R}_A\tilde{R}_B}(\Phi_{AB:R_AR_B})\big]\,\Pi^{\tau_{R_AR_B}}_{R_AR_B} = \tau_{CR_AR_B}\ , \quad (105)$$

and, since $\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \rho_{ABR_A} = \mathrm{tr}_{R_B}[\rho'_{ABR_AR_B}]$,

$$\mathscr{E}^{(2)}_{X'\to X''}(\mathscr{E}^{(1)}_{X\to X'}(\sigma_{XR_X})) = \mathrm{tr}_{R_B}\big[\mathscr{E}^{(2)}_{X'\to X''}(\rho'_{ABR_AR_B})\big] = \tau_{CR_A}\ . \quad (106)$$

All conditions for Proposition 20 are fulfilled, and the claim follows. ■

E. Alternative smoothing of the coherent relative entropy

There is another possible way to define the smooth coherent relative entropy (i.e., for $\varepsilon > 0$), based on optimizing its non-smooth version (for $\varepsilon = 0$) over all states which are ε-close to the requested state. This smoothing method is the method used traditionally in the smooth entropy framework [4, 5, 12]. The disadvantage of this alternative definition is that it can no longer be formulated as a semidefinite program. However, in the regime of small ε, it turns out that both definitions are equivalent up to factors which depend only on ε, and which do not scale with the dimension of the system (Proposition 26 below). In particular, both quantities behave in the same way in the i.i.d. limit.

Alternative smoothing. *For a normalized state $\rho_{X'R_X}$, positive semidefinite $\Gamma_X,\Gamma_{X'}$, and for $\varepsilon \geqslant 0$, we define the quantity*

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \max_{\hat{\rho}_{X'R_X}\approx_\varepsilon \rho_{X'R_X}} \hat{D}_{X\to X'}(\hat{\rho}_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ , \quad (107)$$

where the maximization in (107) is taken over (normalized) quantum states which are in the support of $\Gamma_X\otimes\Gamma_{X'}$ and which are close to $\rho_{X'R_X}$ in the purified distance, $P(\hat{\rho}_{X'R_X},\rho_{X'R_X}) \leqslant \varepsilon$.

Some properties of $\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$ carry over immediately to $\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$, which we summarize here without explicit proof. These propositions are straightforwardly proven by applying the relevant property to the inner coherent relative entropy in (107).

Proposition 22 (cf. Proposition 6)**.** *For any $0 \leqslant \varepsilon < 1$,*

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant -\log\mathrm{tr}\,\Gamma_X - \log\|\Gamma_{X'}^{-1}\|_\infty\ ; \quad (108a)$$

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant \log\|\Gamma_X^{-1}\|_\infty + \log\mathrm{tr}\,\Gamma_{X'}\ . \quad (108b)$$

Proposition 23 (cf. Proposition 8)**.** *For any $a,b \geqslant 0$,*

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,a\Gamma_X, b\Gamma_{X'}) = \bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) + \log\frac{b}{a}\ . \quad (109)$$

Proposition 24 (cf. Proposition 7)**.** *Let $\tilde{X}$, $\tilde{X}'$ be new systems. Suppose there exist partial isometries $V_{X\to\tilde{X}}$ and $V'_{X'\to\tilde{X}'}$ such that both $t_{R_X\to X}(\rho_{R_X})$ and Γ_X are in the support of $V_{X\to\tilde{X}}$, and both $\rho_{X'}$ and $\Gamma_{X'}$ are in the support of $V'_{X'\to\tilde{X}'}$. Then*

$$\bar{D}^{\varepsilon}_{\tilde{X}\to\tilde{X}'}((V'\otimes V)\,\rho_{X'R_X}\,(V'\otimes V)^\dagger\,\|\,V\Gamma_XV^\dagger, V'\Gamma_{X'}V'^\dagger) = \bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ . \quad (110)$$

We now give a loose equivalent of Proposition 16 for the alternative smoothing of the coherent relative entropy. The error term is relatively loose (it scales proportionally to n and to ε), and it does not disappear in the i.i.d. limit unless the limit $\varepsilon\to 0$ is taken explicitly. For this reason, for small ε, it might be advantageous to use Proposition 16 in conjunction with Proposition 26.

Proposition 25. *Let $P_X, P'_{X'}$ be projectors such that $[\Gamma_X,P_X] = 0$ and $[\Gamma_{X'},P'_{X'}] = 0$. Let $\rho_{X'R_X}$ be such that $\rho_{R_X} = t_{X\to R_X}(P_X\Gamma_XP_X/\mathrm{tr}\,P_X\Gamma_X)$ and $\rho_{X'} = P'_{X'}\Gamma_{X'}P'_{X'}/\mathrm{tr}\,P'_{X'}\Gamma_{X'}$. Let $\varepsilon \geqslant 0$. Then*

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \log\frac{\mathrm{tr}\,P'_{X'}\Gamma_{X'}}{\mathrm{tr}\,P_X\Gamma_X} + f(\varepsilon,\Gamma_X,\Gamma_{X'})\ , \quad (111)$$

where the error term $f(\varepsilon,\Gamma_X,\Gamma_{X'})$ is bounded as

$$0 \leqslant f(\varepsilon,\Gamma_X,\Gamma_{X'}) \leqslant f_0(\varepsilon,\Gamma_X) + f_0(\varepsilon,\Gamma_{X'})\ , \quad (112)$$

where $f_0(\varepsilon,\Gamma) = \varepsilon\log(\mathrm{rank}\,\Gamma - 1) + \varepsilon\|\log\Gamma\|_\infty + h(\varepsilon)$ with the binary entropy $h(\varepsilon) = -\varepsilon\log\varepsilon - (1-\varepsilon)\log(1-\varepsilon)$.

Proof of Proposition 25. The lower bound is given simply as

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant \bar{D}^{\varepsilon=0}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \log\frac{\mathrm{tr}\,P'_{X'}\Gamma_{X'}}{\mathrm{tr}\,P_X\Gamma_X}\ , \quad (113)$$

where the latter expression is provided by Proposition 16, recalling that for $\varepsilon = 0$ both versions of the smooth coherent relative entropy coincide exactly. For the upper bound, let $\hat{\rho}_{X'R_X}$ be the optimal state such that $P(\hat{\rho}_{X'R_X},\rho_{X'R_X}) \leqslant \varepsilon$ and

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \hat{D}_{X\to X'}(\hat{\rho}_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ , \quad (114)$$

and invoke Proposition 31 to get

$$(114) \leqslant D(\hat{\rho}_X\,\|\,\Gamma_X) - D(\hat{\rho}_{X'}\,\|\,\Gamma_{X'})\ . \quad (115)$$

We have $D(\hat{\rho}_{R_X},\rho_{R_X}) \leqslant P(\hat{\rho}_{R_X},\rho_{R_X}) \leqslant \varepsilon$ and analogously $D(\hat{\rho}_{X'},\rho_{X'}) \leqslant \varepsilon$. By continuity of the relative entropy given in Lemma 48, we get

$$|D(\hat{\rho}_{R_X}\,\|\,\Gamma_{R_X}) - D(\rho_{R_X}\,\|\,\Gamma_{R_X})| \leqslant f_0(\varepsilon,\Gamma_{R_X})\ ;$$
$$|D(\hat{\rho}_{X'}\,\|\,\Gamma_{X'}) - D(\rho_{X'}\,\|\,\Gamma_{X'})| \leqslant f_0(\varepsilon,\Gamma_{X'})\ , \quad (116a)$$

where $f_0(\varepsilon,\Gamma)$ is as given in the claim. On the other hand,

$$D(\rho_{R_X}\,\|\,\Gamma_{R_X}) - D(\rho_{X'}\,\|\,\Gamma_{X'}) = \log\mathrm{tr}\,P'_{X'}\Gamma_{X'} - \log\mathrm{tr}\,P_{R_X}\Gamma_{R_X}\ , \quad (117)$$

because $\rho_{R_X} = P_{R_X}\Gamma_{R_X}P_{R_X}/\mathrm{tr}\,P_{R_X}\Gamma_{R_X}$ and $\rho_{X'} = P'_{X'}\Gamma_{X'}P'_{X'}/\mathrm{tr}\,P'_{X'}\Gamma_{X'}$, as given by (92). This means that

$$(115) \leqslant \log\mathrm{tr}\frac{P'_{X'}\Gamma_{X'}}{P_{R_X}\Gamma_{R_X}} + f_0(\varepsilon,\Gamma_{R_X}) + f_0(\varepsilon,\Gamma_{X'})\ .$$ ■

Crucially, this alternative smoothing method does not alter the quantity much in the regime of small ε. In fact, both versions of the smooth coherent relative entropy are related by a simple adjustment of the ε parameter, and up to an error term which depends only on ε and doesn't scale with the system size.

Proposition 26. *Let $\rho_{X'R_X}$ be any quantum state. Then for any $\varepsilon \geqslant 0$ with $3\sqrt{\varepsilon} < 1$,*

$$\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant \hat{D}^{3\sqrt{\varepsilon}}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ . \quad (118)$$

Conversely, for any $\varepsilon > 0$ with $9\varepsilon^{1/4} < 1$,

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant \bar{D}^{9\varepsilon^{1/4}}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) + \log(1/\varepsilon)\ . \quad (119)$$

We need to prove the following lemma first.

Lemma 27. *Let $\Gamma_X, \Gamma_{X'} \geqslant 0$. Let $\mathscr{T}_{X\to X'}$ be a completely positive, trace-nonincreasing map. Let $Q_X = \mathscr{T}^\dagger(\mathbb{1}_{X'})$. Assume that the support of Q_X lies within the support of Γ_X, and that $\mathscr{T}_{X\to X'}(\Gamma_X)$ lies within the support of $\Gamma_{X'}$. Then*

$$\min\left\{\alpha:\ \mathscr{T}_{X\to X'}(\Gamma_X) \leqslant \alpha\,\Gamma_{X'}\right\} \geqslant \frac{\operatorname{tr}(Q_X\Gamma_X)}{\operatorname{tr}\Gamma_{X'}}\ . \quad (120)$$

Proof of Lemma 27. The optimal α is given by

$$\begin{aligned}\alpha &= \big\|\Gamma_{X'}^{-1/2}\,\mathscr{T}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\big\|_\infty \\ &\geqslant \operatorname{tr}\left[\left(\frac{\Gamma_{X'}}{\operatorname{tr}\Gamma_{X'}}\right)\Gamma_{X'}^{-1/2}\,\mathscr{T}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\right] \\ &= (\operatorname{tr}\Gamma_{X'})^{-1}\operatorname{tr}[\mathscr{T}_{X\to X'}(\Gamma_X)] = (\operatorname{tr}\Gamma_{X'})^{-1}\operatorname{tr}[Q_X\Gamma_X]\ ,\end{aligned} \quad (121)$$

where we have used that $\|\cdot\|_\infty = \max_\gamma \operatorname{tr}[\gamma(\cdot)]$ with γ ranging over all density operators. ∎

Proof of Proposition 26. First we prove (118). Let $\tilde{\rho}_{X'R}$ be the state which achieves the optimum in $\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$, and let $T_{X'R_X}$, α be optimal primal variables for $2^{-\hat{D}_{X\to X'}(\tilde{\rho}_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$ for the semidefinite program in Proposition 13, and denote by $\mathscr{T}_{X\to X'}$ the completely positive, trace-nonincreasing map corresponding to $T_{X'R_X}$. Write $|\sigma\rangle_{XR_X} = \rho_{R_X}^{1/2}|\Phi\rangle_{X:R_X}$ and $|\tilde{\sigma}\rangle_{XR_X} = \tilde{\rho}_{R_X}^{1/2}|\Phi\rangle_{X:R_X}$. Since $P(\sigma_{R_X},\tilde{\sigma}_{R_X}) \leqslant \varepsilon$, we see using Lemma 47 that $P(\sigma_{XR_X},\tilde{\sigma}_{XR_X}) \leqslant 2\sqrt{\varepsilon}$. The purified distance may not increase under the action of the trace nonincreasing map $\mathscr{T}_{X\to X'}$, and hence

$$\begin{aligned}&P\big(\mathscr{T}_{X\to X'}(\sigma_{XR_X}),\rho_{X'R_X}\big)\\ &\quad\leqslant P\big(\mathscr{T}_{X\to X'}(\sigma_{XR_X}),\tilde{\rho}_{X'R_X}\big) + P(\tilde{\rho}_{X'R_X},\rho_{X'R_X})\\ &\quad\leqslant P\big(\mathscr{T}_{X\to X'}(\sigma_{XR_X}),\mathscr{T}_{X\to X'}(\tilde{\sigma}_{XR_X})\big) + \varepsilon\\ &\quad\leqslant 2\sqrt{\varepsilon}+\varepsilon \leqslant 3\sqrt{\varepsilon}\ .\end{aligned} \quad (122)$$

Hence, $\mathscr{T}_{X\to X'}$ is an optimization candidate for $2^{-\hat{D}^{3\sqrt{\varepsilon}}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$ with the same achieved value, proving (118).

Now we prove (119). In the remainder of this proof, we use the shorthand system name $R \equiv R_X$. Let $\hat{T}_{X'RE}$, $\hat{\alpha}$ be the optimal primal variables for $2^{-\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R}\,\|\,\Gamma_X,\Gamma_{X'})}$. We will construct an explicit $\tilde{\rho}_{X'R}$ close to $\rho_{X'R}$, as well as feasible candidates $\tilde{T}_{X'R}$ and $\tilde{\alpha}$ in the optimization for $\bar{D}_{X\to X'}(\tilde{\rho}_{X'R}\,\|\,\Gamma_X,\Gamma_{X'})$ as given by Proposition 13. We denote by $\hat{\mathscr{T}}_{X\to X'}$ the completely positive, trace nonincreasing map corresponding to $\hat{T}_{X'RE}$. Let $\sigma_{XR} = \rho_R^{1/2}\,\Phi_{X:R}\,\rho_R^{1/2}$ and define

$$\hat{\rho}_{X'R} = \hat{\mathscr{T}}_{X\to X'}(\sigma_{XR})\ . \quad (123)$$

By assumption, $P(\hat{\rho}_{X'R},\rho_{X'R}) \leqslant \varepsilon$ and hence $D(\hat{\rho}_{X'R},\rho_{X'R}) \leqslant \varepsilon$. Using the fact that $\hat{\rho}_{X'R} = \rho_{X'R} + \Delta^+_{X'R} - \Delta^-_{X'R}$ for some $\Delta^{\pm}_{X'R} \geqslant 0$ with $\operatorname{tr}\Delta^+_{X'R} = \operatorname{tr}\Delta^-_{X'R} = D(\hat{\rho}_{X'R},\rho_{X'R}) \leqslant \varepsilon$, we see that $\operatorname{tr}\hat{\rho}_{X'R} \geqslant \operatorname{tr}\rho_{X'R} - \varepsilon = 1-\varepsilon$.

Define $Q = \hat{\mathscr{T}}^\dagger(\mathbb{1}_{X'})$ and note that $0 \leqslant Q \leqslant \mathbb{1}$. For any $0 < \eta < 1$, let P^η be the projector onto the eigenspaces of Q for which the corresponding eigenvalues are greater or equal to η; clearly P^η and Q commute. Define $R^\eta = P^\eta - P^\eta Q P^\eta$, noting that P^η, Q, R^η all commute. By definition, $\eta P^\eta \leqslant P^\eta Q P^\eta$, and hence $R^\eta \leqslant (\eta^{-1}-1)\,P^\eta Q P^\eta \leqslant (\eta^{-1}-1)\,Q$. We may now define

$$\tilde{\mathscr{T}}_{X\to X'}(\cdot) = \hat{\mathscr{T}}_{X\to X'}(\cdot) + \operatorname{tr}(R^\eta\,(\cdot))\,\frac{\Gamma_{X'}}{\operatorname{tr}\Gamma_{X'}}\ . \quad (124)$$

The mapping $\tilde{\mathscr{T}}_{X\to X'}$ is trace non-increasing,

$$\tilde{\mathscr{T}}^\dagger_{X\leftarrow X'}(\mathbb{1}_{X'}) = Q + R^\eta = P^\eta + P^{\eta,\perp}\,Q\,P^{\eta,\perp} \leqslant \mathbb{1}\ , \quad (125)$$

where $P^{\eta,\perp} = \mathbb{1} - P^\eta$, keeping in mind that $Q = P^\eta Q P^\eta + P^{\eta,\perp} Q P^{\eta,\perp}$ and that $R^\eta + P^\eta Q P^\eta = P^\eta$. Furthermore $\tilde{\mathscr{T}}_{X\to X'}$ is trace-preserving on the subspace spanned by P^η, i.e. $P^\eta\,\tilde{\mathscr{T}}^\dagger_{X\leftarrow X'}(\mathbb{1}_{X'})\,P^\eta = P^\eta$. This means that for any state τ lying in the support of P^η, it holds that $\operatorname{tr}[\tilde{\mathscr{T}}_{X\to X'}(\tau)] = 1$. The map $\tilde{\mathscr{T}}_{X\to X'}$ moreover satisfies

$$\begin{aligned}\tilde{\mathscr{T}}_{X\to X'}(\Gamma_X) &\leqslant \hat{\alpha}\,\Gamma_{X'} + \frac{\operatorname{tr}R^\eta\Gamma_X}{\operatorname{tr}\Gamma_{X'}}\,\Gamma_{X'}\\ &\leqslant \left(\hat{\alpha} + (\eta^{-1}-1)\,\frac{\operatorname{tr}Q\Gamma_X}{\operatorname{tr}\Gamma_{X'}}\right)\Gamma_{X'} \leqslant \eta^{-1}\hat{\alpha}\,\Gamma_{X'}\ ,\end{aligned} \quad (126)$$

where in the last inequality we have used Lemma 27 to see that $\hat{\alpha} \geqslant \operatorname{tr}(Q\Gamma_X)/\operatorname{tr}\Gamma_{X'}$. We are led to define (surprise!) $\tilde{\alpha} = \eta^{-1}\hat{\alpha}$.

