Quantum thermodynamics of correlated-catalytic state conversion at small-scale

Naoto Shiraishi[1] and Takahiro Sagawa[2]
[1]*Department of Physics, Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo 171-8588, Japan and*
[2]*Department of Applied Physics and Quantum-Phase Electronics Center (QPEC), The University of Tokyo, 7-3-1 Hongo, Byunkyo-ku, Tokyo 113-8656, Japan*
(Dated: October 22, 2020)

The class of possible thermodynamic conversions can be extended by introducing an auxiliary system called catalyst, which assists state conversion while remaining its own state unchanged. We reveal a complete characterization of catalytic state conversion in quantum and single-shot thermodynamics by allowing an infinitesimal correlation between the system and the catalyst. Specifically, we prove that a single thermodynamic potential, which provides the necessary and sufficient condition for the correlated-catalytic state conversion, is given by the standard nonequilibrium free energy defined with the Kullback-Leibler divergence. This resolves the conjecture raised by Wilming, Gallego, and Eisert [Entropy 19, 241 (2017)] and by Lostaglio and Müller [Phys. Rev. Lett. 123, 020403 (2019)] in positive. Moreover, we show that, with the aid of the work storage, any quantum state can be converted into another one by paying the work cost equal to the difference of the nonequilibrium free energy. Our result would serve as a step towards establishing resource theories of catalytic state conversion in the fully quantum regime.

Introduction.— Extension of thermodynamics to small-scale quantum systems has attracted attention in various research fields. Variety of the second laws employing the Rényi entropies and divergences [1–4] or majorization [5–8] naturally arise in the small-scale, which is contrastive to conventional thermodynamics where only a single thermodynamic potential such as the equilibrium free energy characterizes state convertibility [9]. Recent studies pushing toward this direction are developed in terms of resource theories [6, 7, 10]. The resource theory of athermality [4, 11–13] paves the way for establishing the information-theoretic foundation of thermodynamics.

In resource theories, an auxiliary system called *catalyst* plays a key role, which assists the state conversion while the catalyst itself does not change. In the context of thermodynamics, the concept of catalyst is motivated by the fact that the conventional second law identifies the possibility of state conversion without leaving any change in the environment. To formulate the catalytic state conversion (also called *trumping*) in general, we suppose the composite system of the system and the catalyst, and consider a state conversion $\rho \otimes c \to \sigma \otimes c$, where ρ, σ are states of the system and c is a state of the catalyst. On one hand, if we require that the return of the catalyst is exact, an infinite family of Rényi entropies or divergences characterizes possible catalytic state conversion [1–4]. On the other hand, if we allow a small finite error in the final state of the catalyst (i.e., the final state of the catalyst can be slightly different from the initial one), any state conversion is possible and the resource theory becomes trivial, which is called *embezzling* [4, 14, 15]. Here our focus lies in their intermediate regime, where another nontrivial characterization of state convertibility emerges.

Specifically, we consider the situation that the catalyst returns to its initial state exactly but with a negligibly small correlation between the system and the catalyst. As observed in Refs [16, 17], correlations are resources of thermodynamic state conversions. Along this idea, Wilming, Gallego, and Eisert [18] conjectured that the nonequilibrium free energy defined by the quantum Kullback-Leibler (KL) divergence gives the unique criterion of correlated-catalytic state conversion via a Gibbs-preserving map with a negligibly small correlation. In the classical case, this conjecture has been solved in positive by Müller [19] and generalized by Rethinasamy and Wilde [20]. However, these results cannot apply to the quantum case, because unlike the classical case known criteria for quantum relative majorization are highly complicated [21–23]. Therefore, the original conjecture raised in Ref [18] (also raised in Ref. [24] in a rigorous manner) for the quantum cases has still been left as a highly-nontrivial open problem, despite its physical significance highlighted by the recent development of experimental techniques manipulating quantum systems [25–31].

In this Letter, we solve this problem for the quantum case [18, 24] in the affirmative: We prove that the KL divergence indeed characterizes quantum correlated-catalytic state conversion in a necessary and sufficient manner. That is, the correlated-catalytic state conversion between any two given quantum states by a Gibbs-preserving map is possible *if and only if* the nonequilibrium free energy defined by the KL divergence does not increase. We further prove that even if the final free energy is larger than the initial one, we can still convert the initial state to the final one by adding a two-level work storage and paying the work cost equal to or greater than the free energy difference.

Our result implies that the conventional form of the second law given by the KL divergence is restored even in the quantum regime, if the catalyst is allowed to correlate with the system. We thus regard our result as a conceptual foundation that bridges conventional ther-

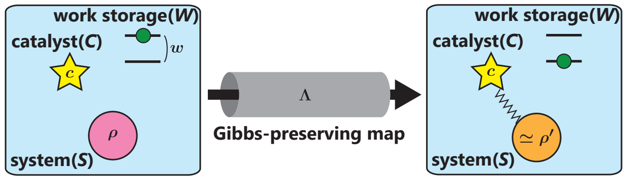

FIG. 1. Schematic of our setup. We convert the system S from ρ to ρ' with the aid of the catalyst C and the work storage W. The catalyst C returns to its original state while it can correlate with the system. The work storage W changes its state with energy difference $w \geq F(\rho') - F(\rho)$ with probability arbitrarily close to unity.

modynamics and the resource theory of thermodynamics. Furthermore, our result would have implications not only to quantum thermodynamics but also to various quantum resource theories [10], as our proof techniques are applicable to quantum state conversion that are not necessarily thermodynamic.

Setup and the main claim.— Consider a finite-dimensional quantum system with Hamiltonian H. We investigate state conversion through a particular class of the completely-positive and trace-preserving (CPTP) maps, called *Gibbs-preserving maps* Λ, which keep the Gibbs state invariant: $\Lambda(\rho_{\mathrm{Gibbs}}) = \rho_{\mathrm{Gibbs}}$. Here, $\rho_{\mathrm{Gibbs}} := e^{-\beta H}/Z$ is the Gibbs state with the inverse temperature β of the environment. We set the Boltzmann constant to unity. In terms of the resource theory of athermality, the Gibbs state is a free state (with zero athermality), and Gibbs-preserving maps do not generate any non-free state (with nonzero athermality) from a free state.

We employ an external system called catalyst denoted by C, which assists state conversion of the system S while the state of C itself does not change (see also Fig. 1). As in Refs. [4, 13, 16, 17, 19, 20, 32–34], we allow a negligibly small error on the final state of S, while the marginal state of C exactly goes back to the initial state. The most crucial assumption is to allow a negligibly small correlation between S and C in the final state. This assumption is motivated by the fact that negligibly small correlations are always allowed between the system and the environment in conventional thermodynamics. In terms of resource theories, a catalyst for a system can be reused as a catalyst for other systems even when a correlation with the first system remains.

We define the nonquilibrium free energy as $F(\rho) := S_1(\rho||\rho_{\mathrm{Gibbs}})$, where $S_1(\rho||\rho_{\mathrm{Gibbs}}) := \mathrm{Tr}[\rho \ln \rho] - \mathrm{Tr}[\rho \ln \rho_{\mathrm{Gibbs}}]$ is the KL divergence [35]. We now state our first main theorem:

Theorem 1– Consider two quantum states of S; ρ and ρ'. Then, $F(\rho) \geq F(\rho')$ is satisfied if and only if there exist a catalyst C and its state c, and a Gibbs-preserving map Λ satisfying $\Lambda(\rho \otimes c) = \tau$ such that (i) $\mathrm{Tr}_S[\tau] = c$, (ii) $\mathrm{Tr}_C[\tau]$ is arbitrarily close to ρ', (iii) the correlation between S and C in the final state is arbitrarily small.

The fully rigorous statement of the above theorem and its proof are presented in Supplemental Material [36]. Here, we only remark that the closeness between states is quantified by the trace distance $d_1(\rho', \rho'') := \frac{1}{2}\mathrm{Tr}[|\rho' - \rho''|]$ and the amount of the correlation is quantified by the mutual information $I_{\mathrm{SC}}[\tau] := S_1(\tau||\rho'' \otimes c)$, where $\rho'' = \mathrm{Tr}_C[\tau]$ is the reduced state of τ on S. This theorem manifests the fact that the free energy $F(\rho)$ serves as the single monotone of quantum thermodynamics at the small-scale if we allow a negligibly small correlation between the system and the catalyst. Note that Theorem 1 is proved as a corollary of a more general theorem, Theorem 3, presented later.

In the case of $F(\rho) < F(\rho')$, Theorem 1 implies that we cannot convert ρ to ρ' through any Gibbs-preserving map. However, even in this case we can convert ρ to ρ' with the aid of the work storage W (see Fig. 1). The work storage is a two-level system which compensates the energy change in S by investing the work cost. The initial state of W is an energy eigenstate $|a\rangle$ with energy E_a, and the final state is arbitrarily close to another energy eigenstate $|b\rangle$ with energy E_b. The work value is thus almost deterministic, which is an approximate version of the single-shot scenario [8, 19, 37].

By applying Theorem 1 to the composite system SW, we find that $\rho \otimes |a\rangle\langle a|$ can be converted to a state close to $\rho' \otimes |b\rangle\langle b|$ with a catalyst if we allow a correlation between SW and C. Further to that, we can prove a much stronger statement: the desired state conversion is possible even when there is no correlation between W and the remaining part SC:

Theorem 2– Consider two quantum states ρ and ρ' of the system S with $F(\rho) - F(\rho') < 0$. Then, $F(\rho) - F(\rho') \geq w$ is satisfied if and only if there exist a catalyst C and its state c, a work storage W with $E_b - E_a \geq w$, and a Gibbs-preserving map Λ satisfying $\Lambda(\rho \otimes c \otimes |a\rangle\langle a|) = \tau \otimes \omega$ with τ and ω being states of SC and W, such that (i) $\mathrm{Tr}_S[\tau] = c$, (ii) $\mathrm{Tr}_C[\tau]$ is arbitrarily close to ρ', (iii) ω is arbitrarily close to $|b\rangle\langle b|$, (iv) the correlation between S and C in τ is arbitrarily small.