It remains to find a state $\tilde{\rho}_{X'R}$ which is close to $\rho_{X'R}$ such that $\tilde{\rho}_R^{1/2}\,\tilde{\mathscr{T}}_{X\to X'}(\Phi_{X:R})\,\tilde{\rho}_R^{1/2} = \tilde{\rho}_{X'R}$. First define

$$\tilde{\sigma}_X = \frac{P^\eta\,\sigma_X\,P^\eta}{\operatorname{tr}(P^\eta\,\sigma_X)}\ . \quad (127)$$

Observe that $\operatorname{tr}(P^\eta\sigma_X) \geqslant \operatorname{tr}(P^\eta Q P^\eta \sigma_X) = \operatorname{tr}(Q\,\sigma_X) - \operatorname{tr}\big(P^{\eta,\perp}\,Q\,P^{\eta,\perp}\sigma_X\big) \geqslant 1-\varepsilon-\eta$, where $P^{\eta,\perp} = \mathbb{1} - P^\eta$, using the fact that all eigenvalues of Q within $P^{\eta,\perp}$ are less than η and that $\operatorname{tr}(Q\,\sigma_X) = \operatorname{tr}\big(\hat{\mathscr{T}}(\sigma_X)\big) = \operatorname{tr}\hat{\rho}_{X'} \geqslant 1-\varepsilon$. Then, using Lemma 45,

$$P(\tilde{\sigma}_X,\sigma_X) \leqslant \frac{\sqrt{2(\varepsilon+\eta)}}{\sqrt{1-\varepsilon-\eta}} =: \bar{\varepsilon}\ . \quad (128)$$

Write $\tilde{\sigma}_{XR} = \tilde{\sigma}_X^{1/2}\,\Phi_{X:R}\,\tilde{\sigma}_X^{1/2}$. Using Lemma 47 we see that $P(\tilde{\sigma}_{XR},\sigma_{XR}) \leqslant 2\sqrt{D(\tilde{\sigma}_R,\rho_R)} \leqslant 2\sqrt{P(\tilde{\sigma}_R,\rho_R)} \leqslant 2\sqrt{\bar{\varepsilon}}$. At this point, define

$$\tilde{\rho}_{X'R} = \tilde{\mathscr{T}}_{X\to X'}(\tilde{\sigma}_{XR})\ ; \quad (129a)$$

$$\bar{\rho}_{X'R} = \tilde{\mathscr{T}}_{X\to X'}(\sigma_{XR})\ . \quad (129b)$$

Because $\tilde{\sigma}_X$ lies within the support of P^η, we have $\operatorname{tr}_{X'}\tilde{\rho}_{X'R} = \operatorname{tr}_X\big(\tilde{\mathscr{T}}^\dagger(\mathbb{1}_{X'})\,\tilde{\sigma}_{XR}\big) = \operatorname{tr}_X\big(\tilde{\mathscr{T}}^\dagger(\mathbb{1}_{X'})\,P^\eta\,\tilde{\sigma}_{XR}\,P^\eta\big) = \tilde{\sigma}_R$, and hence we have $\tilde{\rho}_R^{1/2}\,\tilde{\mathscr{T}}_{X\to X'}(\Phi_{X:R})\,\tilde{\rho}_R^{1/2} = \tilde{\rho}_{X'R}$ as required. Furthermore, the purified distance cannot increase under the action of $\tilde{\mathscr{T}}_{X\to X'}$, so we have $P(\tilde{\rho}_{X'R},\bar{\rho}_{X'R}) \leqslant 2\sqrt{\bar{\varepsilon}}$. Also, $\bar{\rho}_{X'R} = \hat{\mathscr{T}}_{X\to X'}(\sigma_{XR}) + D_{X'R} = \hat{\rho}_{X'R} + D_{X'R}$ with $D_{X'R} = \operatorname{tr}(R^\eta\,\sigma_{XR})\,(\operatorname{tr}\Gamma_{X'})^{-1}\,\Gamma_{X'}$, noting that $\operatorname{tr}D_{X'R} \leqslant \operatorname{tr}(\bar{\rho}_{X'R}) - \operatorname{tr}(\hat{\rho}_{X'R}) \leqslant 1-(1-\varepsilon) \leqslant \varepsilon$; hence $D(\bar{\rho}_{X'R},\hat{\rho}_{X'R}) \leqslant \varepsilon$ and thus $P(\bar{\rho}_{X'R},\hat{\rho}_{X'R}) \leqslant \sqrt{2\varepsilon}$. We deduce that $P(\tilde{\rho}_{X'R},\rho_{X'R}) \leqslant P(\tilde{\rho}_{X'R},\bar{\rho}_{X'R}) + P(\bar{\rho}_{X'R},\hat{\rho}_{X'R}) + P(\hat{\rho}_{X'R},\rho_{X'R}) \leqslant 2\sqrt{\bar{\varepsilon}} + \sqrt{2\varepsilon} + \varepsilon$.

Let's summarize: We now have a state $\tilde{\rho}_{X'R}$ satisfying $P(\tilde{\rho}_{X'R},\rho_{X'R}) \leqslant 2\sqrt{\bar{\varepsilon}} + \sqrt{2\varepsilon} + \varepsilon$, as well as a trace-nonincreasing map $\tilde{\mathscr{T}}_{X\to X'}$ satisfying $\tilde{\rho}_R^{1/2}\,\tilde{\mathscr{T}}_{X\to X'}(\Phi_{X:R})\,\tilde{\rho}_R^{1/2} = \tilde{\rho}_{X'R}$ and $\tilde{\mathscr{T}}_{X\to X'}(\Gamma_X) \leqslant \alpha\,\eta^{-1}\,\Gamma_{X'}$. The claim follows by choosing $\eta = \varepsilon$ and calculating the bounds $\bar{\varepsilon} \leqslant \sqrt{8\varepsilon}$ (using the assumption $\varepsilon < 1/4$) as well as $2\sqrt{\bar{\varepsilon}} + \sqrt{2\varepsilon} + \varepsilon \leqslant \big(4\sqrt{2}+\sqrt{2}+1\big)\,\varepsilon^{1/4} \leqslant 9\,\varepsilon^{1/4}$. ∎

F. Recovering known entropy measures

An interesting aspect of the coherent relative entropy is that it reduces to various previously-known entropy measures, including the min- and max-relative entropies [14], as well as the conditional min- and max-entropy [4, 12]. These measures are already known to be relevant in counting the work cost of specific processes in quantum thermodynamics [9, 15–18].

First we present some definitions. Given a (normalized) quantum state ρ_{AB}, we define the *(conditional) von Neumann entropy*, the *(conditional alternative) max-entropy*, and the *(conditional alternative) min-entropy* respectively as,[4]

$$H(A\,|\,B)_\rho = -\operatorname{tr}(\rho_{AB}\log\rho_{AB}) + \operatorname{tr}(\rho_B\log\rho_B)\ ;$$
$$\hat{H}_{\max}(A\,|\,B)_\rho = \log\left\|\operatorname{tr}_A \Pi^{\rho_{AB}}_{AB}\right\|_\infty\ ;\text{ and}$$
$$\hat{H}_{\min}(A\,|\,B)_\rho = -\log\left\|\rho_B^{-1/2}\rho_{AB}\rho_B^{-1/2}\right\|_\infty\ .$$

For any $\varepsilon > 0$, we define the *smooth (conditional alternative) max-entropy* and *smooth (conditional alternative) min-entropy* respectively as

$$\hat{H}^\varepsilon_{\max}(A\,|\,B)_\rho = \min_{\hat\rho_{AB}\approx_\varepsilon\rho_{AB}} \hat{H}_{\max}(A\,|\,B)_{\hat\rho}\ ;$$
$$\hat{H}^\varepsilon_{\min}(A\,|\,B)_\rho = \max_{\hat\rho_{AB}\approx_\varepsilon\rho_{AB}} \hat{H}_{\min}(A\,|\,B)_{\hat\rho}\ ,$$

where the optimizations range over (normalized[5]) states $\hat\rho_{AB}$ and where $\hat\rho_{AB}\approx_\varepsilon\rho_{AB}$ denotes proximity in the purified distance, i.e., $P(\hat\rho_{AB},\rho_{AB})\leqslant\varepsilon$.

For a (normalized) quantum state ρ_X, and any $\Gamma_X\geqslant 0$, we define the *quantum relative entropy*, the *relative min-entropy*, and the *relative max-entropy* respectively as,

$$D(\rho_X\,\|\,\Gamma_X) = \operatorname{tr}\big[\rho_X(\log_2\rho_X - \log_2\Gamma_X)\big]\ ;$$
$$D_{\min,0}(\rho_X\,\|\,\Gamma_X) = -\log\operatorname{tr}\big[\Pi^{\rho_X}_X\Gamma_X\big]\ ;$$
$$D_{\max}(\rho_X\,\|\,\Gamma_X) = \log\|\Gamma_X^{-1/2}\rho_X\Gamma_X^{-1/2}\|_\infty\ ,$$

recalling that $\Pi^{\rho_X}_X$ denotes the projector onto the support of ρ_X. We define the smoothed versions of the relative min- and max-entropies as

$$D^\varepsilon_{\min,0}(\rho\,\|\,\Gamma) = \max_{\hat\rho\approx_\varepsilon\rho} D_{\min,0}(\hat\rho\,\|\,\Gamma)\ ;$$
$$D^\varepsilon_{\max}(\rho\,\|\,\Gamma) = \min_{\hat\rho\approx_\varepsilon\rho} D_{\max}(\hat\rho\,\|\,\Gamma)\ .$$

where the optimizations range over normalized[6] states $\hat\rho_{AB}$ such that $P(\hat\rho_{AB},\rho_{AB})\leqslant\varepsilon$.

We furthermore define the *hypothesis testing relative entropy* [19–24] for $0<\eta\leqslant 1$ as

$$D^\eta_{\mathrm{H}}(\rho\,\|\,\Gamma) = -\frac{1}{\eta}\log\min_{\substack{0\leqslant Q\leqslant\mathbb{1}\\ \operatorname{tr}[Q\rho]\geqslant\eta}}\operatorname{tr}[Q\Gamma]\ .$$

We now show that we can recover the max-entropy in the case where for both input and output systems we have $\Gamma=\mathbb{1}$.

Proposition 28 (Recovering the max-entropy)**.** *Let $|\rho\rangle_{X'R_XE}$ be any pure state on systems R_X, X', and E with $|E|\geqslant|X'R_X|$. Then*

$$\bar{D}^\varepsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'}) = -\hat{H}^\varepsilon_{\max}(E\,|\,X')_\rho = \hat{H}^\varepsilon_{\min}(E\,|\,R_X)_\rho\ . \tag{130}$$

Proof of Proposition 28. Let $|\tilde\rho\rangle_{X'R_XE}$ be any pure quantum state. Considering the semidefinite problem for $2^{-\hat{D}_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$, let $T_{X'RE} = \tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_XE}\tilde\rho_{R_X}^{-1/2}$. Conditions (49a) and (49c) are automatically satisfied. Choosing $\alpha = \|\operatorname{tr}_{R_X}[T_{X'R_X}]\|_\infty = \|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\|_\infty$ ensures that (49b) is satisfied, and hence

$$\hat{D}_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant -\log\left\|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\right\|_\infty\ . \tag{131}$$

Now let $\omega_{X'}\geqslant 0$ with $\operatorname{tr}\omega_{X'}=1$ such that $\operatorname{tr}\big[\omega_{X'}\cdot\operatorname{tr}_R(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2})\big] = \left\|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\right\|_\infty$, and note that condition (50a) is satisfied. Now let $X_{R_X}=0$ and $Z_{X'R_X}=\tilde\rho_{R_X}^{-1}\otimes\omega_{X'}$, and we see that

$$\tilde\rho_{R_X}^{1/2}Z_{X'R_X}\tilde\rho_{R_X}^{1/2} = \Pi^{\tilde\rho_{R_X}}_{R_X}\otimes\omega_{X'}\leqslant\mathbb{1}_{R_X}\otimes\omega_{X'}\ . \tag{132}$$

The attained value is

$$\begin{aligned}\operatorname{tr}\left[Z_{X'R_X}\tilde\rho_{X'R_X}\right] &= \operatorname{tr}\left[\tilde\rho_{R_X}^{-1}\otimes\omega_{X'}\cdot\tilde\rho_{X'R_X}\right]\\ &= \operatorname{tr}\left[\omega_{X'}\cdot\operatorname{tr}_{R_X}(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2})\right]\\ &= \left\|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\right\|_\infty\ ,\end{aligned}$$

providing us with the opposite bound to (131), and hence proving that

$$\hat{D}_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'}) = -\log\left\|\operatorname{tr}_R\tilde\rho_R^{-1/2}\tilde\rho_{X'R}\tilde\rho_R^{-1/2}\right\|_\infty\ . \tag{133}$$

We now use this expression to show that

$$\hat{D}_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'}) = -\hat{H}_{\max}(E\,|\,X')_{\tilde\rho} = \hat{H}_{\min}(E\,|\,R_X)_{\tilde\rho}\ . \tag{134}$$

Consider the bipartition $EX':R$ of the pure state $|\tilde\rho\rangle_{EX'R}$, and write the Schmidt decomposition $|\tilde\rho\rangle_{EX'R_X} = \tilde\rho_{EX'}^{1/2}|\Phi^{\tilde\rho}\rangle_{EX':R_X} = \tilde\rho_{R_X}^{1/2}|\Phi^{\tilde\rho}\rangle_{EX':R_X}$, with $\operatorname{tr}_{R_X}\Phi^{\tilde\rho}_{EX':R_X} = \Pi^{\tilde\rho_{EX'}}_{EX'}$. Then

$$\begin{aligned}(133) &= -\log\left\|\operatorname{tr}_{ER_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{EX'R_X}\tilde\rho_{R_X}^{-1/2}\right\|_\infty\\ &= -\log\left\|\operatorname{tr}_{ER_X}|\Phi\rangle\langle\Phi|^{\tilde\rho}_{EX'R_X}\right\|_\infty\\ &= -\log\left\|\operatorname{tr}_E\Pi^{\tilde\rho_{EX'}}_{EX'}\right\|_\infty\\ &= -\hat{H}_{\max}(E\,|\,X')_{\tilde\rho}\ .\end{aligned}$$

Similarly,

$$\begin{aligned}(133) &= -\log\left\|\operatorname{tr}_{ER_X}(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{EX'R_X}\tilde\rho_{R_X}^{-1/2})\right\|_\infty\\ &= -\log\left\|\operatorname{tr}_{X'}(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{EX'R_X}\tilde\rho_{R_X}^{-1/2})\right\|_\infty\\ &= -\log\left\|\tilde\rho_{R_X}^{-1/2}\tilde\rho_{ER_X}\tilde\rho_{R_X}^{-1/2}\right\|_\infty = \hat{H}_{\min}(E\,|\,R_X)_{\tilde\rho}\ ,\end{aligned}$$

---

[4] There exist several different variants of the min- and max-entropy [4, 12]; however, all the max-entropies as well as all the min-entropies are equivalent up to terms of order $\log\varepsilon$ after smoothing with a parameter ε.

[5] One easily notices that the normalization of the state doesn't affect these quantities, so smoothing may be restricted to normalized states (in contrast to, e.g., Refs. [4, 5]).

[6] These smooth quantities were introduced in Ref. [14] using the trace distance and optimizing over subnormalized states. The two distances are tightly related and a simple adjustment of the ε parameter is required. Furthermore we restrict to normalized states for our convenience; the $D^\varepsilon_{\min,0}$ is not affected and the $D^\varepsilon_{\max}$ is at most shifted by a factor depending on $\log(1-\varepsilon)$ only.

where the second equality holds because the argument of the partial trace is pure, and hence has the same spectrum on ER as on X' (by Schmidt decomposition).

We now see that

$$\begin{aligned}
&\bar{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \mathbb{1}_X, \mathbb{1}_{X'}) \\
&\quad = \max_{P\left(\tilde{\rho}_{X'R_X}, \rho_{X'R_X}\right)\leqslant\varepsilon} \hat{D}_{X\to X'}(\tilde{\rho}_{X'R_X} \,\|\, \mathbb{1}_X, \mathbb{1}_{X'}) \\
&\quad = \max_{P\left(|\tilde{\rho}\rangle_{X'R_XE}, |\rho\rangle_{X'R_XE}\right)\leqslant\varepsilon} \hat{D}_{X\to X'}(\tilde{\rho}_{X'R_X} \,\|\, \mathbb{1}_X, \mathbb{1}_{X'}) \\
&\quad = \max_{P\left(|\tilde{\rho}\rangle_{X'R_XE}, |\rho\rangle_{X'R_XE}\right)\leqslant\varepsilon} \hat{H}_{\max}(E\,|\,X')_{\tilde{\rho}} \\
&\quad = \hat{H}^{\varepsilon}_{\max}(E\,|\,X')_{\rho}\ ,
\end{aligned}$$

where the second equality holds by properties of the purified distance (Uhlmann's theorem). An analogous argument holds for $\hat{H}^{\varepsilon}_{\min}(E\,|\,R_X)_{\rho}$. ∎

The min- and max-relative entropies already have known connections to thermodynamics [8, 9, 17] in terms of work cost of erasure and work yield of formation of a state in the presence of a heat bath. These results are recovered here, in a fully information-theoretic context.

Proposition 29 (Recovering the min- and max-relative entropies)**.** *The min-relative entropy is recovered with a trivial output state:*

$$\bar{D}^{\varepsilon}_{X\to\emptyset}(\rho_{R_X} \,\|\, \Gamma_X, 1) = D^{\varepsilon}_{\min,0}(\sigma_X \,\|\, \Gamma_X)\ , \tag{135}$$

writing $\sigma_X = t_{R_X\to X}(\rho_{R_X})$. Furthermore the max-relative entropy is recovered with a trivial input state:

$$\bar{D}^{\varepsilon}_{\emptyset\to X'}(\rho_{X'} \,\|\, 1, \Gamma'_X) = -D^{\varepsilon}_{\max}(\rho_{X'} \,\|\, \Gamma_{X'})\ . \tag{136}$$

Proof of Proposition 29. For any state $\tilde{\rho}_{R_X}$, consider the semidefinite program given in Proposition 13 for $2^{-\hat{D}_{X\to\emptyset}(\tilde{\rho}_{R_X}\,\|\,\Gamma_X,1)}$. The choice $T_{R_X} = \Pi^{\tilde{\rho}_{R_X}}_{R_X}$ along with $\alpha = \operatorname{tr}\left(\Pi^{\tilde{\rho}_{R_X}}_{R_X}\Gamma_{R_X}\right)$ is primal feasible, hence

$$2^{-\hat{D}_{X\to\emptyset}(\tilde{\rho}_{R_X}\,\|\,\Gamma_X,1)} \leqslant 2^{-D_{\min,0}(\tilde{\rho}_{R_X}\,\|\,\Gamma_{R_X})}\ . \tag{137}$$

In the dual problem, for any $\mu > 0$ let $Z_R = \mu\Pi^{\tilde{\rho}_{R_X}}_{R_X}$ and $\omega_{X'} = 1$. Let P_{R_X} be the projector onto the eigenspaces associated with the positive (or null) eigenvalues of $(\mu\tilde{\rho}_{R_X} - \Gamma_{R_X})$, and let $X_{R_X} = P_{R_X}(\mu\tilde{\rho}_{R_X} - \Gamma_{R_X})P_{R_X}$. Then the dual constraints (61a) and (61b) are clearly satisfied. The attained value is

$$\begin{aligned}
\operatorname{tr}(Z_{R_X}\tilde{\rho}_{R_X}) - \operatorname{tr}(X_R) &= \mu\operatorname{tr}\tilde{\rho}_{R_X} - \mu\operatorname{tr}(P_{R_X}\tilde{\rho}_{R_X}) + \operatorname{tr}(P_{R_X}\Gamma_{R_X}) \\
&\geqslant \operatorname{tr}(P_{R_X}\Gamma_{R_X}) \geqslant \operatorname{tr}\left(\Pi^{\tilde{\rho}_{R_X}}_{R_X}\Gamma_R\right) - O(1/\mu)\ ,
\end{aligned} \tag{138}$$

where we have used Lemma 43 in the last step. If we take $\mu\to\infty$ we get successive feasible dual candidates whose attained value approaches $2^{-D_{\min,0}(\tilde{\rho}_R\,\|\,\Gamma_R)}$; hence this is the optimal value of the semidefinite program. Finally, we have

$$\begin{aligned}
\bar{D}^{\varepsilon}_{X\to\emptyset}(\rho_{R_X} \,\|\, \Gamma_X, 1) &= \max_{\tilde{\rho}_{R_X}\approx_\varepsilon\rho_{R_X}} \hat{D}_{X\to\emptyset}(\tilde{\rho}_{R_X} \,\|\, \Gamma_X, 1) \\
&= \max_{\tilde{\rho}_{R_X}\approx_\varepsilon\rho_{R_X}} D_{\min,0}(\tilde{\rho}_{R_X} \,\|\, \Gamma_{R_X})\ , \\
&= D^{\varepsilon}_{\min,0}(\sigma_X \,\|\, \Gamma_X)\ .
\end{aligned}$$

Let's now prove equality (136). For any state $\tilde{\rho}_{X'}$, consider the semidefinite program given in Proposition 13 for $2^{-\hat{D}_{\emptyset\to X'}(\tilde{\rho}_{X'}\,\|\,1,\Gamma_{X'})}$. The choice $T_{X'} = \rho_{X'}$ and $\alpha = \left\|\Gamma_{X'}^{-1/2}\tilde{\rho}_{X'}\Gamma_{X'}^{-1/2}\right\|_\infty = 2^{D_{\max}(\tilde{\rho}_{X'}\,\|\,\Gamma_{X'})}$ clearly satisfies the primal constraints, and thus

$$2^{-\hat{D}_{\emptyset\to X'}(\tilde{\rho}_{X'}\,\|\,1,\Gamma_{X'})} \leqslant 2^{D_{\max}(\tilde{\rho}_{X'}\,\|\,\Gamma_{X'})}\ . \tag{139}$$

By properties of the infinity norm, there exists a $\tau_{X'} \geqslant 0$ with $\operatorname{tr}\tau_{X'} = 1$ such that $\left\|\Gamma_{X'}^{-1/2}\tilde{\rho}_{X'}\Gamma_{X'}^{-1/2}\right\|_\infty = \operatorname{tr}\left[\tau_{X'}\cdot\Gamma_{X'}^{-1/2}\tilde{\rho}_{X'}\Gamma_{X'}^{-1/2}\right]$. Let $\omega_{X'} = \Gamma_{X'}^{-1/2}\tau_{X'}\Gamma_{X'}^{-1/2}$, $Z_{X'} = \omega_{X'}$ and $X = 0$. Then the dual constraints are trivially satisfied and the attained value is

$$\operatorname{tr}[Z_{X'}\tilde{\rho}_{X'}] = \operatorname{tr}\left[\Gamma_{X'}^{-1/2}\tau_{X'}\Gamma_{X'}^{-1/2}\tilde{\rho}_{X'}\right] = 2^{D_{\max}(\tilde{\rho}_{X'}\,\|\,\Gamma_{X'})}\ . \tag{140}$$

The primal and dual candidates achieve the same value, and hence this is the optimal solution to the semidefinite program. We then have

$$\begin{aligned}
\bar{D}^{\varepsilon}_{\emptyset\to X'}(\rho_{X'} \,\|\, 1, \Gamma'_X) &= \max_{\tilde{\rho}_{X'}\approx_\varepsilon\rho_{X'}} \hat{D}_{\emptyset\to X'}(\tilde{\rho}_{X'R} \,\|\, 1, \Gamma'_X) \\
&= \max_{\tilde{\rho}_{X'}\approx_\varepsilon\rho_{X'}} -D_{\max}(\tilde{\rho}_{X'} \,\|\, \Gamma_{X'}) \\
&= -D^{\varepsilon}_{\max}(\rho_{X'} \,\|\, \Gamma_{X'})\ .
\end{aligned}$$

∎

It is clear that in Proposition 29 in the case of $\varepsilon = 0$, we may replace the trivial system with $\Gamma = 1$ by a nontrivial system with arbitrary Γ, as long as it is in a pure eigenstate of the Γ operator.

Corollary 30. *Let $\Gamma_X, \Gamma_{X'} \geqslant 0$. Both following statements hold:*

(a) Let $|\mathrm{f}\rangle_{X'}$ be an eigenstate of $\Gamma_{X'}$ with eigenvalue g_{f}, and let σ_X be any quantum state in the support of Γ_X. Then:

$$\begin{aligned}
&\hat{D}_{X\to X'}(t_{X\to R_X}(\sigma_X)\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'} \,\|\, \Gamma_X, \Gamma_{X'}) \\
&\qquad = D_{\min,0}(\sigma_X \,\|\, \Gamma_X) + \log g_{\mathrm{f}}\ .
\end{aligned} \tag{141}$$

(b) Let $|\mathrm{i}\rangle_X$ be an eigenstate of Γ_X with eigenvalue g_{i}, and let $\rho_{X'}$ be any quantum state in the support of $\Gamma_{X'}$. Then:

$$\begin{aligned}
&\hat{D}_{X\to X'}(t_{X\to R_X}(|\mathrm{i}\rangle\langle\mathrm{i}|_X)\otimes\rho_{X'} \,\|\, \Gamma_X, \Gamma_{X'}) \\
&\qquad = -\log g_{\mathrm{i}} - D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'})\ .
\end{aligned} \tag{142}$$

Proof of Corollary 30. First consider claim (a). Invoking successively Proposition 10, Proposition 8, and Proposition 7, we have (writing $\sigma_{R_X} = t_{X\to R_X}(\sigma_X)$):

$$\begin{aligned}
&\hat{D}_{X\to X'}(\sigma_{R_X}\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'} \,\|\, \Gamma_X, \Gamma_{X'}) \\
&\quad = \hat{D}_{X\to X'}(\sigma_{R_X}\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'} \,\|\, \Gamma_X, g_{\mathrm{f}}|\mathrm{f}\rangle\langle\mathrm{f}|_{X'}) \\
&\quad = \hat{D}_{X\to X'}(\sigma_{R_X}\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'} \,\|\, \Gamma_X, |\mathrm{f}\rangle\langle\mathrm{f}|_{X'}) + \log g_{\mathrm{f}} \\
&\quad = \hat{D}_{X\to\emptyset}(\sigma_{R_X} \,\|\, \Gamma_X, 1) + \log g_{\mathrm{f}}\ ,
\end{aligned} \tag{143}$$

at which point we may apply Proposition 29. Claim (b) follows analogously. ∎

Finally, we will see that the usual quantum relative entropy can also be recovered in the regime where we consider states of the form $\rho^{\otimes n}_{X'^nR^n}$ for $n\to\infty$. We defer this case to Section III H, as the proof of this property requires some additional bounds we have yet to present.

G. Bounds on the coherent relative entropy

At this point, we further characterize the coherent relative entropy with bounds in terms of simpler quantities depending only on the input and output states. The main goal of this section is to prove Proposition 33 and Proposition 36, which will allow us to understand our quantity's asymptotic behavior in the i.i.d. regime.

We begin with a few upper bounds on the coherent relative entropy, given in terms of a difference of relative entropies.

Proposition 31. *We have the upper bound*

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'})\ , \tag{144}$$

writing $\sigma_X = t_{R_X\to X}(\rho_{R_X})$

Proof of Proposition 31. Consider the optimal solution $T_{X'R_X}$ and α to the primal semidefinite program of Proposition 13, and let $\mathscr{T}_{X\to X'}$ be the completely positive map corresponding to $T_{X'R_X}$, i.e. defined by $\mathscr{T}_{X\to X'}(\cdot) = \mathrm{tr}_{R_X}[T_{X'R_X}\, t_{X\to R_X}(\cdot)]$. The mapping defined in this way is completely positive since $T_{X'R_X}\geqslant 0$ and is trace-nonincreasing thanks to condition (49a).

The map $\mathscr{T}_{X\to X'}$ thus satisfies the conditions of item (i) of Proposition 4. Hence, invoking item (ii) of that proposition, let $\tilde{\Phi}_{XA\to X'A'}$ be a trace non-increasing Γ-sub-preserving map for large enough A, A', with $\Gamma_A = \mathbb{1}_A$, $\Gamma_{A'} = \mathbb{1}_{A'}$, satisfying

$$\tilde{\Phi}_{XA\to X'A'}\Big(\sigma_{XR_X}\otimes\Big(2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}\Big)\Big) = \rho_{X'R_X}\otimes\Big(2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}\Big)\ , \tag{145}$$

with $\alpha = 2^{-(\lambda_1-\lambda_2)}$ and $|\sigma\rangle_{XR_X} = \rho_{R_X}^{1/2}|\Phi\rangle_{X:R_X}$. (If α is irrational, the following argument may be applied to arbitrary good rational approximations to α.)

Now, dilate $\tilde{\Phi}_{XA\to X'A'}$ using Proposition 2 to a trace-preserving, Γ-preserving map $\Phi_{XAX'A'Q\to XAX'A'Q}$ with states $|\mathrm{x}\rangle_X, |\mathrm{a}\rangle_A, |\mathrm{i}\rangle_Q, |\mathrm{x}'\rangle_{X'}, |\mathrm{a}'\rangle_{A'}, |\mathrm{f}\rangle_Q$ (all of them being eigenstates of the respective Γ operators), satisfying

$$\Phi_{XAX'A'Q}(\Gamma_{XAX'A'Q}) = \Gamma_{XAX'A'Q}\ ; \tag{146a}$$

$$\begin{aligned}\Phi_{XAX'A'Q}&\Big(\sigma_{XR_X}\otimes\Big(2^{-\lambda_1}\mathbb{1}^A_{2^{\lambda_1}}\Big)\otimes|\mathrm{x'a'i}\rangle\langle\mathrm{x'a'i}|_{X'A'Q}\Big)\\ &= \rho_{X'R_X}\otimes\Big(2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}}\Big)\otimes|\mathrm{xaf}\rangle\langle\mathrm{xaf}|_{XAQ}\ ;\quad\text{and}\end{aligned} \tag{146b}$$

$$\langle\mathrm{x\,a\,f}|\Gamma_{XAQ}|\mathrm{x\,a\,f}\rangle_{XAQ} = \langle\mathrm{x'a'i}|\Gamma_{X'A'Q}|\mathrm{x'a'i}\rangle_{X'A'Q}\ . \tag{146c}$$

Using Proposition 17 recalling that $\Gamma_A = \mathbb{1}_A$, we see that

$$D\Big(2^{-\lambda_1}\mathbb{1}^A_{2^{\lambda_1}}\,\Big\|\,\Gamma_A\Big) = -\log\mathrm{tr}\Big(\mathbb{1}^A_{2^{\lambda_1}}\,\Gamma_A\Big) = -\lambda_1\ ; \tag{147a}$$

$$D\Big(2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}}\,\Big\|\,\Gamma_{A'}\Big) = -\log\mathrm{tr}\Big(\mathbb{1}^A_{2^{\lambda_2}}\,\Gamma_A\Big) = -\lambda_2\ , \tag{147b}$$

as well as for any pure eigenstate y of any positive semidefinite Γ,

$$D(|\mathrm{y}\rangle\langle\mathrm{y}|\,\|\,\Gamma) = -\log\mathrm{tr}\langle\mathrm{y}|\Gamma|\mathrm{y}\rangle\ . \tag{147c}$$

Then, by the data processing inequality for the relative entropy and with (146b),

$$\begin{aligned}0 &\leqslant D\Big(\sigma_X\otimes(2^{-\lambda_1}\mathbb{1}^A_{2^{\lambda_1}})\otimes|\mathrm{x'a'i}\rangle\langle\mathrm{x'a'i}|_{X'A'Q}\,\Big\|\,\Gamma_{XAX'A'Q}\Big)\\ &\quad - D\Big(\rho_{X'}\otimes(2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}})\otimes|\mathrm{xaf}\rangle\langle\mathrm{xaf}|_{XAQ}\,\Big\|\,\Gamma_{XAX'A'Q}\Big)\\ &= D(\sigma_X\,\|\,\Gamma_X) + D(2^{-\lambda_1}\mathbb{1}^A_{2^{\lambda_1}}\,\|\,\Gamma_A) + D(|\mathrm{x'a'i}\rangle\langle\mathrm{x'a'i}|_{X'A'Q}\,\|\,\Gamma_{X'A'Q})\\ &\quad - D(\rho_{X'}\,\|\,\Gamma_{X'}) - D(2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}}\,\|\,\Gamma_{A'}) - D(|\mathrm{x\,a\,f}\rangle\langle\mathrm{x\,a\,f}|_{XAQ}\,\|\,\Gamma_{XAQ})\\ &= D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'}) - \lambda_1 + \lambda_2\\ &\quad - \log\langle\mathrm{x'a'i}|\Gamma_{X'A'Q}|\mathrm{x'a'i}\rangle + \log\langle\mathrm{x\,a\,f}|\Gamma_{XAQ}|\mathrm{x\,a\,f}\rangle\\ &= D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'}) - \lambda_1 + \lambda_2\ ,\end{aligned} \tag{148}$$

where we invoked the condition (146c) in the last step. We then have

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \lambda_1 - \lambda_2 \leqslant D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'})\ . \qquad\blacksquare$$

The following upper bound is easy to prove, although it has not found tremendous use.