This theorem reveals the minimum work cost when C correlates only with S as depicted in Fig. 1, and in the conventional terminology, represents *the principle of maximum work* (i.e., the maximum work is given by the free energy difference) [38–40]. The foregoing two theorems together provide the second law of quantum thermodynamics in the small-scale, yet in the apparently same form as conventional macroscopic thermodynamics.

We note that Theorem 2 only applies the case of the work investment ($w < 0$), and does not cover the case of the work extraction ($w > 0$). Although the work investment and the work extraction are almost symmetric in conventional thermodynamics, the treatment of the work extraction is often a much harder task than that of the work investment in small-scale thermodynamics [19, 34]. We will, however, discuss a sufficient condition for the case of work extraction in Supplemental Material (Lemma 3) [36].

Outline of the proof.— We here summarize the outline of the proof. The detailed idea is demonstrated along with a simple example soon later, and the full proofs are presented in Supplemental Material [36]. We mainly treat Theorem 1 and briefly comment on the case of Theorem 2. Our proof consists of three steps: deriving a sufficient condition for quantum state conversion by a Gibbs-preserving map (Step 1), applying the quantum Stein's lemma (Step 2), and the reduction from asymptotic state conversion to catalytic state conversion (Step 3).

In Step 1, we provide a sufficient condition to convert a quantum state σ to another state σ' via a Gibbs-preserving map by explicitly constructing the desired map. We first perform a binary quantum measurement to determine whether the state is σ or the Gibbs state σ_{Gibbs}, and then prepare two quantum states depending on the measurement outcome. In the case of Theorem 2 (with work storage), we consider not binary but ternary measurements, which requires more careful treatment. The derived sufficient condition employs two kinds of divergences; the quantum hypothesis testing divergence [41] and the quantum Rényi divergence, whose rigorous definitions are provided in the explanation of the toy example below. These techniques are inspired by Refs. [8, 33, 34, 42].

In Step 2, we apply the quantum Stein's lemma, which claims the convergence of the quantum hypothesis testing divergence rate to the KL divergence rate in the limit of infinitely many copies of given quantum systems [43, 44]. The reason why quantum hypothesis testing appears in quantum thermodynamics is that the ε-smoothed Rényi-∞ divergence (introduced later) is bounded from both above and below by two quantum hypothesis testing divergences, and hence it also converges to the KL divergence [45, 46].

In Step 3, we reduce the result on asymptotic state conversion (with multiple copies of states) of Step 1 and 2 to catalytic state conversion. Although this type of reduction has been discussed in some literature [1, 4, 47], we need some modification on the existing technique in order to keep the catalyst at the same state, because our asymptotic state conversion accompanies errors. Combining these three steps, we arrive at the desired result.

Toy example of Theorem 1.— We demonstrate a toy example of Theorem 1, where S is a two-level system spanned by $\{|0\rangle, |1\rangle\}$. The following argument also serves as a proof of Theorem 1 in this specific case. We will construct two states that are not convertible from one to the other without catalyst, but are convertible with a correlated-catalyst. For this purpose,

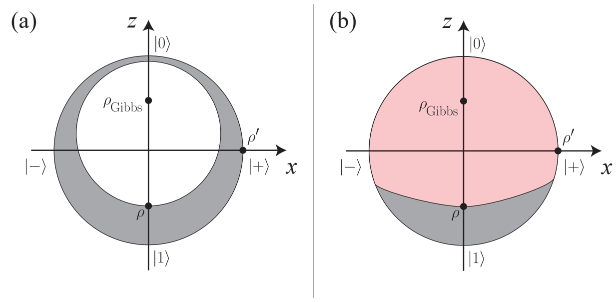

FIG. 2. (a) Schematic of the criterion $S_\infty(\rho'||\rho_{\mathrm{Gibbs}}) < S_\infty(\rho||\rho_{\mathrm{Gibbs}})$ in the x-z plane of the Bloch sphere, where we draw the states inconvertible from ρ in gray. For example, any state conversion $\rho \to \rho'$ by a Gibbs-preserving map is impossible. (b) Schematic of the criterion $S_1(\rho'||\rho_{\mathrm{Gibbs}}) < S_1(\rho||\rho_{\mathrm{Gibbs}})$ in the x-z plane of the Bloch sphere, where we draw the convertible and inconvertible states with a correlated catalyst from ρ in red and gray, respectively. In particular, there exists a Gibbs-preserving map with correlated catalyst converting $\rho \to \rho'$.

we set $\rho = \frac{3}{200}|0\rangle\langle 0| + \frac{197}{200}|1\rangle\langle 1|$, $\rho' = |+\rangle\langle +|$ with $|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $\beta = 1$, $E_0 = 0$ and $E_1 = \ln 4$. The Gibbs state is given by $\rho_{\mathrm{Gibbs}} = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|$. We set the upper bound of the error and the correlation strength as $\varepsilon = 0.01$ and $\delta = 0.06$, respectively. This state conversion is fully quantum because ρ' is not diagonal in the energy eigenbasis. We remark that any Gibbs-preserving map without catalyst cannot convert ρ to ρ'. To see this, we introduce the Rényi-∞ divergence $S_\infty(\sigma||\kappa) := \ln(\min[\lambda : \sigma \le \lambda\kappa])$ and its ε-smoothing $S^\varepsilon_\infty(\sigma||\kappa) := \min_{d(\sigma',\sigma)\le\varepsilon} S_\infty(\sigma||\kappa)$ with $\varepsilon > 0$. The Rényi divergence satisfies the monotonicity under CPTP maps, and hence $S^\varepsilon_\infty(\rho'||\rho_{\mathrm{Gibbs}}) \le S_\infty(\rho||\rho_{\mathrm{Gibbs}})$ is a necessary (but not sufficient) condition for state conversion without catalyst. However, we can show $S^\varepsilon_\infty(\rho'||\rho_{\mathrm{Gibbs}}) > S_\infty(\rho||\rho_{\mathrm{Gibbs}})$ in the above parameter setting (see Fig. 2(a)).

We treat Step 1 and Step 2 in parallel. Consider a composite system of 8 copies of the two-level system: $\{|0\rangle, |1\rangle\}^{\otimes 8}$. We construct a CPTP map which converts $\rho^{\otimes 8}$ to a state Ξ satisfying $d_1(\Xi, \rho'^{\otimes 8}) < \varepsilon$ while keeping ${\rho_{\mathrm{Gibbs}}}^{\otimes 8}$ unchanged. We introduce the projection operator Q onto the subspace of $\{|0\rangle, |1\rangle\}^{\otimes 8}$ spanned by a subset of the computational basis that contains at most one $|0\rangle$. We perform the binary measurement with $\{Q, 1-Q\}$ in order to distinguish $\rho^{\otimes 8}$ and $\rho^{\otimes 8}_{\mathrm{Gibbs}}$. By this measurement, $\rho^{\otimes 8}$ outputs Q with probability $0.994\cdots > 1-\varepsilon$, and $\rho^{\otimes 8}_{\mathrm{Gibbs}}$ outputs $1-Q$ with probability $1 - \frac{25}{4^8}$. Their differences from 1 (i.e., $0.005\cdots$ and $\frac{25}{4^8}$) correspond to the error of the first and the second kind, respectively. We then prepare quantum states depending on the measurement outcome. If the outcome is Q, we prepare the state $|+\rangle\langle +|$, and if the outcome is $1-Q$, we prepare the state ζ expressed as

$$\zeta = \frac{1}{1-\frac{25}{4^8}}\left(\rho_{\mathrm{Gibbs}}^{\otimes 8} - \frac{25}{4^8}\left|+\right\rangle\left\langle+\right|^{\otimes 8}\right), \tag{1}$$

which is positive-semidefinite because $\rho_{\mathrm{Gibbs}}^{\otimes 8} - \lambda\left|+\right\rangle\left\langle+\right|^{\otimes 8} \geq 0$ for $\lambda \leq \left(\frac{3}{8}\right)^8$ [36] and $\frac{25}{4^8} < \left(\frac{3}{8}\right)^8$. This measurement-and-preparation procedure indeed converts ρ_{Gibbs} to ρ_{Gibbs} by construction and converts ρ to $\Xi := (0.994\cdots)\left|+\right\rangle\left\langle+\right|^{\otimes 8} + (0.005\cdots)\zeta$, which satisfies $d_1(\Xi, \left|+\right\rangle\left\langle+\right|^{\otimes 8}) < \varepsilon$. We denote this CPTP map by Λ.