Proposition 32. *The coherent relative entropy may be upper bounded as*

$$\begin{aligned}&\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\qquad\leqslant D_{\max}(\sigma_X\,\|\,\Gamma_X) - D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})\ ,\end{aligned} \tag{149}$$

writing $\sigma_X = t_{X\to R_X}(\rho_{R_X})$

Proof of Proposition 32. Consider an optimal solution $T_{X'R_X}$ and α for the primal semidefinite program. Then we have via the semidefinite constraints

$$\begin{aligned}\rho_{X'} = \mathrm{tr}_{R_X}\big[T_{X'R_X}\rho_{R_X}\big] &\leqslant 2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\,\mathrm{tr}_{R_X}\big[T_{X'R_X}\Gamma_{R_X}\big]\\ &\leqslant \alpha\,2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\,\Gamma_{X'}\ .\end{aligned} \tag{150}$$

By definition, we have

$$2^{D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})} = \min\{\mu:\ \mu\Gamma_{X'}\geqslant\rho_{X'}\}\ , \tag{151}$$

and thus we see that $\alpha 2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}$ is a candidate μ in this minimization. Hence $2^{D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})}\leqslant\alpha\,2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}$ and

$$\alpha\geqslant 2^{D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'}) - D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\ . \tag{152}$$

The claim then follows from $\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = -\log\alpha$. $\blacksquare$

The last of the upper bounds holds for the smooth coherent relative entropy. The present upper bound will be used to prove one direction of the asymptotic equipartition property.

Proposition 33. *Let $\rho_{X'R_X}$ be any quantum state, and denote the corresponding input state by $\sigma_X = t_{R_X\to X}(\rho_{R_X})$. Then for any $\epsilon,\epsilon',\epsilon''\geqslant 0$ such that $\bar\epsilon := \epsilon+\epsilon'+2\epsilon'' < 1$,*

$$\begin{aligned}&\bar{D}^{\epsilon''}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\leqslant D^{\epsilon}_{\max}(\sigma_X\,\|\,\Gamma_X) - D^{\epsilon'}_{\min,0}(\rho_{X'}\,\|\,\Gamma_{X'}) - \log(1-\bar\epsilon)\ .\end{aligned} \tag{153}$$

Proof of Proposition 33. Let $\tilde\rho_{X'R_X}$ be the quantum state which achieves the optimum for $\bar{D}^{\epsilon''}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$, i.e., satisfying $P(\tilde\rho_{X'R},\rho_{X'R})\leqslant\epsilon''$ and $\bar{D}^{\epsilon''}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \bar{D}_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$. The proof proceeds by constructing dual candidates for $2^{-\bar{D}_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$ in (59) achieving the value in the claim. Define the quantum states $\tilde\sigma_X$, $\tilde\rho_{X'}$ as the optimal ones in the optimizations defining the smooth min and max relative entropies, i.e., satisfying $P(\tilde\sigma_X,\sigma_X)\leqslant\epsilon$, $P(\tilde\rho_{X'},\rho_{X'})\leqslant\epsilon'$, as well as

$$D^{\epsilon}_{\max}(\sigma_X\,\|\,\Gamma_X) = D_{\max}(\tilde\sigma_X\,\|\,\Gamma_X)\ ; \tag{154a}$$

$$D^{\epsilon'}_{\min,0}(\rho_{X'}\,\|\,\Gamma_{X'}) = D_{\min,0}(\tilde\rho_{X'}\,\|\,\Gamma_{X'})\ . \tag{154b}$$

Let

$$\mu = 2^{-D_{\max}(\tilde\sigma_X\,\|\,\Gamma_X)}\left(\mathrm{tr}\,\Pi^{\tilde\rho_{X'}}_{X'}\,\Gamma_{X'}\right)^{-1}\ ; \tag{155a}$$

$$Z_{X'R_X} = \mu\,\Pi^{\tilde\rho_{X'}}_{X'}\otimes\mathbb{1}_{R_X}\ ; \tag{155b}$$

$$\omega_{X'} = \big[\mathrm{tr}(\Pi^{\tilde\rho_{X'}}_{X'}\,\Gamma_{X'})\big]^{-1}\,\Pi^{\tilde\rho_{X'}}_{X'}\ . \tag{155c}$$

Condition (61a) is automatically satisfied. Writing $\tilde{\sigma}_{R_X} = t_{X\to R_X}(\tilde{\sigma}_X)$, we have $D(\bar{\rho}_{R_X}, \tilde{\sigma}_{R_X}) \leqslant P(\bar{\rho}_{R_X}, \tilde{\sigma}_{R_X}) \leqslant P(\bar{\rho}_{R_X}, \rho_{R_X}) + P(\rho_{R_X}, \tilde{\sigma}_{R_X}) \leqslant \varepsilon'' + \varepsilon$; hence, there exists $\Delta_{R_X} \geqslant 0$ such that $\bar{\rho}_{R_X} \leqslant \tilde{\sigma}_{R_X} + \Delta_{R_X}$ with $\mathrm{tr}\,\Delta_{R_X} \leqslant \varepsilon'' + \varepsilon$. Then,

$$\begin{aligned}\bar{\rho}_{R_X}^{1/2} Z_{X'R_X} \bar{\rho}_{R_X}^{1/2} &= \mu\, \Pi_{X'}^{\tilde{\rho}_{X'}} \otimes \bar{\rho}_{R_X} \\ &\leqslant \mu\, \Pi_{X'}^{\tilde{\rho}_{X'}} \otimes (\tilde{\sigma}_{R_X} + \Delta_{R_X}) \\ &\leqslant \mu\, \Pi_{X'}^{\tilde{\rho}_{X'}} \otimes \big(2^{D_{\max}(\tilde{\sigma}_{R_X} \,\|\, \Gamma_{R_X})} \Gamma_{R_X} + \Delta_{R_X}\big) \\ &\leqslant \omega_{X'} \otimes \Gamma_{R_X} + \mu\, \mathbb{1}_{X'} \otimes \Delta_{R_X} \ , \end{aligned} \tag{156}$$

and we may define $X_{R_X} = \mu\, \Delta_R$ in order for constraint (61b) to be also satisfied. The attained dual objective value is

$$\text{obj.} = \mathrm{tr}(Z_{X'R_X}\, \bar{\rho}_{X'R_X}) - \mathrm{tr}(X_{R_X}) = \mu\, \big(\mathrm{tr}\big(\Pi_{X'}^{\tilde{\rho}_{X'}}\, \bar{\rho}_{X'}\big) - \varepsilon'' - \varepsilon\big) \ . \tag{157}$$

Analogously to the input state, now we have for the output state $D(\bar{\rho}_{X'}, \tilde{\rho}_{X'}) \leqslant P(\bar{\rho}_{X'}, \tilde{\rho}_{X'}) \leqslant P(\bar{\rho}_{X'}, \rho_{X'}) + P(\rho_{X'}, \tilde{\rho}_{X'}) \leqslant \varepsilon'' + \varepsilon'$; there must exist $\Delta_{X'} \geqslant 0$ with $\bar{\rho}_{X'} \geqslant \tilde{\rho}_{X'} - \Delta_{X'}$ and $\mathrm{tr}\,\Delta_{X'} \leqslant \varepsilon'' + \varepsilon'$. Hence, $\mathrm{tr}\big(\Pi_{X'}^{\tilde{\rho}_{X'}}\, \bar{\rho}_{X'}\big) \geqslant \mathrm{tr}\big(\Pi_{X'}^{\tilde{\rho}_{X'}}\, \tilde{\rho}_{X'}\big) - \mathrm{tr}\big(\Pi_{X'}^{\tilde{\rho}_{X'}}\, \Delta_{X'}\big) \geqslant 1 - \varepsilon'' - \varepsilon'$. Thus,

$$(157) \geqslant \mu\, (1 - \varepsilon - \varepsilon' - 2\varepsilon'') \ . \tag{158}$$

The claim follows by noting that $-\log \mu = D_{\max}^{\varepsilon}(\sigma_X \,\|\, \Gamma_X) - D_{\min,0}^{\varepsilon'}(\rho_{X'} \,\|\, \Gamma_{X'})$. ■

In order to formulate lower bounds on the coherent relative entropy, we introduce a generalization of the *Rob entropy* or *smooth S-entropy* [25]:

$$\begin{aligned} D_{\mathrm{r}}(\rho \,\|\, \Gamma) &= -\log \big\| \rho^{-1/2} \Gamma \rho^{-1/2} \big\|_\infty \\ &= -\log \min\{ \nu :\ \nu\rho \geqslant \Pi^\rho \Gamma \Pi^\rho \} \ ; \end{aligned} \tag{159}$$

$$D_{\mathrm{r}}^{\varepsilon}(\rho \,\|\, \Gamma) = \max_{\hat{\rho} \approx_\varepsilon \rho} D_{\mathrm{r}}(\hat{\rho} \,\|\, \Gamma) \ . \tag{160}$$

Proposition 34. *We have the lower bound*

$$\hat{D}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant D_{\mathrm{r}}(\sigma_X \,\|\, \Gamma_X) - D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) \ , \tag{161}$$

with $\sigma_X = t_{X\to R_X}(\rho_{R_X})$.

Proof of Proposition 34. Choose the primal candidate $T_{X'R_X} = \rho_{R_X}^{-1/2} \rho_{X'R_X} \rho_{R_X}^{-1/2}$. We have $\mathrm{tr}_{X'}\, T_{X'R_X} = \rho_{R_X}^{-1/2} \rho_{R_X} \rho_{R_X}^{-1/2} = \Pi_{R_X}^{\rho_{R_X}} \leqslant \mathbb{1}_{R_X}$ so our candidate satisifes (60a). Also (60c) is satisfied by construction, and $\mathrm{tr}_{R_X}\big(T_{X'R_X} \Gamma_{R_X}\big)$ is in the support of $\rho_{X'}$ and hence it lies in the support of $\Gamma_{X'}$. According to Proposition 12 we choose $\alpha = \big\|\Gamma_{X'}^{-1/2}\, \mathrm{tr}_{R_X}\big[T_{X'R_X} \Gamma_{R_X}\big] \Gamma_{X'}^{-1/2}\big\|_\infty$ and

$$\begin{aligned} &2^{-\hat{D}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})} \\ &\qquad \leqslant \alpha = \big\|\Gamma_{X'}^{-1/2}\, \mathrm{tr}_{R_X}\big[T_{X'R_X} \Gamma_{R_X}\big] \Gamma_{X'}^{-1/2}\big\|_\infty \\ &\qquad = \big\|\Gamma_{X'}^{-1/2}\, \mathrm{tr}_{R_X}\Big[T_{X'R_X} \Pi_{R_X}^{\rho_{R_X}} \Gamma_{R_X} \Pi_{R_X}^{\rho_{R_X}}\Big] \Gamma_{X'}^{-1/2}\big\|_\infty \\ &\qquad \leqslant 2^{-D_{\mathrm{r}}(\rho_{R_X} \,\|\, \Gamma_{R_X})} \big\|\Gamma_{X'}^{-1/2}\, \mathrm{tr}_{R_X}\big[T_{X'R_X} \rho_{R_X}\big] \Gamma_{X'}^{-1/2}\big\|_\infty \ , \end{aligned} \tag{162}$$

since by definition $\rho_{R_X}^{-1/2} \Gamma_{R_X} \rho_{R_X}^{-1/2} \leqslant 2^{-D_{\mathrm{r}}(\rho_{R_X} \,\|\, \Gamma_{R_X})} \mathbb{1}$ and thus $\Pi_{R_X}^{\rho_{R_X}} \Gamma_{R_X} \Pi_{R_X}^{\rho_{R_X}} \leqslant 2^{-D_{\mathrm{r}}(\rho_{R_X} \,\|\, \Gamma_{R_X})} \rho_{R_X}$. Then

$$\begin{aligned} (162) &= 2^{-D_{\mathrm{r}}(\rho_{R_X} \,\|\, \Gamma_{R_X})} \big\|\Gamma_{X'}^{-1/2} \rho_{X'} \Gamma_{X'}^{-1/2}\big\|_\infty \\ &= 2^{-D_{\mathrm{r}}(\sigma_X \,\|\, \Gamma_X)}\, 2^{D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'})} \ . \end{aligned}$$

■

The quantity $D_{\mathrm{r}}(\cdot \,\|\, \cdot)$, when smoothed, is essentially equal to the min-relative entropy: These two differ by a term which is logarithmic in the failure probability. In this way, the smooth quantity $D_{\mathrm{r}}^{\varepsilon}(\cdot \,\|\, \cdot)$ may be related to a better known quantity with an operational interpretation.

Proposition 35. *Let* $\varepsilon > 0$. *Then*

$$D_{\mathrm{r}}^{\varepsilon}(\rho \,\|\, \Gamma) \geqslant D_{\min,0}(\rho \,\|\, \Gamma) + \log \varepsilon' \ , \tag{163}$$

where $\varepsilon' = \varepsilon^2 / \big(2 + \varepsilon^2\big)$, *or equivalently,* $\varepsilon = \sqrt{2\varepsilon'/(1-\varepsilon')}$.

Proof of Proposition 35. The proof of this proposition proceeds via the hypothesis testing relative entropy, $D_{\mathrm{H}}^{\eta}(\rho \,\|\, \Gamma)$. Let $\varepsilon' = \varepsilon^2 / \big(2 + \varepsilon^2\big)$ and let $\eta = 1 - \varepsilon'$. The hypothesis testing relative entropy can be written as the solution of a semidefinite program [23]. Specifically, there exists $Q \geqslant 0$, $\mu \geqslant 0$ and $X \geqslant 0$ such that

$$2^{-D_{\mathrm{H}}^{\eta}(\rho \,\|\, \Gamma)} = \frac{1}{\eta}\, \mathrm{tr}\,[Q\Gamma] = \mu - \frac{\mathrm{tr}\,X}{\eta} \ , \tag{164}$$

with Q, μ and X satisfying the conditions

$$Q \leqslant \mathbb{1} \ ; \tag{165a}$$

$$\mathrm{tr}\,[Q\rho] \geqslant \eta \ ; \tag{165b}$$

$$\mu\rho \leqslant \Gamma + X \ . \tag{165c}$$

In addition, the complementary slackness relations for these variables read

$$XQ = X \ ; \tag{166a}$$

$$\mathrm{tr}\,(Q\rho) = \eta \ ; \tag{166b}$$

$$Q\,(\mu\rho - \Gamma - X) = 0 \ . \tag{166c}$$

Define $\bar{\rho} = \Pi^Q \rho \Pi^Q$, where Π^Q is the projector onto the support of Q. Apply $Q^{-1}\,(\cdot)\,\Pi^Q$ onto (166c) to obtain

$$\mu\bar{\rho} = \Pi^Q \Gamma \Pi^Q + \Pi^Q X \Pi^Q \geqslant \Pi^Q \Gamma \Pi^Q \ . \tag{167}$$

In addition, because $\Pi^Q \Gamma \Pi^Q$ has support on Π^Q, then $\bar{\rho}$ must also have support on the full of Π^Q, i.e. $\Pi^{\bar{\rho}} = \Pi^Q$. So, by definition of $D_{\mathrm{r}}(\bar{\rho} \,\|\, \Gamma)$ have that

$$2^{-D_{\mathrm{r}}(\bar{\rho} \,\|\, \Gamma)} \leqslant \mu \ . \tag{168}$$

Also, define $\bar{\rho}' = \bar{\rho}/\mathrm{tr}\,\bar{\rho}$, and we can see by Lemma 45 that $P(\rho, \bar{\rho}') \leqslant \sqrt{2\varepsilon'/(1-\varepsilon')} = \varepsilon$. Also, $2^{-D_{\mathrm{r}}(\bar{\rho}' \,\|\, \Gamma)} \leqslant 2^{-D_{\mathrm{r}}(\bar{\rho} \,\|\, \Gamma)}$ by definition of $D_{\mathrm{r}}(\cdot \,\|\, \cdot)$. Then $\bar{\rho}'$ is a valid optimization candidate in the definition of $D_{\mathrm{r}}^{\tilde{\varepsilon}}(\rho \,\|\, \Gamma)$ and

$$2^{-D_{\mathrm{r}}^{\tilde{\varepsilon}}(\rho \,\|\, \Gamma)} \leqslant 2^{-D_{\mathrm{r}}(\bar{\rho}' \,\|\, \Gamma)} \leqslant \mu \ . \tag{169}$$

It thus remains to show that $\mu \leqslant \varepsilon'^{-1}\, 2^{-D_{\min,0}(\rho \,\|\, \Gamma)}$. Apply $\mathrm{tr}\,(\Pi^\rho\,(\cdot))$ onto the constraint (165c) to obtain

$$\mu \leqslant \mathrm{tr}\,(\Pi^\rho \Gamma) + \mathrm{tr}\,(\Pi^\rho X) \leqslant \mathrm{tr}\,(\Pi^\rho \Gamma) + \mathrm{tr}\,(X) \ . \tag{170}$$

Now, because of (164), we have $0 \leqslant \mathrm{tr}\,[Q\Gamma] = \mu\eta - \mathrm{tr}\,X$, and thus $\mathrm{tr}\,X \leqslant \mu\eta$. Combining with (170) gives

$$\mu\,(1 - \eta) \leqslant \mathrm{tr}\,(\Pi^\rho \Gamma) \ ; \tag{171}$$

since $\varepsilon' = 1 - \eta$ and $\mathrm{tr}\,(\Pi^\rho \Gamma) = 2^{-D_{\min,0}(\rho \,\|\, \Gamma)}$ we have $\mu \leqslant (1/\varepsilon')\, 2^{-D_{\min,0}(\rho \,\|\, \Gamma)}$ and the claim follows. ■

The following proposition gives a lower bound to the smooth coherent relative entropy. This will prove crucial to the proof of the asymptotic equipartition theorem.

Proposition 36. *Let* $\varepsilon', \varepsilon'' \geqslant 0$ *and* $\varepsilon''' > 0$. *Let* $\varepsilon \geqslant 2\sqrt{2\varepsilon'} + 2\sqrt{2(\varepsilon'' + \varepsilon''')}$. *Then*

$$\begin{aligned} &\bar{D}_{X\to X'}^{\varepsilon}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\qquad \geqslant D_{\min,0}^{\varepsilon''}(\sigma_X \,\|\, \Gamma_X) - D_{\max}^{\varepsilon'}(\rho_{X'} \,\|\, \Gamma_{X'}) + \log \frac{\varepsilon'''^2}{2 + \varepsilon'''^2} \ , \end{aligned} \tag{172}$$

where $\sigma_X = t_{X\to R_X}(\rho_{R_X})$. *(Proof on page 19.)*

Proof of Proposition 36. Let $\tilde{\rho}_{R_X}, \tilde{\rho}_{X'}$ be quantum states which are optimal smoothed states for the quantities

$$D^{\varepsilon''}_{\min,0}(\rho_{R_X} \,\|\, \Gamma_{R_X}) = D_{\min,0}(\tilde{\rho}_{R_X} \,\|\, \Gamma_{R_X}) \ . \tag{173a}$$

$$D^{\varepsilon'}_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) = D_{\max}(\tilde{\rho}_{X'} \,\|\, \Gamma_{X'}) \ . \tag{173b}$$

With $\varepsilon''' > 0$ and using Proposition 35, we know that

$$D^{\varepsilon'''}_{\mathrm{r}}(\tilde{\rho}_{R_X} \,\|\, \Gamma_{R_X}) \geqslant D_{\min,0}(\tilde{\rho}_{R_X} \,\|\, \Gamma_{R_X}) + \log \frac{\varepsilon'''^2}{2+\varepsilon'''^2} \ . \tag{174}$$

Let $\tilde{\tilde{\rho}}_{R_X}$ be the optimal smoothed state for $D^{\varepsilon'''}_{\mathrm{r}}(\tilde{\rho}_{R_X} \,\|\, \Gamma_{R_X})$, such that

$$D_{\mathrm{r}}(\tilde{\tilde{\rho}}_{R_X} \,\|\, \Gamma_{R_X}) = D^{\varepsilon'''}_{\mathrm{r}}(\tilde{\rho}_{R_X} \,\|\, \Gamma_{R_X}) \ . \tag{175}$$

At this point, we have

$$\begin{aligned} &D_{\mathrm{r}}(\tilde{\tilde{\rho}}_{R_X} \,\|\, \Gamma_{R_X}) - D_{\max}(\tilde{\rho} \,\|\, \Gamma_{X'}) \\ &\qquad \geqslant D^{\varepsilon''}_{\min,0}(\rho_{R_X} \,\|\, \Gamma_{R_X}) - D^{\varepsilon'}_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) + \log \frac{\varepsilon'''^2}{2+\varepsilon'''^2} \ , \end{aligned} \tag{176}$$

with

$$P(\tilde{\rho}_{X'}, \rho_{X'}) \leqslant \varepsilon' \ ; \quad P(\tilde{\rho}_{R_X}, \rho_{R_X}) \leqslant \varepsilon'' \ ; \quad P(\tilde{\tilde{\rho}}_{R_X}, \tilde{\rho}_{R_X}) \leqslant \varepsilon''' \ . \tag{177}$$

Now, we'll apply Lemma 46 twice to construct a state close to $\rho_{X'R_X}$ which has marginals $\tilde{\rho}_{X'}$ and $\tilde{\tilde{\rho}}_{R_X}$ exactly. Let $\tau_{X'R_X}$ be the quantum state given by Lemma 46 satisfying

$$\tau_{X'} = \tilde{\rho}_{X'} \ ; \qquad \tau_{R_X} = \rho_{R_X} \ ; \qquad P(\tau_{X'R_X}, \rho_{X'R_X}) \leqslant 2\sqrt{2\varepsilon'} \ . \tag{178}$$

Applying Lemma 46 again, let $\tau'_{X'R}$ be a quantum state close to $\tau_{X'R}$ such that

$$\tau'_{X'} = \tilde{\rho}_{X'} \ ; \quad \tau'_{R_X} = \tilde{\tilde{\rho}}_{R_X} \ ; \quad P(\tau'_{X'R_X}, \tau_{X'R_X}) \leqslant 2\sqrt{2(\varepsilon''+\varepsilon''')} \ . \tag{179}$$

We thus have by triangle inequality

$$P(\tau'_{X'R_X}, \rho_{X'R_X}) \leqslant 2\sqrt{2\varepsilon'} + 2\sqrt{2(\varepsilon''+\varepsilon''')} \ . \tag{180}$$

By Proposition 34 we can now write

$$\begin{aligned} \hat{D}_{X\to X'}(\tau'_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) &\geqslant D_{\mathrm{r}}(\tau'_{R_X} \,\|\, \Gamma_{R_X}) - D_{\max}(\tau'_{X'} \,\|\, \Gamma_{X'}) \\ &= D_{\mathrm{r}}(\tilde{\tilde{\rho}}_{R_X} \,\|\, \Gamma_{R_X}) - D_{\max}(\tilde{\rho}_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned} \tag{181}$$

Observe now that $\tau'_{X'R_X}$ is a valid optimization candidate for $\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$. Hence

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant \hat{D}_{X\to X'}(\tau'_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \ . \tag{182}$$

Finally, inequality (182) followed by (181) and (176) provides us the seeked lower bound. ■

We also have a bound which applies to product states, given in terms of min- and max-relative entropies of input and output. Physically, it asserts that a possible strategy for implementing the product state process matrix is to completely erase the input state (at a cost given by the min-relative entropy), and subsequently prepare the required output state (at a yield given by the max-relative entropy).

Proposition 37 (coherent relative entropy for product states)**.** *For states σ_X and $\rho_{X'}$, we have*

$$\begin{aligned} &\hat{D}_{X\to X'}(t_{X\to R_X}(\sigma_X) \otimes \rho_{X'} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\qquad \geqslant D_{\min,0}(\sigma_X \,\|\, \Gamma_X) - D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned} \tag{183}$$

Proof of Proposition 37. Write $\sigma_{R_X} = t_{X\to R_X}(\sigma_X)$. Choose $T_{X'R_X} = \Pi^{\sigma_{R_X}}_{R_X} \otimes \rho_{X'}$. This choice trivially satisfies (60a). Also, $\sigma^{1/2}_{R_X} T_{X'R_X} \sigma^{1/2}_{R_X} = \sigma_{R_X} \otimes \rho_{X'}$ so (60c) is also satisfied. We have that $\mathrm{tr}_{R_X} T_{X'R_X} \Gamma_{R_X}$ lies in the support of $\Gamma_{X'}$ because $\rho_{X'}$ does so, and as per Proposition 12 the optimal value of α corresponding to this $T_{X'R_X}$ is given by

$$\begin{aligned} \alpha &= \big\| \Gamma^{-1/2}_{X'} \, \mathrm{tr}_{R_X}\big[T_{X'R_X} \Gamma_{R_X}\big] \Gamma^{-1/2}_{X'} \big\|_\infty \\ &= \big\| \Gamma^{-1/2}_{X'} \, \mathrm{tr}_{R_X}\Big[(\Pi^{\sigma_{R_X}}_{R_X} \otimes \rho_{X'})\Gamma_{R_X}\Big] \Gamma^{-1/2}_{X'} \big\|_\infty \\ &= \mathrm{tr}_{R_X}[\Pi^{\sigma_{R_X}}_{R_X} \Gamma_{R_X}] \, \big\| \Gamma^{-1/2}_{X'} \rho_{X'} \Gamma^{-1/2}_{X'} \big\|_\infty \\ &= 2^{-D_{\min,0}(\sigma_{R_X} \,\|\, \Gamma_{R_X})} \, 2^{D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'})} \ . \end{aligned} \tag{184}$$

This choice of α and $T_{X'R_X}$ is feasible for $2^{-\hat{D}_{X\to X'}(\sigma_{R_X} \otimes \rho_{X'} \,\|\, \Gamma_X, \Gamma_{X'})}$, hence

$$\begin{aligned} &\hat{D}_{X\to X'}(\sigma_{R_X} \otimes \rho_{X'} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\qquad \geqslant D_{\min,0}(\sigma_X \,\|\, \Gamma_X) - D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned}$$

■

H. Asymptotic equipartition property

Finally, the coherent relative entropy also obeys an asymptotic equipartition property in the i.i.d. limit. In this limit, the coherent relative entropy converges to the difference of relative entropies of the input and the output to the respective Γ operators.

Both versions of the coherent relative entropy we have introduced have the same asymptotic behavior for small ε. For completeness we present the detailed statements, including the ranges of ε for which the property is proven for each quantity.

Proposition 38 (Asymptotic equipartition property)**.** *For any $\Gamma_X, \Gamma_{X'} \geqslant 0$, for any quantum state $\rho_{X'R_X}$, and for any $0 < \varepsilon < 1/2$,*

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \bar{D}^{\varepsilon}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big) \\ &\qquad = D(\sigma_X \,\|\, \Gamma_X) - D(\rho_{X'} \,\|\, \Gamma_{X'}) \ , \end{aligned} \tag{185}$$

where $\sigma_X = t_{X\to X'}(\rho_{R_X})$.

Similarly, for any $0 < \varepsilon < (18)^{-4}$,

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \hat{D}^{\varepsilon}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big) \\ &\qquad = D(\sigma_X \,\|\, \Gamma_X) - D(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned} \tag{186}$$

(Proof on page 20.)

While the original asymptotic equipartition statements in the context of smooth entropies (e.g., refs. [12, 14]) considered first the limit $n\to\infty$, and then $\varepsilon\to 0$, the above proposition is slightly more general in that the limit $\varepsilon\to 0$ is not necessary (in line with, e.g., refs. [4, 22, 23]). However, one may ask if it is possible to take $\varepsilon\to 0$ simultaneously with $n\to\infty$. We may indeed prove such a statement using recent results on moderate deviation analysis [26, 27].

Proposition 39 (Asymptotic equipartition property, take 2)**.** *Consider any $\Gamma_X, \Gamma_{X'} \geqslant 0$, and any quantum state $\rho_{X'R_X}$. Let*

(ε_n) *be a sequence such that* $\varepsilon_n \to 0$ *and* $-(1/n)\ln(\varepsilon_n) \to 0$. *Then*[7]

$$\lim_{n\to\infty} \frac{1}{n} \bar{D}^{\varepsilon_n}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big) = D(\sigma_X \,\|\, \Gamma_X) - D(\rho_{X'} \,\|\, \Gamma_{X'}) \;, \tag{187}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \hat{D}^{\varepsilon_n}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big) = D(\sigma_X \,\|\, \Gamma_X) - D(\rho_{X'} \,\|\, \Gamma_{X'}) \;. \tag{188}$$

where $\sigma_X = t_{X\to X'}(\rho_{R_X})$. *(Proof on page 21.)*

The proof of the asymptotic equipartition follows from bounds we have derived above using the min- and max-relative entropies. The latter have known asymptotic behavior, and they converge to the usual quantum relative entropy [14]. Note that our definitions of the smooth min- and max-relative entropy differ in minor details from the ones originally introduced in ref. [14]. For completeness, we hence provide an adapted proof of the asymptotic equipartition property for the min- and max-relative entropy. Our proof proceeds via the hypothesis testing entropy, whose asymptotic behavior has been thoroughly studied [22, 23, 28–31]. This will allow us to prove Proposition 39 via direct application of the results in refs. [26, 27].

Lemma 40 (Bounds for min- and max-relative entropy in terms of hypothesis testing entropy)**.** *The following bounds hold for any* $0<\varepsilon<1/2$*:*

$$D^{\varepsilon}_{\min,0}(\sigma \,\|\, \Gamma) \leqslant D^{1-\varepsilon}_{\mathrm{H}}(\sigma \,\|\, \Gamma) - \log(1-\varepsilon) \;; \tag{189a}$$

$$D^{\varepsilon}_{\min,0}(\sigma \,\|\, \Gamma) \geqslant D^{1-\varepsilon'}_{\mathrm{H}}(\sigma \,\|\, \Gamma) - \log\frac{1-\varepsilon'}{\varepsilon'} \;; \tag{189b}$$

$$D^{\varepsilon}_{\max}(\rho \,\|\, \Gamma) \geqslant D^{2\varepsilon}_{\mathrm{H}}(\rho \,\|\, \Gamma) - 1 \;; \tag{189c}$$

$$D^{\varepsilon}_{\max}(\rho \,\|\, \Gamma) \leqslant D^{\varepsilon^2/2}_{\mathrm{H}}(\rho \,\|\, \Gamma) - \log(1-\varepsilon) \;, \tag{189d}$$

for any $0<\varepsilon'\leqslant \varepsilon^2/(4+2\varepsilon^2)$*; we may choose, e.g.,* $\varepsilon'=\varepsilon^2/6$.

Recall that, as a direct consequence of Quantum Stein's lemma [23, 28, 29], we have that for all $0<\varepsilon<1$,

$$\lim_{n\to\infty} \frac{1}{n} D^{\varepsilon}_{\mathrm{H}}(\sigma^{\otimes n} \,\|\, \Gamma^{\otimes n}) = D(\sigma \,\|\, \Gamma) \;. \tag{190}$$

As an immediate consequence of Lemma 40 and of (190), we find that for any $0<\varepsilon<1/2$,

$$\lim_{n\to\infty} \frac{1}{n} D^{\varepsilon}_{\min,0}(\sigma^{\otimes n}_X \,\|\, \Gamma^{\otimes n}_X) = D(\sigma_X \,\|\, \Gamma_X) \;; \tag{191a}$$

$$\lim_{n\to\infty} \frac{1}{n} D^{\varepsilon}_{\max}(\rho^{\otimes n}_{X'} \,\|\, \Gamma^{\otimes n}_{X'}) = D(\rho_{X'} \,\|\, \Gamma_{X'}) \;, \tag{191b}$$

noting that terms which scale sublinearly in n, for instance $\log(1-\varepsilon)$, disappear because of the factor $1/n$ in the limit $n\to\infty$.

Proof of Lemma 40. Let $\tilde\sigma$ be optimal for $D^{\varepsilon}_{\min,0}(\sigma \,\|\, \Gamma)$, i.e., $D^{\varepsilon}_{\min,0}(\sigma \,\|\, \Gamma) = D_{\min,0}(\tilde\sigma \,\|\, \Gamma) = -\log \mathrm{tr}[\Pi^{\tilde\sigma}\Gamma]$ with $P(\sigma,\tilde\sigma)\leqslant\varepsilon$. As $\sigma \geqslant \tilde\sigma - \Delta$ for some $\Delta\geqslant 0$ with $\mathrm{tr}\Delta\leqslant\varepsilon$, we have that $\mathrm{tr}[\Pi^{\tilde\sigma}\sigma] \geqslant 1-\mathrm{tr}(\Pi^{\tilde\sigma}\Delta) \geqslant 1-\varepsilon$. Then $\Pi^{\tilde\sigma}$ is feasible in the primal program for $2^{-D^{1-\varepsilon}_{\mathrm{H}}(\sigma\|\Gamma)}$, achieving the value $(1-\varepsilon)^{-1}\,\mathrm{tr}(\Pi^{\tilde\sigma}\Gamma)$. Hence, for any $0<\varepsilon<1$,

$$D^{\varepsilon}_{\min,0}(\sigma \,\|\, \Gamma) \leqslant D^{1-\varepsilon}_{\mathrm{H}}(\sigma \,\|\, \Gamma) - \log(1-\varepsilon) \;. \tag{192}$$