We next move to Step 3. We identify the system S to S_1 and the catalyst C to $S_2\otimes\cdots\otimes S_8\otimes A$, where A is an auxiliary system spanned by a basis $\{\left|1\right\rangle, \left|2\right\rangle, \ldots, \left|8\right\rangle\}$. The Hamiltonian of A is set to be trivial (i.e., all the states in A take the same energy). Using Ξ on $S_1\otimes\cdots\otimes S_8$ introduced above, we define Ξ_i $(i = 1, \ldots, 8)$ as the reduced state of Ξ on $S_1\otimes\cdots\otimes S_i$. We set $\Xi_0 := 1 \in \mathbb{R}$ (i.e., the trivial state) for convenience. Using these states, we define the state of C as

$$c := \frac{1}{8}\sum_{k=1}^{8} \rho^{\otimes k-1}\otimes\Xi_{8-k}\otimes\left|k\right\rangle\left\langle k\right|, \tag{2}$$

where $\rho^{\otimes k-1}$ is the state of $S_2\otimes\cdots\otimes S_k$, and Ξ_{8-k} is now the state of $S_{k+1}\otimes\cdots\otimes S_8$. The initial state of the composite system is $\rho\otimes c = \frac{1}{8}\sum_{k=1}^{8}\rho^{\otimes k}\otimes\Xi_{8-k}\otimes\left|k\right\rangle\left\langle k\right|$ (see Fig. 3(a)). We now construct the desired CPTP map as follows: If the auxiliary system A is $\left|8\right\rangle\left\langle 8\right|$, then we apply Λ to $S_1\otimes\cdots\otimes S_8$, and leave it unchanged otherwise. Through this process, $\frac{1}{8}\sum_{k=1}^{8}\rho^{\otimes k}\otimes\Xi_{8-k}\otimes\left|k\right\rangle\left\langle k\right|$ is converted into $\tau' = \frac{1}{8}\sum_{k=1}^{8}\rho^{\otimes k-1}\otimes\Xi_{9-k}\otimes\left|k\right\rangle\left\langle k\right|$ (see Fig. 3.(b)). Then, we shift the auxiliary system A as $\left|8\right\rangle \to \left|1\right\rangle$ and $\left|n\right\rangle \to \left|n+1\right\rangle$. Remarkably, the partial trace of τ' with respect to S_8 recovers the initial state of the catalyst c. In addition, by defining ξ_k as the reduced state of Ξ on S_k, the reduced state of τ' on S_8 is expressed as $\frac{1}{8}\sum_{k=1}^{8}\xi_k$, which is ε-close to the desired state $\rho' = \left|+\right\rangle\left\langle+\right|$ because $d_1(\Xi, \left|+\right\rangle\left\langle+\right|^{\otimes 8}) < \varepsilon$. Moreover, since Λ is a Gibbs-preserving map, the constructed CPTP map is also Gibbs-preserving. Thus, swapping the two-level systems as $S_n \to S_{n+1}$ and $S_8 \to S_1$ after the above CPTP map, we arrive at the desired Gibbs-preserving map: $\rho\otimes c$ is converted into τ with $\mathrm{Tr}_S[\tau] = c$, $d_1(\mathrm{Tr}_C[\tau], \rho') < \varepsilon$, and $I_{\mathrm{SC}}(\tau) < \delta$. Here, since S is a two-level system, $d_1(\tau, \rho'\otimes c) < \varepsilon$ implies $I_{\mathrm{SC}}(\tau) < -\varepsilon\ln\varepsilon - (1-\varepsilon)\ln(1-\varepsilon) = 0.056\cdots < 0.06$.

Discussion.— The obtained results solve in positive the conjecture raised by Wilming, Gallego, and Eisert [18] and by Lostaglio and Müller [24]. Note that Müller [19] proved this conjecture for classical systems by showing an elaborate way to explicitly construct a catalyst, which is completely different from our approach. Thus, our proof restricted to the classical regime serves as an alternative proof of Müller's.

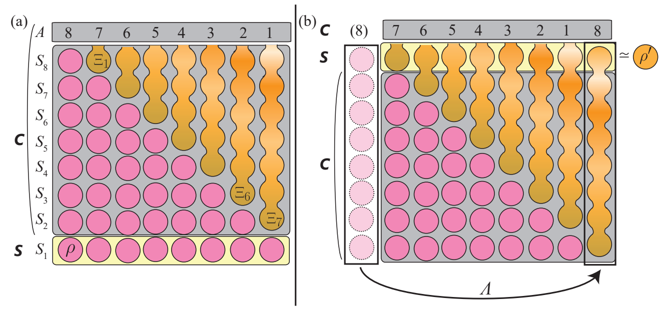

FIG. 3. Schematic of Step 3 of the proof. (a) The initial state of the composite SC. The vertical direction represents different systems $S_1, \ldots, S_8$, and the horizontal direction means their classical mixture. (b) Schematic of how the CPTP map Λ gives the desired catalytic state conversion.

In this work, we have considered Gibbs-preserving maps as thermodynamic processes instead of thermal operations. Gibbs-preserving maps and thermal operations are equivalent in the classical regime [12, 48], while the set of Gibbs-preserving maps is strictly larger than the set of thermal operations in the quantum regime [49]. The original conjecture of Ref. [18] is about Gibbs-preserving maps, while a stronger conjecture with quantum thermal operations was raised in Ref. [19]. However, Refs. [24, 50] solved the latter stronger conjecture in negative by proving that coherence cannot be broadcast even under the presence of correlations with catalyst. Coherence is known to prevent the reduction from Gibbs-preserving maps to thermal operations. Therefore, in the present work, we focus on Gibbs-preserving maps that could still give a positive answer to the conjectures [18, 24].

We also note that Müller [19] performs trivialization of the catalyst Hamiltonian, but we did not. Here, we say a catalyst trivialized when the Hamiltonian of the catalyst is trivial. Hardness of trivialization in the fully quantum regime comes from the fact that merging and splitting of states are irreversible due to decoherence in quantum systems. Due to this difficulty, we did not trivialize the catalyst in the present work, which remains as a future work.

Besides the correlated classical cases [16, 17, 19, 51], in some setups of quantum thermodynamics and resource theories a single thermodynamic potential with the KL divergence also appears [4, 13, 32–34, 52–55]. However, those previous results are different from our result in some important aspects: Some of them allow small changes in the catalyst [4] or other external systems [52–55] (instead they consider more restricted classes of operations compared to Gibbs-preserving maps), and some others consider asymptotic (macroscopic) conversion [13, 32–34]. It is yet interesting to see that the same thermodynamic potential appears in these various setups.

Meanwhile, we can further extend Theorem 1 to general CPTP maps by employing a similar idea of the proof. This is about the quantum counterpart of catalytic d-majorization (also called relative majorization) [6, 8]:

Theorem 3– Consider four quantum states ρ, ρ', η, and η' of the system S. Then, $S(\rho||\eta) \geq S(\rho'||\eta')$ holds if and only if there exists a catalyst C and its two states c, d, and a CPTP map which converts $\eta \otimes d$ to $\eta' \otimes d$ and $\rho \otimes c$ to τ satisfying (i) $\mathrm{Tr}_S[\tau] = c$, (ii) $\mathrm{Tr}_C[\tau]$ is arbitrarily close to ρ', (iii) the correlation between S and C in τ is arbitrarily small.

We present the proof of this theorem in Supplemental Material [36]. This theorem almost solves the conjecture of Rethinasamy and Wilde [20], who proved the classical case. The only difference between our result and the conjecture of them is that we did not trivialize catalyst c.

Finally, our approach sheds new light on other single-shot resource theories with catalyst. Brandao and Gour [56] have established that various resource theories concerning asymptotic state conversion with a small error are characterized by the KL divergence. Their result applies to the resource theories of entanglement, coherence, contextuality, and stabilizer computation. By employing our technique (in particular, Step 3 of the proof), we see that the KL divergence also serves as a single monotone in these single-shot resource theories in the presence of a catalyst with a negligibly small correlation.

Acknowledgement.— NS is supported by JSPS KAKENHI Grants-in-Aid for Early-Career Scientists Grant Number JP19K14615. TS is supported by JSPS KAKENHI Grant Numbers JP16H02211 and JP19H05796. TS is also supported by institute of AI and Beyond of the University of Tokyo.

[1] G. Aubrun and I. Nechita, *Catalytic Majorization and ℓ^p Norms*. Comm. Math. Phys. **278**, 133 (2008).

[2] M. Klimesh, *Inequalities that Collectively Completely Characterize the Catalytic Majorization Relation*, arXiv:0709.3680 (2007).

[3] S. Turgut, *Catalytic transformations for bipartite pure states*. J. Phys. A Math. Theor. **40** 12185 (2007).

[4] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, *The second laws of quantum thermodynamics*. Proc. Nat. Acad. Sci. **112**, 3275 (2015).

[5] A. W. Marshall, I. OlkinBarry, and C. Arnold, *Inequalities: theory of majorization and its applications*. Springer (1979).

[6] G. Gour, M. P. Müller, V. Narasimhachar. R. W. Spekkens, N. Y. Halpern, *The resource theory of informational nonequilibrium in thermodynamics*. Phys. Rep. **583**, 1 (2015).

[7] M. Lostaglio, *An introductory review of the resource theory approach to thermodynamics*. Rep. Prog. Phys. **82** 114001 (2019).

[8] T, Sagawa, *Entropy, Divergence, and Majorization in Classical and Quantum Thermodynamics*. arXiv:2007.09974 (2020).

[9] E. H.Lieb and J. Yngvason, *The physics and mathematics of the second law of thermodynamics*. Phys. Rep. **310**, 1 (1999).

[10] E. Chitambar and G. Gour, *Quantum resource theories*. Rev. Mod. Phys. 91, 025001 (2019).

[11] D. Janzing, P. Wocjan, R. Zeier, R. Geiss and Th. Beth, *Thermodynamic Cost of Reliability and Low Temperatures: Tightening Landauer's Principle and the Second Law*. Int. J. Theor. Phys. **39**, 2717 (2000).

[12] M. Horodecki and J. Oppenheim, *Fundamental limitations for quantum and nanoscale thermodynamics*. Nat. Comm. **4**, 2059 (2013).

[13] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, *Resource Theory of Quantum States Out of Thermal Equilibrium*. Phys. Rev. Lett. **111**, 250404 (2013).

[14] W. van Dam and P. Hayden, *Universal entanglement transformations without communication*. Phys. Rev. A **67**, 060302 (2003).

[15] N. H. Y. Ng, L. Manĉinska, C. Cirstoiu, J. Eisert, and S. Wehner, *Limits to catalysis in quantum thermodynamics*. New J. Phys. **17** 081003 (2015).

[16] M. Lostaglio, M. P. Müller, and M. Pastena, *Stochastic Independence as a Resource in Small-Scale Thermodynamics*. Phys. Rev. Lett. **115**, 150402 (2015).

[17] P. Müller and M. Pastena, *A Generalization of Majorization that Characterizes Shannon Entropy*. IEEE Transactions on Information Theory **62**, 1711 (2016).

[18] H. Wilming, R. Gallego, and J. Eisert, *Axiomatic Characterization of the Quantum Relative Entropy and Free Energy*. Entropy **19**, 241 (2017).

[19] M. P. Müller, *Correlating Thermal Machines and the Second Law at the Nanoscale*. Phys. Rev. X **8**, 041051 (2018).

[20] S. Rethinasamy and M. M. Wilde, *Relative Entropy and Catalytic Relative Majorization*. Phys. Rev. Research **2**, 033455 (2020).

[21] F. Buscemi and G. Gour, *Quantum relative Lorenz curves*. Phys. Rev. A **95**, 012110 (2017).

[22] F. Buscemi, *Fully quantum second-law–like statements from the theory of statistical comparisons*. arXiv:1505.00535 (2015).

[23] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, *Quantum majorization and a complete set of entropic conditions for quantum thermodynamics*. Nat. Comm. **9**, 5352 (2018).

[24] M. Lostaglio and M. P. Müller, *Coherence and Asymmetry Cannot be Broadcast*. Phys. Rev. Lett. **123**, 020403 (2019).

[25] R. Dorner, S. R. Clark, L. Heaney, R. Fazio, J. Goold, and V. Vedral, *Extracting Quantum Work Statistics and Fluctuation Theorems by Single-Qubit Interferometry*. Phys. Rev. Lett. **110**, 230601 (2013).

[26] S. An, J. Zhang, M. Um, D. Lv, Y. Lu, J. Zhang, Z. Yin, H. T. Quan, and K. Kim, *Experimental Test of the Quantum Jarzynski Equality with a Trapped-ion System*, Nat. Phys. **11**, 193 (2014).

[27] T. B. Batalhao, A. M. Souza, L. Mazzola, R. Auccaise,R. S. Sarthour, I. S. Oliveira, J. Goold, G. DeChiara, M. Paternostro, R. M. Serra, *Experimental Reconstruction of Work Distribution and Study of Fluctuation Relations in a Closed Quantum System*, Phys. Rev. Lett., **113**, 140601

(2014).

[28] J. Rosnagel, S. T. Dawkins, K. N. Tolazzi, O. Abah, E. Lutz, F. Schmidt-Kaler, and K. Singer, *A singleatom heat engine*, Science **352**, 325 (2016).

[29] J. Klatzow, J. N. Becker, P. M. Ledingham, C. Weinzetl, K. T. Kaczmarek, D. J. Saunders, J. Nunn, I. A. Walmsley, R. Uzdin, and E. Poem, *Experimental Demonstration of Quantum Effects in the Operation of Microscopic Heat Engines*, Phys. Rev. Lett. **122**, 110601 (2017).

[30] N. Cottet, S. Jezouin, L. Bretheau, P. Campagne-Ibarcq, Q. Ficheux, J. Anders, A. Auffves, R. Azouitd, Pi. Rouchond, and B. Huard *Observing a Quantum Maxwell Demon at Work* PNAS, 114, 7561 (2017).

[31] Y. Masuyama, K. Funo, Y. Murashita, A. Noguchi, S. Kono, Y. Tabuchi, R. Yamazaki, M. Ueda and Y. Nakamura, *Information-to-work conversion by Maxwell's demon in a superconducting circuit quantum electrodynamical system.* Nat. Comm. **9**, 1291 (2018).

[32] R. van der Meer, N. H. Y. Ng, and S. Wehner, *Smoothed generalized free energies for thermodynamics.* Phys. Rev. A **96**, 062135 (2017).

[33] P. Faist, T. Sagawa, K. Kato, H. Nagaoka, and F. G. S. L. Brandão, *Macroscopic Thermodynamic Reversibility in Quantum Many-Body Systems.* Phys. Rev. Lett. **123**, 250601 (2019).

[34] T. Sagawa, P. Faist, K. Kato, K. Matsumoto, H. Nagaoka, and F. G. S. L. Brandao, *Asymptotic Reversibility of Thermal Operations for Interacting Quantum Spin Systems via Generalized Quantum Stein's Lemma.* arXiv:1907.05650

[35] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information.* Cambridge university press, (2010).

[36] See Supplemental Material at [proper url will appear here].

[37] M. P. Woods,, N. H. Y. Ng, and S. Wehner, *The maximum efficiency of nano heat engines depends on more than temperature.* Quantum 3, 177 (2019).

[38] E. Fermi, *Thermodynamics.* New York: Prentice Hall. (1937).

[39] H. Callen, *Thermodynamics: An Introduction to the Physical Theories of Equilibrium Thermostatics and Irreversible Thermodynamics.* New York: John Wiley and Sons (1960).

[40] L. D. Landau and E. M. Lifshitz, *Statistical Physics: Theory of the Condensed State.* Butterworth-Heinemann (1980).

[41] L. Wang and R. Renner, *One-Shot Classical-Quantum Capacity and Hypothesis Testing.* Phys. Rev. Lett. **108**, 200501 (2012).

[42] P. Faist and R. Renner, *Fundamental Work Cost of Quantum Processes.* Phys. Rev. X **8**, 021011 (2018).

[43] F. Hiai and D. Petz, *The proper formula for relative entropy and its asymptotics in quantum probability.* Comm. Math. Phys. **143**, 99 (1991).

[44] T. Ogawa and H. Nagaoka, *Strong Converse and Stein's Lemma in Quantum Hypothesis Testing.* Asymptotic Theory of Quantum Statistical Inference, 28 (2005).

[45] H. Nagaoka and M. Hayashi, *An Information-Spectrum Approach to Classical and Quantum Hypothesis Testing for Simple Hypotheses.* IEEE Transactions on Information Theory **53**, 534 (2007).

[46] N. Datta, *Min- and Max-Relative Entropies and a New Entanglement Monotone.* IEEE Transactions on Information Theory **55**, 2816 (2009).

[47] R. Duan, Y. Feng, X. Li, and M. Ying, *Multiple-copy entanglement transformation and entanglement catalysis.* Phys. Rev. A **71**, 042319 (2005).

[48] N. Shiraishi, *Two constructive proofs on d-majorization and thermo-majorization.* arXiv:2004.14751 (2020).

[49] P. Faist, J. Oppenheim, and R. Renner, *Gibbs-preserving maps outperform thermal operations in the quantum regime.* New J. Phys. **17**, 043003 (2015).

[50] I. Marvian and R. W. Spekkens, *No-Broadcasting Theorem for Quantum Asymmetry and Coherence and a Trade-off Relation for Approximate Broadcasting.* Phys. Rev. Lett. **123**, 020404 (2019).

[51] F. Sapienza, F. Cerisola and A. J. Roncaglia, *Correlations as a resource in quantum thermodynamics*, Nat. Comm. **10**, 2492 (2019).

[52] P. Skrzypczyk, A. J. Short, and S. Popescu, *Extracting work from quantum systems.* arXiv:1302.2811 (2013).

[53] J. Åberg, *Catalytic Coherence.* Phys. Rev. Lett. **113**, 150402 (2014).

[54] P. Skrzypczyk, A. J. Short, and S. Popescu, *Work extraction and thermodynamics for individual quantum systems.* Nat. Comm. **5**, 4185 (2014).

[55] K. Korzekwa, M. Lostaglio, J. Oppenheim, and D. Jennings, *The extraction of work from quantum coherence.* New J. Phys. **18**, 023045 (2016).

[56] F. Brandão and G. Gour, *Reversible Framework for Quantum Resource Theories.* Phys. Rev. Lett. **115**, 070503 (2015).

Supplemental Material for
"Quantum thermodynamics of correlated-catalytic state conversion at small-scale"

Naoto Shiraishi[1] and Takahiro Sagawa[2]

[1]Department of Physics, Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo 171-8588, Japan
[2]Department of Applied Physics and Quantum-Phase Electronics Center (QPEC), The University of Tokyo, 7-3-1 Hongo, Byunkyo-ku, Tokyo 113-8656, Japan

In this Supplemental Material, we provide rigorous proofs of the theorems presented in the main text. This Supplemental Material has the reference numbers in common with the main text.

A. Proof of Theorem 1 and Theorem 3

We first describe the fully rigorous statement of Theorem 3. We note that Theorem 3 includes Theorem 1.

Theorem 3– Consider four quantum states ρ, ρ', η, η' of the system S. The following two are equivalent.

1). $S_1(\rho||\eta) \geq S_1(\rho'||\eta')$.

2). For any $\varepsilon > 0$ and $\delta > 0$, there exist a catalyst system C, its state c and d, and a CPTP map $\mathcal{N}$ on the composite system SC such that $\mathcal{N}(\rho \otimes c) = \tau$ and $\mathcal{N}(\eta \otimes d) = \eta' \otimes d$, where τ satisfies $\mathrm{Tr}_S[\tau] = c$, $d_1(\mathrm{Tr}_C[\tau], \rho') < \varepsilon$, and $I_{SC}[\tau] < \delta$.