Conversely, for any $0<\varepsilon'<1/2$ to be fixed later, let Q be primal optimal for $2^{-D^{1-\varepsilon'}_{\mathrm{H}}(\sigma\|\Gamma)} = (1-\varepsilon')^{-1}\,\mathrm{tr}(Q\Gamma)$ with $\mathrm{tr}(Q\sigma)\geqslant 1-\varepsilon'$. For $\eta=\varepsilon'$, Let P^η be the projector onto the eigenspaces of Q associated to eigenvalues greater than or equal to η, and hence satisfying $\eta P^\eta \leqslant Q$. It follows that $\mathrm{tr}(Q\Gamma)\geqslant \eta\,\mathrm{tr}(P^\eta\Gamma)$. Now, define $\tilde\sigma = P^\eta\sigma P^\eta/\mathrm{tr}(P^\eta\sigma)$, noting that $\mathrm{tr}(P^\eta\sigma) \geqslant \mathrm{tr}(P^\eta Q P^\eta \sigma) \geqslant \mathrm{tr}(Q\sigma) - \mathrm{tr}((\mathbb{1}-P^\eta)\,Q\,(\mathbb{1}-P^\eta)\,\sigma) \geqslant 1-\varepsilon'-\eta$ (recall that all eigenvalues of $(\mathbb{1}-P^\eta)\,Q\,(\mathbb{1}-P^\eta)$ are less than η). Using Lemma 45, we see that $P(\tilde\sigma,\sigma) \leqslant \sqrt{2(\varepsilon'+\eta)}/\sqrt{1-\varepsilon'-\eta} = \sqrt{4\varepsilon'/(1-2\varepsilon')}$. Now $\tilde\sigma$ is a valid candidate for the smoothing in $D^{\sqrt{4\varepsilon'/(1-2\varepsilon')}}_{\min,0}(\sigma \,\|\, \Gamma)$, and hence $D^{\sqrt{4\varepsilon'/(1-2\varepsilon')}}_{\min,0}(\sigma \,\|\, \Gamma) \geqslant -\log\mathrm{tr}(P^\eta\Gamma) \geqslant -\log[\eta^{-1}\,\mathrm{tr}(Q\Gamma)] = -\log[((1-\varepsilon')/\varepsilon')\,(1-\varepsilon')^{-1}\,\mathrm{tr}(Q\Gamma)] = D^{1-\varepsilon'}_{\mathrm{H}}(\sigma \,\|\, \Gamma) - \log[(1-\varepsilon')/\varepsilon']$. Now consider any $0<\varepsilon<1/2$ and assume that $0<\varepsilon'\leqslant\varepsilon^2/(4+2\varepsilon^2)$, noting that $0<\varepsilon'<1/2$. We have $\varepsilon'(4+2\varepsilon^2)\leqslant\varepsilon^2 \Leftrightarrow 4\varepsilon'\leqslant\varepsilon^2(1-2\varepsilon') \Leftrightarrow \varepsilon\geqslant\sqrt{4\varepsilon'/(1-2\varepsilon')}$, and thus $D^{\varepsilon}_{\min,0}(\sigma \,\|\, \Gamma) \geqslant D^{\sqrt{4\varepsilon'/(1-2\varepsilon')}}_{\min,0}(\sigma \,\|\, \Gamma)$. Hence, for any $0<\varepsilon<1/2$ and for any $0<\varepsilon'\leqslant\varepsilon^2/(4+2\varepsilon^2)$, we have:

$$D^{\varepsilon}_{\min,0}(\sigma \,\|\, \Gamma) \geqslant D^{1-\varepsilon'}_{\mathrm{H}}(\sigma \,\|\, \Gamma) - \log\frac{1-\varepsilon'}{\varepsilon'} \;. \tag{193}$$

For the max-relative entropy, for any ρ,Γ and for any $0<\varepsilon<1/2$, let $\tilde\rho$ be a normalized quantum state such that $D^{\varepsilon}_{\max}(\rho \,\|\, \Gamma) = D_{\max}(\tilde\rho \,\|\, \Gamma)$. Let Q be primal optimal for $2^{-D^{2\varepsilon}_{\mathrm{H}}(\rho\|\Gamma)} = (2\varepsilon)^{-1}\,\mathrm{tr}(Q\Gamma)$, such that $\mathrm{tr}(Q\rho)\geqslant 2\varepsilon$. But $\tilde\rho\geqslant\rho-\Delta$ for a $\Delta\geqslant 0$ with $\mathrm{tr}\Delta\leqslant\varepsilon$, since $D(\tilde\rho,\rho)\leqslant\varepsilon$, and thus $\mathrm{tr}(Q\tilde\rho)\geqslant 2\varepsilon-\varepsilon=\varepsilon$. Then Q is primal feasible also for $D^{\varepsilon}_{\mathrm{H}}(\tilde\rho \,\|\, \Gamma)$ and $2^{-D^{\varepsilon}_{\mathrm{H}}(\tilde\rho\|\Gamma)} \leqslant \varepsilon^{-1}\,\mathrm{tr}(Q\Gamma) = 2\cdot 2^{-D^{2\varepsilon}_{\mathrm{H}}(\rho\|\Gamma)}$. Then, using [23, Prop. 4.1], $D_{\max}(\tilde\rho \,\|\, \Gamma) \geqslant D^{\varepsilon}_{\mathrm{H}}(\tilde\rho \,\|\, \Gamma) \geqslant D^{2\varepsilon}_{\mathrm{H}}(\rho \,\|\, \Gamma) - 1$, and hence

$$D^{\varepsilon}_{\max}(\rho \,\|\, \Gamma) \geqslant D^{2\varepsilon}_{\mathrm{H}}(\rho \,\|\, \Gamma) - 1 \;. \tag{194}$$

For a lower bound on $D^{\varepsilon}_{\max}$, we invoke [23, Prop. 4.1]; however the quantity called $D^{\varepsilon}_{\max}$ there optimizes over subnormalized states whereas we optimize over normalized states only, so we have to work a little more. For any subnormalized state $\tilde\rho$ with $\mathrm{tr}\,\tilde\rho\geqslant 1-\varepsilon$, we have by definition that $2^{D_{\max}(\tilde\rho\|\Gamma)} = \|\Gamma^{-1/2}\,\tilde\rho\,\Gamma^{-1/2}\|_\infty = \mathrm{tr}(\tilde\rho)\,2^{D_{\max}(\tilde\rho/\mathrm{tr}\tilde\rho\|\Gamma)} \geqslant (1-\varepsilon)\cdot 2^{D_{\max}(\tilde\rho/\mathrm{tr}\tilde\rho\|\Gamma)}$, and hence

$$\min_{\substack{\tilde\rho:\ \mathrm{tr}\tilde\rho\leqslant 1\\ P(\tilde\rho,\rho)\leqslant\varepsilon}} D_{\max}(\tilde\rho \,\|\, \Gamma) \geqslant \min_{\substack{\tilde\rho:\ \mathrm{tr}\tilde\rho\leqslant 1\\ P(\tilde\rho,\rho)\leqslant\varepsilon}} D_{\max}(\tilde\rho/\mathrm{tr}\,\tilde\rho \,\|\, \Gamma) + \log(1-\varepsilon) = D^{\varepsilon}_{\max}(\rho \,\|\, \Gamma) + \log(1-\varepsilon) \;. \tag{195}$$

Then, invoking [23, Prop. 4.1] for any $0<\varepsilon<1$, and chaining with the above inequality,

$$D^{\varepsilon^2/2}_{\mathrm{H}}(\rho \,\|\, \Gamma) \geqslant D^{\varepsilon}_{\max}(\rho \,\|\, \Gamma) + \log(1-\varepsilon) \;. \quad \blacksquare$$

Proof of Proposition 38. We start by upper bounding the coherent relative entropy $\bar{D}^{\varepsilon}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big)$. Thanks to Proposition 33, choosing $\tilde\varepsilon = \tilde\varepsilon' = (1-2\varepsilon)/137414920$ with $\tilde{\tilde\varepsilon} = \tilde\varepsilon + \tilde\varepsilon' + 2\varepsilon$, and then using (191),

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \bar{D}^{\varepsilon}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big) \\
&\quad\leqslant \lim_{n\to\infty} \frac{1}{n} \Big[D^{\tilde\varepsilon}_{\max}\big(\rho^{\otimes n}_{R_X} \,\big\|\, \Gamma^{\otimes n}_{R_X}\big) - D^{\tilde\varepsilon'}_{\min,0}\big(\rho^{\otimes n}_{X'} \,\big\|\, \Gamma^{\otimes n}_{X'}\big) - \log\big[\varepsilon(1-\tilde{\tilde\varepsilon})\big]\Big] \\
&\quad= D(\rho_{R_X} \,\|\, \Gamma_{R_X}) - D(\rho_{X'} \,\|\, \Gamma_{X'}) \;.
\end{aligned} \tag{196}$$

[7] The condition on the sequence (ε_n) corresponds to requiring that (ε_n) results from a *moderate sequence* as defined in [26]. It is equivalent to requiring that the sequence (ε_n) converges to zero slower than $\exp(-n)$. (For example, this is satisfied if $\varepsilon_n \sim 1/\mathrm{poly}(n)$.)

The lower bound is given by Proposition 36: Choosing $\hat{\varepsilon}' = \hat{\varepsilon}'' = \hat{\varepsilon}''' = \varepsilon^2/197334000868$,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\bar{D}^{\varepsilon}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X}\,\big\|\,\Gamma^{\otimes n}_X,\Gamma^{\otimes n}_{X'}\big)\\ \geqslant \lim_{n\to\infty}\frac{1}{n}\left[D^{\hat{\varepsilon}''}_{\min,0}\big(\rho^{\otimes n}_{R_X}\,\|\,\Gamma^{\otimes n}_{R_X}\big) - D^{\hat{\varepsilon}'}_{\max}\big(\rho^{\otimes n}_{X'}\,\|\,\Gamma^{\otimes n}_{X'}\big) + \log\frac{\hat{\varepsilon}'''^2}{2+\hat{\varepsilon}'''^2}\right]\\ = D(\rho_{R_X}\,\|\,\Gamma_{R_X}) - D(\rho_{X'}\,\|\,\Gamma_{X'})\ .\end{aligned}$$

Equation (186) follows directly from (185), using the relations given by Proposition 26. ■

Proof of Proposition 39. Moderate deviation analysis provides a full characterization of the second-order asymptotic behavior of the hypothesis testing entropy [26, 27] in cases where $\varepsilon\to 0$ simultaneously with $n\to\infty$. For our purposes and for simplicity we consider the leading order only: For any sequence $(\hat{\varepsilon}_n)$ such that $\hat{\varepsilon}_n\to 0$ and $-(1/n)\ln(\hat{\varepsilon}_n)\to 0$, it holds that

$$\lim_{n\to\infty}\frac{1}{n}D^{\hat{\varepsilon}_n}_{\mathrm{H}}\big(\sigma^{\otimes n}\,\|\,\Gamma^{\otimes n}\big) = D(\sigma\,\|\,\Gamma)\ ; \tag{197a}$$

$$\lim_{n\to\infty}\frac{1}{n}D^{1-\hat{\varepsilon}_n}_{\mathrm{H}}\big(\sigma^{\otimes n}\,\|\,\Gamma^{\otimes n}\big) = D(\sigma\,\|\,\Gamma)\ . \tag{197b}$$

So, we proceed analogously to the proof of Proposition 38 via the bounds we determined on the coherent relative entropy in terms of the min- and max-relative entropies.

We invoke Proposition 33 choosing $\tilde{\varepsilon}_n = \tilde{\varepsilon}'_n = \min(\varepsilon_n,(1-2\varepsilon_n)/3)$ and $\tilde{\tilde{\varepsilon}}_n = \tilde{\varepsilon}_n + \tilde{\varepsilon}'_n + 2\varepsilon_n$, further observing that $\tilde{\tilde{\varepsilon}}_n < 1$ and $\tilde{\tilde{\varepsilon}}_n \leqslant 4\varepsilon_n$. Then

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\bar{D}^{\varepsilon_n}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X}\,\big\|\,\Gamma^{\otimes n}_X,\Gamma^{\otimes n}_{X'}\big)\\ \leqslant \lim_{n\to\infty}\frac{1}{n}\left[D^{\tilde{\varepsilon}_n}_{\max}\big(\rho^{\otimes n}_{R_X}\,\|\,\Gamma^{\otimes n}_{R_X}\big) - D^{\tilde{\varepsilon}'_n}_{\min,0}\big(\rho^{\otimes n}_{X'}\,\|\,\Gamma^{\otimes n}_{X'}\big) - \log\left[\varepsilon_n(1-\tilde{\tilde{\varepsilon}}_n)\right]\right]\\ \leqslant \lim_{n\to\infty}\frac{1}{n}\Big[D^{\tilde{\varepsilon}_n^2/2}_{\mathrm{H}}\big(\rho^{\otimes n}_{R_X}\,\|\,\Gamma^{\otimes n}_{R_X}\big) - D^{1-\tilde{\varepsilon}'^2_n/6}_{\mathrm{H}}\big(\rho^{\otimes n}_{X'}\,\|\,\Gamma^{\otimes n}_{X'}\big)\\ - \log(1-\tilde{\varepsilon}_n) + \log\left[\frac{1-\tilde{\varepsilon}'^2_n/6}{\tilde{\varepsilon}'^2_n/6}\right] - \log\left[\varepsilon_n(1-\tilde{\tilde{\varepsilon}}_n)\right]\Big]\\ \leqslant \lim_{n\to\infty}\frac{1}{n}\Big[D^{\tilde{\varepsilon}_n^2/2}_{\mathrm{H}}\big(\rho^{\otimes n}_{R_X}\,\|\,\Gamma^{\otimes n}_{R_X}\big) - D^{1-\tilde{\varepsilon}'^2_n/6}_{\mathrm{H}}\big(\rho^{\otimes n}_{X'}\,\|\,\Gamma^{\otimes n}_{X'}\big)\\ + \log(\mathrm{poly}(\varepsilon_n)/\mathrm{poly}(\varepsilon_n))\Big]\\ = D(\rho_{R_X}\,\|\,\Gamma_{R_X}) - D(\rho_{X'}\,\|\,\Gamma_{X'})\ ,\end{aligned} \tag{198}$$

where we used Lemma 40 in the second inequality, where $\mathrm{poly}(\varepsilon_n)$ denotes a polynomial in ε_n of arbitrary but constant degree, and where we used (197) for the last equality, noting that $(1/n)\log(\mathrm{poly}(\varepsilon_n))\to 0$ as $n\to\infty$, using the assumption in the claim that $-(1/n)\ln(\varepsilon_n)\to 0$ as $n\to\infty$.

The other direction follows similarly: We apply Proposition 36 choosing $\bar{\varepsilon}'_n = \bar{\varepsilon}''_n = \bar{\varepsilon}'''_n = \varepsilon_n^2/64$, such that $2\sqrt{2\bar{\varepsilon}'_n} + 2\sqrt{2(\bar{\varepsilon}''_n+\bar{\varepsilon}'''_n)} = 2\sqrt{2\varepsilon_n^2/64} + 2\sqrt{4\varepsilon_n^2/64} \leqslant \varepsilon_n$; then

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\bar{D}^{\varepsilon_n}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X}\,\big\|\,\Gamma^{\otimes n}_X,\Gamma^{\otimes n}_{X'}\big)\\ \geqslant \lim_{n\to\infty}\frac{1}{n}\left[D^{\bar{\varepsilon}''_n}_{\min,0}\big(\rho^{\otimes n}_{R_X}\,\|\,\Gamma^{\otimes n}_{R_X}\big) - D^{\bar{\varepsilon}'_n}_{\max}\big(\rho^{\otimes n}_{X'}\,\|\,\Gamma^{\otimes n}_{X'}\big) + \log\frac{\bar{\varepsilon}'''^2_n}{2+\bar{\varepsilon}'''^2_n}\right]\\ \geqslant \lim_{n\to\infty}\frac{1}{n}\Big[D^{1-\bar{\varepsilon}''^2_n/6}_{\mathrm{H}}\big(\rho^{\otimes n}_{R_X}\,\|\,\Gamma^{\otimes n}_{R_X}\big) - D^{\bar{\varepsilon}'^2_n/2}_{\mathrm{H}}\big(\rho^{\otimes n}_{X'}\,\|\,\Gamma^{\otimes n}_{X'}\big)\\ - \log\left[\frac{1-\bar{\varepsilon}''^2_n/6}{\bar{\varepsilon}''^2_n/6}\right] + \log(1-\bar{\varepsilon}'_n) + \log(\bar{\varepsilon}'''^2_n/3)\Big]\\ \geqslant \lim_{n\to\infty}\frac{1}{n}\Big[D^{1-\bar{\varepsilon}''^2_n/6}_{\mathrm{H}}\big(\rho^{\otimes n}_{R_X}\,\|\,\Gamma^{\otimes n}_{R_X}\big) - D^{\bar{\varepsilon}'^2_n/2}_{\mathrm{H}}\big(\rho^{\otimes n}_{X'}\,\|\,\Gamma^{\otimes n}_{X'}\big)\\ + \log\big(\mathrm{poly}(\varepsilon_n)/\mathrm{poly}(\varepsilon_n)\big)\Big]\\ = D(\rho_{R_X}\,\|\,\Gamma_{R_X}) - D(\rho_{X'}\,\|\,\Gamma_{X'})\ .\end{aligned} \tag{199}$$

Equation (188) follows directly from (187), using the relations given by Proposition 26. ■

IV. ROBUSTNESS OF BATTERY STATES TO SMOOTHING

Because the battery system is a part of the physical implementation of the process, we may ask why it is not included in the definition of the smooth coherent relative entropy (46) in a way which would allow the physical implementation to fail to produce the appropriate output battery state with a small probability. Remarkably, there would have been no difference had we chosen to smooth the battery states as well. This follows from the following proposition, which asserts that optimization candidates which include smoothing on the battery states are in fact already included in the optimization in the definition above. This holds for the general battery states of the form $P_A\Gamma_A P_A/\operatorname{tr}(P_A\Gamma_A)$, for a projector P_A commuting with the Γ_A of the battery (see item (v) of Proposition 4).