Here, $S_1(\rho||\eta) := \mathrm{Tr}[\rho \ln \rho - \rho \ln \eta]$ is the quantum Kullback-Leibler (KL) divergence, $d_1(\rho', \rho'') := \frac{1}{2}\mathrm{Tr}[|\rho' - \rho''|]$ is the trace distance, and $I_{\mathrm{SC}}[\tau] := S_1(\tau||\rho'' \otimes c)$ is the mutual information, where $\rho'' = \mathrm{Tr}_C[\tau]$ is a reduced state of τ on S. In Theorem 3, the map $\mathcal{N}$ exactly converts η to η' with a catalyst, and also converts ρ to ρ' with a catalyst within arbitrary accuracy. By setting $\eta = \eta' = \rho_{\mathrm{Gibbs}}$, and setting the catalyst Hamiltonian such that its Gibbs state is d, Theorem 3 reduces to Theorem 1 in the main text.

Proof. The proof of 2) $\Rightarrow$ 1) is easy. The additivity and the monotonicity of the KL divergence imply that

$$S_1(\rho||\eta) + S_1(c||d) = S_1(\rho \otimes c||\eta \otimes d) \geq S_1(\tau||\eta' \otimes d) \geq S_1(\mathrm{Tr}_C[\tau]||\eta') + S_1(c||d). \tag{S.1}$$

Since $\mathrm{Tr}_C[\tau]$ and ρ' are arbitrarily close, the above relation directly implies $S_1(\rho||\eta) \geq S_1(\rho'||\eta')$.

We now treat the difficult part: 1) $\Rightarrow$ 2). Without loss of generality, we can assume that $S_1(\rho||\eta) > S_1(\rho'||\eta')$, because in the case of $S_1(\rho||\eta) = S_1(\rho'||\eta')$, by replacing $\frac{\varepsilon}{2}$ to ε and ρ' to ρ'' such that $d_1(\rho', \rho'') < \frac{\varepsilon}{2}$ and $S_1(\rho||\eta) > S_1(\rho''||\eta')$, the proof is reduced to the case of $S_1(\rho||\eta) > S_1(\rho'||\eta')$. We prove 1) $\Rightarrow$ 2) in the following three steps.

Step 1: Deriving a sufficient condition for state conversion.

In the first step, we show a useful lemma. This lemma employs two kinds of divergences: the quantum hypothesis testing divergence and the quantum Rényi-∞ divergence.

The quantum hypothesis testing divergence of the first kind is defined as

$$S_{\mathrm{H}}^{1-\varepsilon}(\sigma||\kappa) := -\ln \left(\min_{0 \leq Q \leq I,\ \mathrm{Tr}[\sigma Q] \geq 1-\varepsilon} \mathrm{Tr}[\kappa Q] \right), \tag{S.2}$$

where I is the identity operator. In quantum hypothesis testing, an unknown quantum state which takes σ or κ is provided and the task is to identify it. In our setup, the error of the first (resp. second) kind is the probability that when the actual state is σ(resp. κ), we incorrectly guess that the state is κ(reps. σ). The quantum hypothesis testing divergence is the logarithm of the minimum error of the second kind under a fixed amount of the error of the first kind. The operator Q serves as a measurement operator for guessing that the state is σ (and $1-Q$ for κ), and the condition $\mathrm{Tr}[\sigma Q] \geq 1 - \varepsilon$ implies that the error of the first kind is less than ε (i.e., if the state is σ, we guess that the state is σ with probability at least $1 - \varepsilon$). Under this condition, we minimize $\mathrm{Tr}[\kappa Q]$, which is the probability that we guess that the state is σ when the actual state is κ.

We also introduce the quantum Rényi-∞ divergence $S_\infty(\sigma||\kappa)$ defined as

$$S_\infty(\sigma||\kappa) := \ln(\min[\lambda : \sigma \leq \lambda\kappa]). \tag{S.3}$$

We sometimes use this quantity in the form $\sigma \leq e^{S_\infty(\sigma||\kappa)}\kappa$. We now state the key lemma.

Lemma 1: Suppose that the following relation

$$S_{\mathrm{H}}^{1-\varepsilon}(\sigma||\kappa) \geq S_\infty(\sigma'||\kappa') \tag{S.4}$$

holds for a fixed $0 < \varepsilon < 1$. Then, there exists a CPTP map Λ which satisfies

$$\Lambda(\kappa) = \kappa', \tag{S.5}$$

$$d_1(\Lambda(\sigma), \sigma') < \varepsilon. \tag{S.6}$$

Proof of Lemma 1. The proof of Lemma 1 is based on the measurement-and-prepare method. In this proof, we employ the abbreviation $s := S_{\mathrm{H}}^{1-\varepsilon}(\sigma||\kappa)$ for brevity.

We first perform quantum measurement. By definition of the hypothesis testing divergence, there exists a nonnegative Hermitian operator A satisfying $\mathrm{Tr}[A\sigma] = 1-\varepsilon$ and $\mathrm{Tr}[A\sigma_{\mathrm{Gibbs}}] = e^{-s}$. Thus, by performing the POVM measurement with $(A, 1-A)$ on σ and κ, we map quantum states into classical probability distributions:

$$\sigma \to \begin{pmatrix} 1-\varepsilon \\ \varepsilon \end{pmatrix}, \quad \kappa \to \begin{pmatrix} e^{-s} \\ 1-e^{-s} \end{pmatrix}. \tag{S.7}$$

We next prepare quantum states depending on the classical probability distribution. If the classical state is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, we prepare σ', and if the classical state is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we prepare

$$\kappa'' := \frac{\kappa' - e^{-s}\sigma'}{1-e^{-s}}. \tag{S.8}$$

The relation $s \geq S_\infty(\sigma'||\kappa')$ guarantees that κ'' is nonnegative. Through this preparation, $\begin{pmatrix} 1-\varepsilon \\ \varepsilon \end{pmatrix}$ is converted into $\sigma'' := (1-\varepsilon)\sigma' + \varepsilon\kappa''$, which satisfies $d_1(\sigma'', \sigma') \leq \varepsilon$. In a similar manner, $\begin{pmatrix} e^{-s} \\ 1-e^{-s} \end{pmatrix}$ is converted into κ'. We thus obtain the desired CPTP map. $\square$

Step 2: Applying the quantum Stein's lemma.

We next apply the quantum Stein's lemma, which is one of the most important results in the field of quantum hypothesis testing. Ogawa and Nagaoka [44] proved that for multi-copy systems $\rho^{\otimes n}$ and $\eta^{\otimes n}$ any quantum hypothesis testing divergence rate with $0 < \varepsilon < 1$ converges to the KL divergence in the infinite copy limit:

$$\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{H}}^{1-\varepsilon}(\rho^{\otimes n}||\eta^{\otimes n}) = S_1(\rho||\eta). \tag{S.9}$$

This relation plays a crucial role to derive a single monotone in quantum thermodynamics (see also Ref. [34]).

Datta [46] found the connection between the quantum hypothesis testing divergence and the ε-smoothed Rényi divergence. Define the ε-neighborhood of ρ as $B^\varepsilon(\rho) := \{\tilde{\rho} : d_1(\rho, \tilde{\rho}) \leq \varepsilon\}$, where $\tilde{\rho}$ is a density operator. We define the ε-smoothed Rényi-∞ divergence as

$$S_\infty^\varepsilon(\rho||\eta) := \min_{\tilde{\rho}\in B^\varepsilon(\rho)} S_\infty(\tilde{\rho}||\eta). \tag{S.10}$$

By using the quantum Stein's lemma, Datta [46] showed that for any $0 < \varepsilon < \frac{1}{2}$ the ε-smoothed Rényi-∞ divergence rate also converges to the KL divergence in the infinite copy limit:

$$\lim_{n\to\infty} \frac{1}{n} S_\infty^\varepsilon(\rho^{\otimes n}||\eta^{\otimes n}) = S_1(\rho||\eta). \tag{S.11}$$

Combining Eqs.(S.9), (S.11), and Lemma 1, we arrive at the following key result: If $S_1(\rho||\eta) > S_1(\rho'||\eta')$ is satisfied, then for any $\varepsilon > 0$ there exists a sufficiently large n such that there is a CPTP map Λ which converts

$$\Lambda(\eta^{\otimes n}) = \eta'^{\otimes n} \tag{S.12}$$

and

$$\Lambda(\rho^{\otimes n}) = \Xi \tag{S.13}$$

with $d_1(\rho'^{\otimes n}, \Xi) < \varepsilon$.

Step 3: Reduction of asymptotic state conversion to catalytic conversion.

Our remaining task is to reduce the result of the approximate asymptotic (multi-copy) state conversion to the catalytic state conversion. To this end, we consider a refinement of the existing technique [1, 4, 47]. In Refs. [1, 47], they consider *exact* asymptotic state conversion and reduce it to catalytic conversion, while we consider *approximate* asymptotic state conversion and need to reduce it to catalytic conversion.

We set the Hilbert space of the catalyst system as $S^{\otimes n-1} \otimes A$, where A is spanned by $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$, and n is the large integer introduced in Step 2. We call the system S also as S_1, and $n-1$ copies of the same systems in the catalyst as $S_2, \dots, S_n$. The remaining subsystem of the catalyst A serves as a label.

Using Ξ on $S_1 \otimes S_2 \otimes \cdots \otimes S_n$, we introduce Ξ_i ($i \in \{1, \dots, n\}$) as a reduced state of Ξ on $S_1 \otimes \cdots \otimes S_i$. Note that we later consider the shift of Ξ_i in $S_1 \otimes \cdots \otimes S_n$. We interpret Ξ_0 as a trivial state 1 for convenience. Using these states, we set the state of the catalyst c as

$$c := \frac{1}{n} \sum_{k=1}^{n} \rho^{\otimes k-1} \otimes \Xi_{n-k} \otimes |k\rangle\langle k|, \tag{S.14}$$

where $\rho^{\otimes k-1}$ is on $S_2 \otimes \cdots \otimes S_k$, and Ξ_{n-k} is now on $S_{k+1} \otimes \cdots \otimes S_n$. In a similar manner, we set d as

$$d := \frac{1}{n} \sum_{k=1}^{n} \eta^{\otimes k-1} \otimes \eta'^{\otimes n-k} \otimes |k\rangle\langle k|. \tag{S.15}$$

We now construct the desired CPTP map, which consists of the following three processes:

I. If the auxiliary system A is $|n\rangle\langle n|$, then we apply Λ (obtained in Step 2) to $S_1 \otimes \cdots \otimes S_n$. Otherwise, we leave $S_1 \otimes \cdots \otimes S_n$ as it is.

II. We convert the auxiliary system A as $|n\rangle \to |1\rangle$ and $|i\rangle \to |i+1\rangle$.

III. We shift the systems as $S_i \to S_{i+1}$ and $S_n \to S_1$.

Through the processes I and II, the initial state

$$\rho \otimes c = \frac{1}{n} \sum_{k=1}^{n} \rho^{\otimes k} \otimes \Xi_{n-k} \otimes |k\rangle\langle k| \tag{S.16}$$

is converted into

$$\tau' = \frac{1}{n} \sum_{k=1}^{n} \rho^{\otimes k-1} \otimes \Xi_{n+1-k} \otimes |k\rangle\langle k|. \tag{S.17}$$

We note that the partial trace of τ' with respect to S_n recovers the initial state of the catalyst c. We denote by τ the state after the shift process III on τ'. In a similar manner, through the processes I and II, the initial state

$$\eta \otimes d = \frac{1}{n} \sum_{k=1}^{n} \eta^{\otimes k} \otimes \eta'^{\otimes n-k} \otimes |k\rangle\langle k| \tag{S.18}$$

is converted into

$$\frac{1}{n} \sum_{k=1}^{n} \eta^{\otimes k-1} \otimes \eta'^{\otimes n-k+1} \otimes |k\rangle\langle k|. \tag{S.19}$$

Through the shift process III, the above term (S.19) becomes $d \otimes \eta'$.

We finally show that $\mathrm{Tr}_C[\tau]$ is ε-close to ρ' and the correlation is small. We first treat the former. The state $\mathrm{Tr}_C[\tau]$ is expressed as

$$\mathrm{Tr}_C[\tau] = \frac{1}{n}\sum_{k=1}^{n} \xi_k, \tag{S.20}$$

where we defined ξ_k as the reduced state of Ξ on S_k: $\xi_k := \mathrm{Tr}_{[1,2,\dots,k-1,k+1,\dots,n]}[\Xi]$. Then, using the monotonicity and convexity of the trace distance d_1, we arrive at the desired result

$$d_1\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k, \rho'\right) \le \frac{1}{n}\sum_{k=1}^{n} d_1(\xi_k, \rho') \le \frac{1}{n}\sum_{k=1}^{n} d_1(\Xi, \rho'^{\otimes n}) < \varepsilon. \tag{S.21}$$

We next prove the latter condition (i.e., $I_{SC}[\tau] < \delta$). By construction, Ξ can be written in the form of $\Xi = (1-\varepsilon)\rho'^{\otimes n} + \varepsilon X$, and hence τ is written as

$$\tau = (1-\varepsilon)\rho' \otimes \left(\frac{1}{n}\sum_{k=1}^{n} \rho^{\otimes k} \otimes \rho'^{\otimes n-k-1} \otimes |k\rangle\langle k|\right) + \sum_{i=1}^{d} \varepsilon_i \rho_i \otimes c_i, \tag{S.22}$$

where $\sum_{i=1}^{d} \varepsilon_i = \varepsilon$, d is the dimension of the Hilbert space of S, and ρ_i and c_i are some states of S and C, respectively. Then, using the subadditivity and the Araki-Lieb inequality for the von Neumann entropy [35], the correlation between the system and the catalyst is evaluated as

$$I_{SC}[\tau] \le 2\left[-(1-\varepsilon)\ln(1-\varepsilon) - \sum_{i=1}^{d}\varepsilon_i \ln \varepsilon_i\right] \le 2\left[-(1-\varepsilon)\ln(1-\varepsilon) - \varepsilon \ln\frac{\varepsilon}{d}\right]. \tag{S.23}$$

Thus, by replacing ε to $\tilde{\varepsilon}$ such that $2\left[-(1-\tilde{\varepsilon})\ln(1-\tilde{\varepsilon}) - \tilde{\varepsilon}\ln\tilde{\varepsilon}/d\right] \le \delta$, we obtain the desired inequality $I_{SC}[\tau] < \delta$. This completes the proof of Theorem 3. □

B. Proof of Theorem 2

We first present the rigorous statement of Theorem 2:

Theorem 2– Consider two states ρ and ρ' of the system S with $F(\rho) - F(\rho') < 0$. Then, the following two conditions are equivalent.

1). $F(\rho) - F(\rho') \ge \beta w$.

2). For any $t > 0$, $u > 0$, and $\delta > 0$, there exist a catalyst C, its state c, a work storage W with two eigenstates $|a\rangle$, $|b\rangle$ with energies $E_a = 0$ and $E_b = w < 0$, and a Gibbs-preserving map $\rho \otimes c \otimes |a\rangle\langle a| \to \tau \otimes \omega$ such that (i) $\mathrm{Tr}_S[\tau] = c$, (ii) $d_1(\mathrm{Tr}_C[\tau], \rho') < t$, (iii) $d_1(\omega, |b\rangle\langle b|) < u$, and (iv) $I_{SC}[\tau] < \delta$.

Proof. The proof of 2) ⇒ 1) is straightforward as in the case of Theorem 3. The additivity and the monotonicity of the KL divergence imply that

$$\begin{aligned} S_1(\rho||\rho_{\mathrm{Gibbs}}) + S_1(c||c_{\mathrm{Gibbs}}) + S_1(|a\rangle\langle a|\,||\omega_{\mathrm{Gibbs}}) =& S_1(\rho \otimes c \otimes |a\rangle\langle a|\,||\rho_{\mathrm{Gibbs}} \otimes c_{\mathrm{Gibbs}} \otimes \omega_{\mathrm{Gibbs}}) \\ \ge& S_1(\tau \otimes \omega||\rho_{\mathrm{Gibbs}} \otimes c_{\mathrm{Gibbs}} \otimes \omega_{\mathrm{Gibbs}}) \\ \ge& S_1(\mathrm{Tr}_C[\tau]||\rho_{\mathrm{Gibbs}}) + S_1(c||c_{\mathrm{Gibbs}}) + S_1(\omega||\omega_{\mathrm{Gibbs}}), \end{aligned} \tag{S.24}$$

where c_{Gibbs} is the Gibbs state of C, and we defined

$$\omega_{\mathrm{Gibbs}} := \frac{1}{z}|a\rangle\langle a| + \frac{e^{-\beta w}}{z}|b\rangle\langle b| \tag{S.25}$$

with $z := 1 + e^{-\beta w}$ is the Gibbs state of W. Since ω is arbitrarily close to $|b\rangle\langle b|$ and $S_1(|b\rangle\langle b|\,||\omega_{\mathrm{Gibbs}}) - S_1(|a\rangle\langle a|\,||\omega_{\mathrm{Gibbs}}) = \beta w$, we arrive at the desired relation $F(\rho) - F(\rho') \ge \beta w$.

We now prove 1) ⇒ 2). In order to apply the reduction from asymptotic state conversion to catalytic conversion (Step 3 of Theorem 3), we need the following lemma, which serves as a combination of Step 1 and Step 2:

Lemma 2: Suppose that the following relation holds

$$S_1(\rho||\rho_{\rm Gibbs}) > S_1(\rho'||\rho_{\rm Gibbs}) + \beta w. \tag{S.26}$$

Then, for any $t>0$ and $u>0$, there exists an integer m and a Gibbs-preserving map such that

$$\rho^{\otimes m}\otimes |a\rangle\langle a|^{\otimes m} \to \Xi\otimes\omega^{\otimes m} \tag{S.27}$$

with

$$d_1(\Xi,\rho'^{\otimes m}) < t, \tag{S.28}$$

$$d_1(\omega,|b\rangle\langle b|) < u, \tag{S.29}$$

where we set $E_a=0$ and $E_b=w<0$.

We note that, for the same reason as the proof of Theorem 3, we only have to consider the case of $S_1(\rho||\rho_{\rm Gibbs}) > S_1(\rho'||\rho_{\rm Gibbs}) + \beta w$ (without the equality case).

Proof of Lemma 2. Without loss of generality, we suppose that $t<\frac{1}{2}$ and $u<\min(e^{2w},\frac{e^w}{2})$. We set $\varepsilon=\frac{t}{2}$ and $m=\max(\frac{2e^{2w}}{u},M,2)$, where M is a sufficiently large integer such that $m\geq M$ implies

$$S_{\rm H}^{1-\varepsilon}(\rho^{\otimes m}||\rho_{\rm Gibbs}^{\otimes m}) \geq S_\infty(\rho'^{\otimes m}||\rho_{\rm Gibbs}^{\otimes m}) + \beta m w. \tag{S.30}$$

The existence of such an integer M is guaranteed by the quantum Stein's lemma. We denote $\sigma:=\rho^{\otimes m}$, $\sigma':=\rho'^{\otimes m}$, $\sigma_{\rm Gibbs}:=\rho_{\rm Gibbs}^{\otimes m}$, and $s:=S_{\rm H}^{1-\varepsilon}(\sigma||\sigma_{\rm Gibbs})$, $s':=S_\infty(\sigma'||\sigma_{\rm Gibbs})$ for brevity.