Proposition 41 (Smoothing battery states)**.** *Let A, A' be quantum systems with corresponding Γ_A, $\Gamma_{A'}$. Let P_A, $P'_{A'}$ be projectors such that $[P_A,\Gamma_A]=0$ and $[P_{A'},\Gamma_{A'}]=0$, and let $\Phi_{XA\to X'A'}$ be a trace nonincreasing, completely positive map such that $\Phi_{XA\to X'A'}(\Gamma_X\otimes\Gamma_A)\leqslant\Gamma_{X'}\otimes\Gamma_{A'}$, and such that*

$$P\left[\Phi_{XA\to X'A}\left(\sigma_{XR}\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A}\right),\ \rho_{X'R}\otimes\frac{P'_{A'}\Gamma_{A'}P'_{A'}}{\operatorname{tr}P'_{A'}\Gamma_{A'}}\right]\leqslant\varepsilon\ , \tag{200}$$

Then there exists a trace-nonincreasing, completely positive map $\mathscr{T}_{X\to X'}$ such both the following conditions hold:

$$P(\mathscr{T}_{X\to X'}(\sigma_{XR}),\rho_{X'R})\leqslant\varepsilon\ ; \tag{201a}$$

$$\mathscr{T}_{X\to X'}(\Gamma_X)\leqslant\frac{\operatorname{tr}(P'_{A'}\Gamma_{A'})}{\operatorname{tr}(P_A\Gamma_A)}\Gamma_{X'}\ . \tag{201b}$$

Proof of Proposition 41. Define, for any ω_X,

$$\mathscr{T}_{X\to X'}(\omega_X) = \operatorname{tr}_{A'}\left[P'_{A'}\,\Phi_{XA\to X'A'}\left(\omega_X\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A}\right)\right]\ . \tag{202}$$

Then

$$\begin{aligned}\mathscr{T}_{X\to X'}(\sigma_{XR}) = \operatorname{tr}_{A'}\left[P'_{A'}\,\Phi_{XA\to X'A'}\left(\sigma_{XR}\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A}\right)\right]\\ = \operatorname{tr}_{A'}\left[P'_{A'}\,\tilde{\rho}_{A'X'R}\right]\ ,\end{aligned} \tag{203}$$

where $\tilde{\rho}_{A'X'R} := \Phi_{XA\to X'A'}(\sigma_{XR}\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A})$ satisfies $P(\tilde{\rho}_{A'X'R},\rho_{X'R}\otimes\frac{P'_{A'}\Gamma_{A'}P'_{A'}}{\operatorname{tr}P'_{A'}\Gamma_{A'}})\leqslant\varepsilon$ by assumption. Using the monotonicity of the purified distance [4] in particular under the trace-nonincreasing completely positive map $\operatorname{tr}\left[P'_{A'}(\cdot)\right]$, we have

$$P\big(\mathscr{T}_{X\to X'}(\sigma_{XR}),\rho_{X'R}\big)\leqslant\varepsilon\ . \tag{204}$$

We also have

$$\begin{aligned}\mathscr{T}_{X\to X'}(\Gamma_X) = \operatorname{tr}_{A'}\left[P'_{A'}\,\Phi_{XA\to X'A'}(\Gamma_X\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A})\right]\\ \leqslant\frac{1}{\operatorname{tr}P_A\Gamma_A}\cdot\operatorname{tr}_{A'}\left[P'_{A'}\,\Gamma_{X'}\otimes\Gamma_{A'}\right]\ ,\end{aligned} \tag{205}$$

using the fact that $P_A\Gamma_A P_A = \Gamma_A^{1/2}P_A\Gamma_A^{1/2}\leqslant\Gamma_A$ (because $[P_A,\Gamma_A]=0$) and also with the fact that $\Phi_{XA\to X'A'}$ is Γ-sub-preserving. Then

$$\mathscr{T}_{X\to X'}(\Gamma_X)\leqslant\frac{\operatorname{tr}P'_{A'}\Gamma_{A'}}{\operatorname{tr}P_A\Gamma_A}\Gamma_{X'}\ , \tag{206}$$

which completes the proof. ■

This means that the processes which also allow "fuzziness" on the battery states are *de facto* already included in the optimization defining the smooth coherent relative entropy (46). This is formulated explicitly in the following corollary.

Corollary 42. *Let $\rho_{X'R_X}$ be a subnormalized state, let $\Gamma_X, \Gamma_{X'} \geqslant 0$ and let $\varepsilon > 0$. Then*

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = \max_{A,A',P_A,P'_{A'},\Phi_{XA\to X'A'}} -\log \frac{\operatorname{tr} P'_{A'}\Gamma_{A'}}{\operatorname{tr} P_A \Gamma_A} \ , \tag{207}$$

where the optimization is performed over all systems A, A', all operators Γ_A, $\Gamma_{A'}$, and all projectors P_A, $P'_{A'}$ such that $[P_A, \Gamma_A] = 0$ and $[P'_{A'}, \Gamma_{A'}] = 0$, for which there is a trace nonincreasing, completely positive map $\Phi_{XA\to X'A'}$ satisfying $\Phi_{XA\to X'A'}(\Gamma_X \otimes \Gamma_A) \leqslant \Gamma_{X'} \otimes \Gamma_{A'}$ as well as

$$P\left[\Phi_{XA\to X'A'}\left(\sigma_{XR} \otimes \frac{P_A\Gamma_A P_A}{\operatorname{tr} P_A\Gamma_A}\right), \rho_{X'R} \otimes \frac{P'_{A'}\Gamma_{A'}P'_{A'}}{\operatorname{tr} P'_{A'}\Gamma_{A'}}\right] \leqslant \varepsilon \ . \tag{208}$$

Proof of Corollary 42. First, let A, A', P_A, $P'_{A'}$, Γ_A, $\Gamma_{A'}$ and $\Phi_{XA\to X'A'}$ satisfy the conditions of the maximization (207). Let $\mathscr{T}_{X\to X'}$ the mapping given by Proposition 41. Observe that $\big\|\Gamma_{X'}^{-1/2}\mathscr{T}_{X\to X'}(\Gamma_X)\Gamma_{X'}^{-1/2}\big\|_\infty \leqslant (\operatorname{tr} P'_{A'}\Gamma_{A'})/(\operatorname{tr} P_A\Gamma_A)$. Note also that $P(\mathscr{T}_{X\to X'}(\sigma_{XR}), \rho_{X'R_X}) \leqslant \varepsilon$ as guaranteed by our previous use of Proposition 41. Hence, $\mathscr{T}_{X\to X'}$ is a valid candidate in the optimization given by Proposition 12 for $\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$. Hence

$$\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant -\log \frac{\operatorname{tr} P'_{A'}\Gamma_{A'}}{\operatorname{tr} P_A\Gamma_A} \ . \tag{209}$$

To show that equality is achieved in (207), let $\mathscr{T}_{X\to X'}$ be a valid optimization candidate in (46) for $\hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$ which achieves the optimal value $y = \hat{D}^{\varepsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_R, \Gamma_{X'}) = -\log\big\|\Gamma_{X'}^{-1/2}\mathscr{T}_{X\to X'}(\Gamma_X)\Gamma_{X'}^{-1/2}\big\|_\infty$, with $P(\mathscr{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) \leqslant \varepsilon$. Then $\mathscr{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\Gamma_{X'}$, and this mapping satisfies the conditions of item (i) of Proposition 4. Let $A = A'$ be a qubit system with $P_A = |0\rangle\langle 0|_A$, $P'_{A'} = |1\rangle\langle 1|_{A'}$, and $\Gamma_A = \Gamma_{A'} = g_0\,|0\rangle\langle 0|_A + g_1\,|1\rangle\langle 1|_A$, with $g_0/g_1 = 2^y$. In virtue of item (iii) of Proposition 4, there exists a trace-nonincreasing, completely positive map $\Phi_{XA\to X'A'}$ such that $\Phi_{XA\to X'A'}(\Gamma_X \otimes \Gamma_A) \leqslant \Gamma_{X'} \otimes \Gamma_{A'}$ and which satisfies $\Phi_{XA\to X'A'}((\cdot)\otimes|0\rangle\langle 0|_A) = \mathscr{T}_{X\to X'}(\cdot)\otimes|1\rangle\langle 1|_{A'}$. Then

$$\Phi_{XA\to X'A'}(\sigma_{XR_X} \otimes |0\rangle\langle 0|_A) = \mathscr{T}_{X\to X'}(\sigma_{XR_X}) \otimes |1\rangle\langle 1|_{A'} \ , \tag{210}$$

and hence

$$P(\Phi_{XA\to X'A'}(\sigma_{XR_X} \otimes |0\rangle\langle 0|_A), \rho_{X'R_X} \otimes |1\rangle\langle 1|_{A'}) = P(\mathscr{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) \leqslant \varepsilon \ . \tag{211}$$

Hence, all the conditions of the maximization (207) are satisfied, and the achieved value is indeed $-\log[(\operatorname{tr} P'_{A'}\Gamma_{A'})/(\operatorname{tr} P_A\Gamma_A)] = -\log(g_1/g_0) = y$. ■

V. TECHNICAL UTILITIES

Lemma 43. *Let $A \geqslant 0$, $B \geqslant 0$ and let Π be the projector onto the support of A. Let $\mu > 0$. Define P as the projector onto the eigenspaces associated to nonnegative eigenvalues of the operator $(\mu A - B)$. Then there exists a constant c which is independent of μ such that*

$$\|\Pi - P\Pi P\|_\infty \leqslant \frac{c}{\mu} \ . \tag{212}$$

In particular,

$$\Pi \leqslant P + \frac{c}{\mu}\mathbb{1} \ . \tag{213}$$

Proof of Lemma 43. This lemma follows from a result of perturbation of matrix eigenspaces [32]. We'll consider the operators $A - \frac{1}{\mu}B$ and A. Let $Q = \mathbb{1} - P$ be the projector on the eigenspaces associated to the strictly negative eigenvalues of $A - \frac{1}{\mu}B$. Let $a_{\min} = \|A^{-1}\|_\infty^{-1}$ be the smallest nonzero eigenvalue of A. Recall that Π projects onto the eigenspaces of A associated to eigenvalues larger or equal to $a_{\min}$. We may now invoke [32, Theorem VII.3.1], which asserts that for any unitarily invariant norm $\|\cdot\|_\bullet$,

$$\|Q\Pi\|_\bullet \leqslant \frac{1}{\mu a_{\min}}\|QB\Pi\|_\bullet \leqslant \frac{1}{\mu a_{\min}}\|B\|_\bullet \ . \tag{214}$$

(The gap δ in [32, Theorem VII.3.1] is here the gap between 0 and $a_{\min}$.) In particular, we have $\|Q\Pi\|_\infty \leqslant (\mu a_{\min})^{-1}\,\|B\|_\infty$. We then have

$$\begin{aligned}\|\Pi - P\Pi P\|_\infty &\leqslant \|\Pi - P\Pi\|_\infty + \|P\Pi - P\Pi P\|_\infty \\ &\leqslant \|\Pi - P\Pi\|_\infty + \|P\|_\infty\|\Pi - \Pi P\|_\infty \\ &= 2\|\Pi - P\Pi\|_\infty = 2\|Q\Pi\|_\infty \leqslant \frac{c}{\mu} \ ,\end{aligned} \tag{215}$$

with $c = 2\,(a_{\min})^{-1}\,\|B\|_\infty$. This implies (213) because

$$\Pi - P\Pi P \leqslant \frac{c}{\mu}\mathbb{1} \qquad \Rightarrow \qquad \Pi \leqslant P\Pi P + \frac{c}{\mu}\mathbb{1} \leqslant P + \frac{c}{\mu}\mathbb{1} \ .$$ ■

Lemma 44. *Let ρ and σ be quantum states. The trace distance $D(\rho,\sigma)$ between ρ and σ can be written as the semidefinite program in terms of the variables $\Delta^{\pm} \geqslant 0$:*

$$\text{minimize}: \quad \frac{1}{2}\operatorname{tr}(\Delta^+ + \Delta^-) \tag{216a}$$

$$\text{subject to}: \quad \sigma = \rho + \Delta^+ - \Delta^- \ . \tag{216b}$$

Furthermore, $\operatorname{tr}\Delta^+ = \operatorname{tr}\Delta^- = D(\rho,\sigma)$ for the optimal solution. The dual to this program is an alternate expression of the same quantity, in terms of the Hermitian variable Z:

$$\text{maximize}: \quad \frac{1}{2}\operatorname{tr}(Z(\rho - \sigma)) \tag{217a}$$

$$\text{subject to}: \quad -\mathbb{1} \leqslant Z \leqslant \mathbb{1} \ . \tag{217b}$$

Proof of Lemma 44. Write $D(\rho,\sigma) = \frac{1}{2}\|\rho - \sigma\|_1$ and recall that for any Hermitian A, $\|A\|_1 = \operatorname{tr}|A|$. Choosing $\Delta_\pm \geqslant 0$ as the positive and negative parts of $\rho - \sigma$, i.e. such that $\rho - \sigma = \Delta_+ - \Delta_-$, yields feasible candidates for the primal problem and $\frac{1}{2}\operatorname{tr}(\Delta_+ + \Delta_-) = \frac{1}{2}\operatorname{tr}|\rho - \sigma| = D(\rho,\sigma)$. Now let $\Pi_\pm$ be the projectors onto the strictly positive and strictly negative parts of $\rho - \sigma$, respectively, and choose $Z = \Pi_+ - \Pi_-$. Observe that $\Pi_\pm(\rho - \sigma) = \pm\Delta_\pm$. Then $\frac{1}{2}\operatorname{tr}(Z(\rho - \sigma)) = \frac{1}{2}\operatorname{tr}(\Delta_+ + \Delta_-) = D(\rho,\sigma)$. We have exhibited primal and dual candidates achieving the value $D(\rho,\sigma)$, and hence this is the optimal solution of the semidefinite program. Furthermore (216b) implies that $\operatorname{tr}\Delta_+ = \operatorname{tr}\Delta_-$ and hence $\operatorname{tr}\Delta_+ = \operatorname{tr}\Delta_- = \frac{1}{2}\operatorname{tr}(\Delta_+ + \Delta_-) = D(\rho,\sigma)$. ■

Lemma 45 (Gentle measurement lemma for the purified distance)**.** *Let $\rho \geqslant 0$ with $\operatorname{tr}\rho = 1$. Let $\varepsilon \geqslant 0$. Let Π be a projector such that $\operatorname{tr}(\Pi\rho) \geqslant 1-\varepsilon$. Then*

$$P\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) \leqslant \frac{\sqrt{2\varepsilon}}{\sqrt{1-\varepsilon}} \ . \tag{218}$$

Proof of Lemma 45. Calculate

$$\begin{aligned} P^2\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) &= 1 - F^2\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) \\ &= \frac{1}{\operatorname{tr}(\Pi\rho)}\left[\operatorname{tr}(\Pi\rho) - F^2(\rho,\Pi\rho\Pi)\right] \\ &\leqslant \frac{1}{\operatorname{tr}(\Pi\rho)}\left[1 - F^2(\rho,\Pi\rho\Pi)\right] \\ &\leqslant \frac{1}{1-\varepsilon} P^2(\rho,\Pi\rho\Pi) \ , \end{aligned} \tag{219}$$

noting that the generalized fidelity is $F(\rho,\sigma) = \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1$ as long as one of the states is normalized, and hence for $a > 0$ we have $F^2(\rho, a\sigma) = aF^2(\rho,\sigma)$ if $\operatorname{tr}\rho = 1$. Now, applying [33, Lemma 7], we have

$$P\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) \leqslant \frac{\sqrt{2\varepsilon-\varepsilon^2}}{\sqrt{1-\varepsilon}} = \frac{\sqrt{\varepsilon\,(2-\varepsilon)}}{\sqrt{1-\varepsilon}} \leqslant \frac{\sqrt{2\varepsilon}}{\sqrt{1-\varepsilon}} \ . \qquad \blacksquare$$

Lemma 46 (Smoothing "part of" a state)**.** *Let ρ_{AB} be a bipartite normalized quantum state and let $\tilde\rho_A$ be a normalized quantum state such that $D(\tilde\rho_A,\rho_A) \leqslant \delta$. Then there exists a normalized quantum state $\hat\rho_{AB}$ such that $\operatorname{tr}_B \hat\rho_{AB} = \tilde\rho_A$, $\operatorname{tr}_A \hat\rho_{AB} = \rho_B$ and $P(\hat\rho_{AB},\rho_{AB}) \leqslant 2\sqrt{2\delta}$.*

Proof of Lemma 46. Because $\tilde\rho_A$ and ρ_A are δ-close in trace distance, by Lemma 44 there exists $\Delta_A^\pm \geqslant 0$ such that $\operatorname{tr}\Delta_A^- = \operatorname{tr}\Delta_A^+ = D(\tilde\rho_A,\rho_A) \leqslant \delta$ and

$$\tilde\rho_A = \rho_A + \Delta_A^+ - \Delta_A^- \ . \tag{220}$$

Let $A = \tilde\rho_A + \Delta_A^- \geqslant 0$ and let $M_A = \tilde\rho_A^{1/2} A^{-1/2}$. Observe that $M_A^\dagger M_A = A^{-1/2}\,\tilde\rho_A\,A^{-1/2} \leqslant \mathbb{1}$ since $\tilde\rho_A \leqslant A$. Now define the completely positive map

$$\mathscr{M}_{A\to A}(\cdot) = M_A\,(\cdot)\,M_A^\dagger + \operatorname{tr}\left[(\mathbb{1} - M_A^\dagger M_A)(\cdot)\right]\xi_A \ , \tag{221}$$

with $\xi_A := (M_A\,\Delta_A^+\,M_A^\dagger)/\operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger) \geqslant 0$ except if $\operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger) = 0$, in which case we set $\xi_A := \mathbb{1}_A/|A|$. In any case $\operatorname{tr}\xi_A = 1$ and $\operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger)\,\xi_A = M_A\,\Delta_A^+\,M_A^\dagger$. The mapping $\mathscr{M}_{A\to A}$ is trace preserving:

$$\mathscr{M}^\dagger(\mathbb{1}_A) = M_A^\dagger M_A + (\mathbb{1} - M_A^\dagger M_A)\operatorname{tr}\xi_A = \mathbb{1}_A \ . \tag{222}$$