We first perform not binary but ternary measurement on the composite system. By definition of the hypothesis testing divergence, there exists a Hermitian operator A satisfying ${\rm Tr}[A\sigma]=1-\varepsilon$ and ${\rm Tr}[A\sigma_{\rm Gibbs}]=e^{-s}$. Using this operator, we construct the POVM measurement with $\{A\otimes|a\rangle\langle a|^{\otimes m}, (1-A)\otimes|a\rangle\langle a|^{\otimes m}, 1\otimes(1-|a\rangle\langle a|^{\otimes m})\}$, for which we denote the measurement output by 1, 2, and 3, respectively. Then, if the state is $\sigma\otimes|a\rangle\langle a|^{\otimes m}$, the measurement outcome is 1 with probability $1-\varepsilon$, 2 with probability ε, and 3 with probability 0. If the state is $\sigma_{\rm Gibbs}\otimes\omega_{\rm Gibbs}^{\otimes m}$, the measurement outcome is 1 with probability $e^{-s}\frac{1}{z^m}$, 2 with probability $(1-e^{-s})\frac{1}{z^m}$, and 3 with probability $1-\frac{1}{z^m}$.

Then, we prepare quantum states based on the measurement outcome. If the measurement output is i ($i=1,2,3$), we prepare a quantum state as

$$\zeta^i := \sum_{y\in\{a,b\}^{\otimes m}} \left[v_{1y}^i\sigma'\otimes|y\rangle\langle y| + v_{0y}^i\sigma_{\rm Gibbs}\otimes|y\rangle\langle y|\right], \tag{S.31}$$

where y runs 2^m possible sequences of $\{a,b\}^{\otimes m}$ and v_{jy}^i's are coefficients. The normalization condition requires $\sum_{jy}v_{jy}^i=1$ for any i. In addition, ζ^i is positive-semidefinite if

$$v_{0y}^i e^{-s'} + v_{1y}^i \geq 0 \tag{S.32}$$

holds for any i and y.

We set v_{jy}^i's as

$$v_{1y}^1 = \frac{1-t}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)}, \tag{S.33}$$

$$v_{0y}^1 = \frac{t-\varepsilon}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)}, \tag{S.34}$$

$$v_{1y}^2 = 0, \tag{S.35}$$

$$v_{0y}^2 = u^{N_a(y)}(1-u)^{N_b(y)}, \tag{S.36}$$

$$v_{1y}^3 = -\frac{1}{z^m-1}e^{-s}\frac{1-t}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)}, \tag{S.37}$$

$$v_{0y}^3 = \frac{1}{z^m-1}\left[e^{-wN_b(y)} - e^{-s}\frac{t-\varepsilon}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)} - (1-e^{-s})u^{N_a(y)}(1-u)^{N_b(y)}\right], \tag{S.38}$$

where $N_a(y)$ (resp. $N_b(x)$) is the number of a (resp. b) in the sequence y. Remarkably, these v^i_{jy}'s satisfy

$$(1-\varepsilon)v^1_{1y} + \varepsilon v^2_{1y} =(1-t)u^{N_a(y)}(1-u)^{N_b(y)}, \tag{S.39}$$
$$(1-\varepsilon)v^1_{0y} + \varepsilon v^2_{0y} =tu^{N_a(y)}(1-u)^{N_b(y)}, \tag{S.40}$$

and

$$\frac{e^{-s}}{z^m}v^1_{1y} + \frac{1-e^{-s}}{z^m}v^2_{1y} + \left(1-\frac{1}{z^m}\right)v^3_{1y} = 0, \tag{S.41}$$
$$\frac{e^{-s}}{z^m}v^1_{0y} + \frac{1-e^{-s}}{z^m}v^2_{0y} + \left(1-\frac{1}{z^m}\right)v^3_{0y} = \frac{e^{-wN_b(y)}}{z^m}, \tag{S.42}$$

for any y, which implies that this measurement-and-preparation CPTP map converts

$$\sigma\otimes |a\rangle\langle a|^{\otimes m} \to ((1-t)\sigma' + t\sigma_{\mathrm{Gibbs}})\otimes (u|a\rangle\langle a| + (1-u)|b\rangle\langle b|)^{\otimes m}, \tag{S.43}$$
$$\sigma_{\mathrm{Gibbs}}\otimes \omega^{\otimes m}_{\mathrm{Gibbs}} \to \sigma_{\mathrm{Gibbs}}\otimes \omega^{\otimes m}_{\mathrm{Gibbs}}. \tag{S.44}$$

Namely, the constructed CPTP map is the desired Gibbs-preserving map.

We finally demonstrate that these v^i_{jy}'s satisfy the positive-semidefinite condition (Eq. (S.32)). Since the cases of $i=1,2$ are trivial, below we only treat the case of $i=3$. The positive-semidefinite condition $v^3_{0y}e^{-s'} + v^3_{1y} \geq 0$ can be transformed as

$$e^{-wN_b(y)} \geq e^{s'-s}\frac{1-t}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)} + e^{-s}\frac{t-\varepsilon}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)} + (1-e^{-s})u^{N_a(y)}(1-u)^{N_b(y)}. \tag{S.45}$$

Owing to $s\geq s'+mw$ and $N_a(y)+N_b(y)=m$, it suffices to prove

$$1 \geq e^{-N_a(y)w}\frac{1-t}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)} + e^{-N_a(y)w-s'}\frac{t-\varepsilon}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)} + e^{N_b(y)w}(1-e^{-s})u^{N_a(y)}(1-u)^{N_b(y)} \tag{S.46}$$

for any y. Finally, using $e^{2w}/m < u < e^{2w}$, $u < e^w/2 \leq 1/(1+e^{-w})$, and $s'\geq 0$, we prove Eq. (S.46) as

$$\begin{aligned}
&e^{-N_a(y)w}\frac{1-t}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)} + e^{-N_a(y)w-s'}\frac{t-\varepsilon}{1-\varepsilon}u^{N_a(y)}(1-u)^{N_b(y)} + e^{N_b(y)w}(1-e^{-s})u^{N_a(y)}(1-u)^{N_b(y)}\\
\leq&(1+e^{mw})e^{-N_a(y)w}u^{N_a(y)}(1-u)^{N_b(y)}\\
\leq&(1+e^{mw})(1-u)^m\\
\leq&(1+e^{mw})\left(1-\frac{e^{2w}}{m}\right)^m\\
<&(1+e^{2w})\frac{1}{e^{(e^{2w})}}\\
\leq&(1+e^{2w})\frac{1}{1+e^{2w}}\\
=&1.
\end{aligned} \tag{S.47}$$

In the second line we used $e^{-s'}\leq 1$ and $N_a(y)+N_b(y)=m$, in third line we used $e^{-w}u < \frac{1}{2} < 1-u$, and in the fifth line we used $m\geq 2$ and $(1-a/m)^m < 1/e^a$. This completes the proof. □

Lemma 2 serves as the counterpart of Step 1 and Step 2 of the proof of Theorem 3. By combining the idea of Step 3 of Theorem 3, we prove Theorem 2 as follows.

We set the catalyst as

$$c := \frac{1}{m}\sum_{k=1}^{m}\rho^{\otimes k-1}\otimes \Xi_{m-k}\otimes |a\rangle\langle a|^{\otimes k-1}\otimes \omega^{\otimes m-k}\otimes |k\rangle\langle k|, \tag{S.48}$$

where we set $\Xi = (1-t)\rho'^{\otimes m} + t\rho^{\otimes m}_{\mathrm{Gibbs}}$ and $\omega = u|a\rangle\langle a| + (1-u)|b\rangle\langle b|$, and the definition of Ξ_{m-k} is the same as the proof of Theorem 3. The choices of t, u, and m are the same as Lemma 2.

We follow Step 3 of the proof of Theorem 3 by setting the Gibbs preserving map to that derived in Lemma 2 and by replacing S with $S \otimes W$. Then, the initial state of the composite system SWC is

$$\rho \otimes c \otimes |a\rangle\langle a| = \frac{1}{m}\sum_{k=1}^{m} \rho^{\otimes k} \otimes \Xi_{m-k} \otimes |a\rangle\langle a|^{\otimes k} \otimes \omega^{\otimes m-k} \otimes |k\rangle\langle k| \tag{S.49}$$

and the final state is

$$\frac{1}{m}\sum_{k=1}^{m} \rho^{\otimes k-1} \otimes \Xi_{m-k+1} \otimes |a\rangle\langle a|^{\otimes k-1} \otimes \omega^{\otimes m-k+1} \otimes |k\rangle\langle k|\,. \tag{S.50}$$

We thus have

$$\mathcal{N}(\rho \otimes c \otimes |a\rangle\langle a|) = \tau \otimes \omega \tag{S.51}$$

with $\mathrm{Tr}_S[\tau] = c$, $d_1(\mathrm{Tr}_C[\tau], \rho') < t$, $d_1(\omega, |b\rangle\langle b|) < u$. In addition, by setting t sufficiently small if needed, we obtain $I_{SC}[\tau] < \delta$. □

C. Sufficient condition for work investment/extraction without catalyst

As a side remark, we show an interesting lemma: a necessary condition for state conversion with a two-level work storage W with states a and b. Unfortunately, this lemma does not directly apply to our setup of Theorem 2. However, this proof method inspires the proof of Theorem 2, and this lemma itself has potential applications to quantum thermodynamics. In particular, this lemma applies in the case of work extraction.