We now show that $\mathscr{M}_{A\to A}(\rho_A) = \tilde\rho_A$. On one hand, using $A = \tilde\rho_A + \Delta_A^- = \rho_A + \Delta_A^+$, we have

$$M_A\,\rho_A\,M_A^\dagger = M_A\,A\,M_A^\dagger - M_A\,\Delta_A^+\,M_A^\dagger = \tilde\rho_A - M_A\,\Delta_A^+\,M_A^\dagger \ . \tag{223}$$

while noting that ρ_A lies within the support of A since $A = \rho_A + \Delta_A^+$. We deduce that $\operatorname{tr}(M_A\,\rho_A\,M_A^\dagger) = 1 - \operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger)$. On the other hand,

$$\begin{aligned} \operatorname{tr}\left[(\mathbb{1} - M_A^\dagger M_A)\,\rho_A\right]\xi_A &= \left(1 - \operatorname{tr}(M_A\,\rho_A\,M_A^\dagger)\right)\xi_A \\ &= \operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger)\,\xi_A \\ &= M_A\,\Delta_A^+\,M_A^\dagger \ , \end{aligned} \tag{224}$$

and hence, combining (223) with (224)

$$\mathscr{M}_{A\to A}(\rho_A) = M_A\,\rho_A\,M_A^\dagger + \operatorname{tr}\left[(\mathbb{1} - M_A^\dagger M_A)\,\rho_A\right]\xi_A = \tilde\rho_A \ . \tag{225}$$

Define now the state $\hat\rho_{AB}$ as

$$\hat\rho_{AB} = \mathscr{M}_{A\to A}\left[\rho_{AB}\right] \tag{226}$$

where the identity mapping is understood on system B. By properties of quantum channels the state on B is preserved, i.e. $\operatorname{tr}_A \hat\rho_{AB} = \rho_B$ (and in particular we have $\operatorname{tr}\hat\rho_{AB} = 1$), and we showed above that $\operatorname{tr}_B \hat\rho_{AB} = \tilde\rho_A$. It remains to see that $\hat\rho_{AB}$ and ρ_{AB} are close in purified distance. Let $|\rho\rangle_{ABC}$ be a purification of ρ_{AB}. Apply [23, Lemma A.4]—itself a reformulation of [34, Lemma 15]—with $\rho_{\text{Lem A.4}} = \rho_A$, $\sigma_{\text{Lem A.4}} = \tilde\rho_A$, $\Delta_{\text{Lem A.4}} = \Delta_A^-$, $G_{\text{Lem A.4}} = M_A$ and $|\psi_{\text{Lem A.4}}\rangle = |\rho\rangle_{ABC}$ to obtain

$$P\left(M_A\rho_{ABC}M_A^\dagger, \rho_{ABC}\right) \leqslant \sqrt{\left(2 - \operatorname{tr}\Delta_A^-\right)\operatorname{tr}\Delta_A^-} \leqslant \sqrt{2\delta} \ . \tag{227}$$

This distance can only decrease if we trace out the system C, and thus $P\left(M_A\rho_{AB}M_A^\dagger, \rho_{AB}\right) \leqslant \sqrt{2\delta}$. On the other hand, we have by definition

$$\hat\rho_{AB} = M_A\,\rho_{AB}\,M_A^\dagger + \Delta'_{AB} \ , \tag{228}$$

with $\Delta'_{AB} = \operatorname{tr}_A[(\mathbb{1}_A - M_A^\dagger M_A)\,\rho_{AB}] \otimes \xi_A \geqslant 0$. Calculate $\operatorname{tr}\Delta'_{AB} = \operatorname{tr}\left[(\mathbb{1}_A - M_A^\dagger M_A)\,\rho_A\right] = \operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger) \leqslant \operatorname{tr}\Delta_A^+ \leqslant \delta$, and hence $D\left(M_A\,\rho_{AB}\,M_A^\dagger, \hat\rho_{AB}\right) \leqslant \delta$. Finally, by triangle inequality and using $P(\rho,\rho') \leqslant \sqrt{2D(\rho,\rho')}$,

$$P(\hat\rho_{AB},\rho_{AB}) \leqslant P\left(\hat\rho_{AB}, M_A\,\rho_{AB}\,M_A^\dagger\right) + P\left(M_A\,\rho_{AB}\,M_A^\dagger, \rho_{AB}\right) \leqslant 2\sqrt{2\delta} \ . \qquad \blacksquare$$

Lemma 47. *Let σ_X, $\hat\sigma_X$ be two states. Consider another system $R \simeq X$. Then*

$$P(\sigma_X^{1/2}\,\Phi_{X:R}\,\sigma_X^{1/2}, \hat\sigma_X^{1/2}\,\Phi_{X:R}\,\hat\sigma_X^{1/2}) \leqslant 2\sqrt{D(\sigma_X,\hat\sigma_X)} \ . \tag{229}$$

Proof of Lemma 47. Let $\varepsilon = D(\sigma_X,\hat\sigma_X)$. Using the properties of the trace distance, let $\Delta_X^\pm \geqslant 0$ satisfy $\hat\sigma_X = \sigma_X + \Delta_X^+ - \Delta_X^-$ with $\operatorname{tr}\Delta^+ = \operatorname{tr}\Delta^- = \varepsilon$. Let $|\psi\rangle = (1+\varepsilon)^{-1/2}\,(\sigma_X + \Delta_X^+)^{1/2}|\Phi\rangle_{X:R} = (1+\varepsilon)^{-1/2}\,(\hat\sigma_X + \Delta_X^-)^{1/2}|\Phi\rangle_{X:R}$, noting that $\langle\psi\,|\,\psi\rangle = \operatorname{tr}(\sigma_X + \Delta_X^+)/(1-\varepsilon) = 1$. For any two pure states $|\phi\rangle, |\chi\rangle$ we know that $P(|\phi\rangle\langle\phi|, |\chi\rangle\langle\chi|) = \left(1 - |\langle\phi\,|\,\chi\rangle|^2\right)^{1/2}$. Our strategy for proving the claim is the following: We show that both

$$|\langle\psi_{XR}\,|\,\sigma_X^{1/2}\,|\,\Phi_{X:R}\rangle| \geqslant (1+\varepsilon)^{-1/2} \ ; \tag{230a}$$

$$|\langle\psi_{XR}\,|\,\hat\sigma_X^{1/2}\,|\,\Phi_{X:R}\rangle| \geqslant (1+\varepsilon)^{-1/2} \ , \tag{230b}$$

and the claim will then follow by triangle inequality for the purified distance: $P(\sigma_X^{1/2}\,\Phi_{X:R}\,\sigma_X^{1/2}, \hat\sigma_X^{1/2}\,\Phi_{X:R}\,\hat\sigma_X^{1/2}) \leqslant P(\sigma_X^{1/2}\,\Phi_{X:R}\,\sigma_X^{1/2}, |\psi\rangle\langle\psi|) + P(\hat\sigma_X^{1/2}\,\Phi_{X:R}\,\hat\sigma_X^{1/2}, |\psi\rangle\langle\psi|) \leqslant 2\sqrt{1 - 1/(1+\varepsilon)} \leqslant 2\sqrt{\varepsilon/(1+\varepsilon)} \leqslant 2\sqrt{\varepsilon}$. It remains to show the properties (230). We have $\langle\psi_{XR}\,|\,\sigma_X^{1/2}\,|\,\Phi_{X:R}\rangle = \langle\Phi_{X:R}\,|\,(\sigma_X^{1/2} + \Delta_X^+)^{1/2}\sigma_X^{1/2}\,|\,\Phi_{X:R}\rangle/\sqrt{1+\varepsilon} = \operatorname{tr}\left[(\sigma_X + \Delta_X^+)^{1/2}\sigma_X^{1/2}\right]/\sqrt{1+\varepsilon} \geqslant 1/\sqrt{1+\varepsilon}$, noting that $(\sigma_X + \Delta_X^+)^{1/2} \geqslant (\sigma_X)^{1/2}$, and hence $|\langle\psi_{XR}\,|\,\sigma_X^{1/2}\,|\,\Phi_{X:R}\rangle| \geqslant (1+\varepsilon)^{-1/2}$. Similarly, $\langle\psi_{XR}\,|\,\hat\sigma_X^{1/2}\,|\,\Phi_{X:R}\rangle = \langle\Phi_{X:R}\,|\,(\hat\sigma_X^{1/2} + \Delta_X^-)^{1/2}\hat\sigma_X^{1/2}\,|\,\Phi_{X:R}\rangle/\sqrt{1+\varepsilon} \geqslant (1+\varepsilon)^{-1/2}$. $\blacksquare$

Lemma 48 (Continuity of the relative entropy in its first argument)**.** *Let $\Gamma \geqslant 0$. Let ρ,σ lie within the support of Γ. Assume that $D(\rho,\sigma) \leqslant \varepsilon$. Then*

$$\begin{aligned} &|D(\rho\,\|\,\Gamma) - D(\sigma\,\|\,\Gamma)| \\ &\qquad \leqslant \varepsilon\log(\operatorname{rank}\Gamma - 1) + h(\varepsilon) + \varepsilon\,\|\log\Gamma\|_\infty \ , \end{aligned} \tag{231}$$

where $h(\varepsilon) = -\varepsilon\log\varepsilon - (1-\varepsilon)\log(1-\varepsilon)$ is the binary entropy.

Proof of Lemma 48. First, write

$$D(\rho\,\|\,\Gamma) = \operatorname{tr}\left[\rho\log\rho - \rho\log\Gamma\right] = -H(\rho) - \operatorname{tr}\left[\rho\log\Gamma\right] \ , \tag{232}$$

and so

$$|D(\rho\,\|\,\Gamma) - D(\sigma\,\|\,\Gamma)| \leqslant |H(\sigma) - H(\rho)| + |\mathrm{tr}[\sigma \log\Gamma] - \mathrm{tr}[\rho\log\Gamma]| \ . \quad (233)$$

Using the continuity bound of Audenaert [35], we have

$$|H(\rho) - H(\sigma)| \leqslant \varepsilon \log(\mathrm{rank}\,\Gamma - 1) + h(\varepsilon)\ , \quad (234)$$

where the states ρ and σ can be seen as living in a subspace of the full Hilbert space of dimension at most Γ (because they must both lie within the support of Γ), and where $h(\varepsilon) = -\varepsilon\ln\varepsilon - (1-\varepsilon)\ln(1-\varepsilon)$ is the binary entropy. On the other hand,

$$\begin{aligned}\mathrm{tr}\,\rho\log\Gamma - \mathrm{tr}\,\sigma\log\Gamma &= \|\log\Gamma\|_\infty\,\mathrm{tr}\left[(\rho-\sigma)\frac{\log\Gamma}{\|\log\Gamma\|_\infty}\right]\\ &\leqslant \|\log\Gamma\|_\infty\, D(\rho,\sigma)\ ,\end{aligned}$$

as $\log\Gamma/\|\log\Gamma\|_\infty$ is a valid candidate for Z in Lemma 44. Inverting the roles of ρ and σ in the equation above we finally obtain:

$$|\mathrm{tr}\,\rho\log\Gamma - \mathrm{tr}\,\sigma\log\Gamma| \leqslant \|\log\Gamma\|_\infty D(\rho,\sigma) \leqslant \|\log\Gamma\|_\infty\cdot\varepsilon\ . \qquad \blacksquare$$

[1] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, UK, 2006).

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[3] M. Tomamichel, R. Colbeck, and R. Renner, "Duality between smooth min- and max-entropies," IEEE Transactions on Information Theory **56**, 4674–4681 (2010), arXiv:0907.5238.

[4] M. Tomamichel, *A Framework for Non-Asymptotic Quantum Information Theory*, Ph.D. thesis, ETH Zurich (2012), arXiv:1203.2142.

[5] M. Tomamichel, *Quantum Information Processing with Finite Resources*, SpringerBriefs in Mathematical Physics, Vol. 5 (Springer International Publishing, Cham, 2016) arXiv:1504.00233.

[6] J. Watrous, "Semidefinite programs for completely bounded norms," Theory of Computing **5**, 217–238 (2009), arXiv:0901.4709.

[7] J. Watrous, "Lecture notes on the theory of quantum information," (2011).

[8] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," Proceedings of the National Academy of Sciences **112**, 201411728 (2015), arXiv:1305.5278.

[9] M. Horodecki and J. Oppenheim, "Fundamental limitations for quantum and nanoscale thermodynamics," Nature Communications **4**, 2059 (2013), arXiv:1111.3834.

[10] P. Skrzypczyk, A. J. Short, and S. Popescu, "Work extraction and thermodynamics for individual quantum systems." Nature communications **5**, 4185 (2014), arXiv:1307.1558.

[11] Á. M. Alhambra, L. Masanes, J. Oppenheim, and C. Perry, "Fluctuating work: From quantum thermodynamical identities to a second law equality," Physical Review X **6**, 041017 (2016), arXiv:1601.05799.

[12] R. Renner, *Security of Quantum Key Distribution*, Ph.D. thesis, ETH Zürich (2005), arXiv:quant-ph/0512258.

[13] P. Faist, J. Oppenheim, and R. Renner, "Gibbs-preserving maps outperform thermal operations in the quantum regime," New Journal of Physics **17**, 043003 (2015), arXiv:1406.3618.

[14] N. Datta, "Min- and max-relative entropies and a new entanglement monotone," IEEE Transactions on Information Theory **55**, 2816–2826 (2009), arXiv:0803.2770.

[15] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, "Inadequacy of von Neumann entropy for characterizing extractable work," New Journal of Physics **13**, 53015 (2011), arXiv:0908.0424.

[16] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, "The thermodynamic meaning of negative entropy," Nature **474**, 61–63 (2011), arXiv:1009.1630.

[17] J. Åberg, "Truly work-like work extraction via a single-shot analysis," Nature Communications **4**, 1925 (2013), arXiv:1110.6121.

[18] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, "The minimal work cost of information processing," Nature Communications **6**, 7669 (2015), arXiv:1211.1037.

[19] F. Buscemi and N. Datta, "The quantum capacity of channels with arbitrarily correlated noise," IEEE Transactions on Information Theory **56**, 1447–1460 (2010), arXiv:0902.0158.

[20] F. G. S. L. Brandão and N. Datta, "One-shot rates for entanglement manipulation under non-entangling maps," IEEE Transactions on Information Theory **57**, 1754–1760 (2011), arXiv:0905.2673.

[21] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," Physical Review Letters **108**, 200501 (2012), arXiv:1007.5456.

[22] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," IEEE Transactions on Information Theory **59**, 7693–7710 (2013), arXiv:1208.1478.

[23] F. Dupuis, L. Kraemer, Ph. Faist, J. M. Renes, and R. Renner, "Generalized entropies," in *XVIIth International Congress on Mathematical Physics* (2013) pp. 134–153, arXiv:1211.3141.

[24] W. Matthews and S. Wehner, "Finite blocklength converse bounds for quantum channels," IEEE Transactions on Information Theory **60**, 7317–7329 (2014), arXiv:1210.4722.

[25] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner, "Chain rules for smooth min- and max-entropies," IEEE Transactions on Information Theory **59**, 2603–2612 (2013), arXiv:1205.5231.

[26] C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, "Moderate deviation analysis for classical communication over quantum channels," Communications in Mathematical Physics **355**, 1283–1315 (2017), arXiv:1701.03114.

[27] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," (2017), arXiv:1701.03195.

[28] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," Communications in Mathematical Physics **143**, 99–114 (1991).

[29] H. Nagaoka and T. Ogawa, "Strong converse and stein's lemma

in quantum hypothesis testing," IEEE Transactions on Information Theory **46**, 2428–2433 (2000), arXiv:quant-ph/9906090.

[30] I. Bjelakovic and R. Siegmund-Schultze, "Quantum stein's lemma revisited, inequalities for quantum entropies, and a concavity theorem of lieb," (2003), arXiv:quant-ph/0307170.

[31] M. Mosonyi and T. Ogawa, "Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies," Communications in Mathematical Physics **334**, 1617–1648 (2014), arXiv:1309.3228.

[32] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics (Springer, New York, 1997).

[33] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, "The uncertainty principle in the presence of quantum memory," Nature Physics **6**, 659–662 (2010), arXiv:0909.0950.

[34] M. Tomamichel, R. Colbeck, and R. Renner, "A fully quantum asymptotic equipartition property," IEEE Transactions on Information Theory **55**, 5840–5847 (2009), arXiv:0811.1221.

[35] K. M. R. Audenaert, "A sharp continuity estimate for the von Neumann entropy," Journal of Physics A: Mathematical and Theoretical **40**, 8127–8136 (2007), arXiv:quant-ph/0610146.