Lemma 3: Let σ_{Gibbs} be the Gibbs state of the system S and consider two states σ and σ' of S. Suppose that

$$S_{\mathrm{H}}^{1-\varepsilon}(\sigma||\sigma_{\mathrm{Gibbs}}) \geq S_\infty(\sigma'||\sigma_{\mathrm{Gibbs}}) + \beta w, \tag{S.52}$$

where $\varepsilon > 0$ is sufficiently small (a more detailed condition is given soon later). We denote $s := S_{\mathrm{H}}^{1-\varepsilon}(\sigma||\sigma_{\mathrm{Gibbs}})$ and $s' := S_\infty(\sigma'||\sigma_{\mathrm{Gibbs}})$ for brevity, and define $Z := 1 + e^{-\beta w}$ and $q := 1 - e^{-s'-\beta w}/Z$. We assume that ε satisfies $0 < \varepsilon < \min(\frac{1}{2}, \frac{4q^5}{(1+q)^2})$.

We claim that there exist a two-level system called work storage W with two eigenstates $|a\rangle$, $|b\rangle$ of the Hamiltonian with energies $E_a = 0$ and $E_b = w$ and a Gibbs-preserving map on the composite system SW such that it converts

$$\sigma \otimes |a\rangle\langle a| \to \tilde{\sigma} \otimes \Omega \tag{S.53}$$

with

$$d_1(\tilde{\sigma}, \sigma') < 2\sqrt{\frac{\varepsilon}{q}}, \tag{S.54}$$

$$d_1(\Omega, |b\rangle\langle b|) < 2\sqrt{\frac{\varepsilon}{q}}. \tag{S.55}$$

The proof is very similar to that of Lemma 1 and Lemma 2. We construct the desired Gibbs-preserving map by employing the measurement-and-preparation method as follows.

Proof. We write the Gibbs state of the work storage W as

$$\Omega_{\mathrm{Gibbs}} := \frac{1}{Z}|a\rangle\langle a| + \frac{e^{-\beta w}}{Z}|b\rangle\langle b|\,. \tag{S.56}$$

By definition of the hypothesis testing divergence, there is a nonnegative Hermitian operator A satisfying $\mathrm{Tr}[A\sigma] = 1 - \varepsilon$ and $\mathrm{Tr}[A\sigma_{\mathrm{Gibbs}}] = e^{-s}$. Using this operator, we first perform a POVM measurement with $\{A \otimes |a\rangle\langle a|, 1 - A \otimes |a\rangle\langle a|\}$ on the composite system of the system and the work storage, which maps

$$\sigma \otimes |a\rangle\langle a| \to \begin{pmatrix} 1-\varepsilon \\ \varepsilon \end{pmatrix}, \quad \sigma_{\mathrm{Gibbs}} \otimes \Omega_{\mathrm{Gibbs}} \to \begin{pmatrix} \frac{e^{-s}}{Z} \\ 1 - \frac{e^{-s}}{Z} \end{pmatrix}. \tag{S.57}$$

We then apply a classical stochastic map M such that

$$\begin{pmatrix}1-\varepsilon'\\ \varepsilon'\end{pmatrix}=M\begin{pmatrix}1-\varepsilon\\ \varepsilon\end{pmatrix},\quad \begin{pmatrix}\frac{e^{-s'-\beta w}}{Z}\\ 1-\frac{e^{-s'-\beta w}}{Z}\end{pmatrix}=M\begin{pmatrix}\frac{e^{-s}}{Z}\\ 1-\frac{e^{-s}}{Z}\end{pmatrix} \tag{S.58}$$

with $\varepsilon'\le\varepsilon$, which is realized by

$$M:=\begin{pmatrix}1 & 1-\frac{Z-e^{-s'-\beta w}}{Z-e^{-s}}\\ 0 & \frac{Z-e^{-s'-\beta w}}{Z-e^{-s}}\end{pmatrix}. \tag{S.59}$$

The condition $s\ge s'+\beta w$ guarantees the nonnegativity of matrix elements.

We now consider the preparation step. By defining $p_{\mathrm{w}}^{\mathrm{G}}:=e^{-\beta w}/Z$ and $u:=\sqrt{\varepsilon/q}$, we introduce two states ζ_1 and ζ_2 as

$$\begin{aligned}\zeta_1:=&\frac{(1-u)uq}{q-\varepsilon'}\sigma'\otimes|a\rangle\langle a|+\frac{(1-u)^2q}{q-\varepsilon'}\sigma'\otimes|b\rangle\langle b|\\ &+\frac{u^2q-(1-p_{\mathrm{w}}^{\mathrm{G}})\varepsilon'}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}\otimes|a\rangle\langle a|+\frac{u(1-u)q-p_{\mathrm{w}}^{\mathrm{G}}\varepsilon'}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}\otimes|b\rangle\langle b|,\end{aligned} \tag{S.60}$$

$$\begin{aligned}\zeta_2:=&-\frac{(1-q)(1-u)u}{q-\varepsilon'}\sigma'\otimes|a\rangle\langle a|-\frac{(1-q)(1-u)^2}{q-\varepsilon'}\sigma'\otimes|b\rangle\langle b|\\ &+\frac{(1-\varepsilon')(1-p_{\mathrm{w}}^{\mathrm{G}})-(1-q)u^2}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}\otimes|a\rangle\langle a|+\frac{(1-\varepsilon')p_{\mathrm{w}}^{\mathrm{G}}-(1-q)u(1-u)}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}\otimes|b\rangle\langle b|.\end{aligned} \tag{S.61}$$

Using these states, we prepare a quantum state from a classical distribution as

$$\begin{pmatrix}p_1\\ p_2\end{pmatrix}\to p_1\zeta_1+p_2\zeta_2. \tag{S.62}$$

This preparation realizes the desired state conversion:

$$\begin{pmatrix}1-\varepsilon'\\ \varepsilon'\end{pmatrix}\to((1-u)\sigma'+u\sigma_{\mathrm{Gibbs}})\otimes((1-u)|b\rangle\langle b|+u|a\rangle\langle a|), \tag{S.63}$$

$$\begin{pmatrix}\frac{e^{-s'-\beta w}}{Z}\\ 1-\frac{e^{-s'-\beta w}}{Z}\end{pmatrix}=\begin{pmatrix}1-q\\ q\end{pmatrix}\to\sigma_{\mathrm{Gibbs}}\otimes\Omega_{\mathrm{Gibbs}}. \tag{S.64}$$

We now examine the conditions $0\le\zeta_1\le 1$ and $0\le\zeta_2\le 1$. We first consider ζ_1. It suffices to confirm the following inequalities:

$$\frac{u^2q-(1-p_{\mathrm{w}}^{\mathrm{G}})\varepsilon'}{q-\varepsilon'}\ge 0, \tag{S.65}$$

$$\frac{u(1-u)q-p_{\mathrm{w}}^{\mathrm{G}}\varepsilon'}{q-\varepsilon'}\ge 0. \tag{S.66}$$

Since $\frac{u(1-u)q-p_{\mathrm{w}}^{\mathrm{G}}\varepsilon'}{q-\varepsilon'}\ge\frac{u^2q-(1-p_{\mathrm{w}}^{\mathrm{G}})\varepsilon'}{q-\varepsilon'}$, we only have to show Eq. (S.65), which is easily verified as

$$u^2q-(1-p_{\mathrm{w}}^{\mathrm{G}})\varepsilon'\ge u^2q-\varepsilon'\ge u^2q-\varepsilon=0. \tag{S.67}$$

We next treat ζ_2. It suffices to confirm the following inequalities:

$$\frac{(1-\varepsilon')(1-p_{\mathrm{w}}^{\mathrm{G}})-(1-q)u^2}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}-\frac{(1-q)(1-u)u}{q-u}\sigma'\ge 0, \tag{S.68}$$

$$\frac{(1-\varepsilon')p_{\mathrm{w}}^{\mathrm{G}}-(1-q)u(1-u)}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}-\frac{(1-q)(1-u)^2}{q-u}\sigma'\ge 0. \tag{S.69}$$

We only have to show the latter one (Eq. (S.69)) because

$$\frac{(1-\varepsilon')(1-p_{\mathrm{w}}^{\mathrm{G}})-(1-q)u^2}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}-\frac{(1-q)(1-u)u}{q-u}\sigma'\ge\frac{(1-\varepsilon')p_{\mathrm{w}}^{\mathrm{G}}-(1-q)u(1-u)}{q-\varepsilon'}\sigma_{\mathrm{Gibbs}}-\frac{(1-q)(1-u)^2}{q-u}\sigma'. \tag{S.70}$$

To verify Eq. (S.69), we show the following inequality:

$$\begin{aligned}\frac{q-\varepsilon'}{q-u}\frac{(1-q)(1-u)^2}{(1-\varepsilon')p_{\rm w}^{\rm G}-(1-q)u(1-u)} =&e^{-s'}\frac{q-\varepsilon'}{q-u}\frac{(1-q)(1-u)}{1-\varepsilon'-e^{-s'}u(1-u)}\\ \leq&e^{-s'}\frac{q-\varepsilon'}{q-u}\frac{(1-q)(1-u)}{1-\varepsilon'-u(1-u)}\\ \leq&e^{-s'}\frac{q}{q-u}\frac{1-q}{1-u}\\ \leq&e^{-s'}.\end{aligned} \tag{S.71}$$

In the first line we used $\frac{1-q}{p_{\rm w}^{\rm G}}=e^{-s'}$, in the second line we used $e^{-s'}\leq 1$, and in the third line we used $0<\varepsilon'\leq\varepsilon=u^2q\leq u$. In the last line, we used $q(1-q)\leq(q-u)(1-u)$ for $u\leq\frac{2q^2}{1+q}<\frac{1+q-\sqrt{1+2q-3q^2}}{2}$ and $u^2=\frac{\varepsilon}{q}$. Using $\sigma_{\rm Gibbs}-e^{-s'}\sigma'\geq 0$, which comes from the definition of $s'=S_\infty(\sigma'||\sigma_{\rm Gibbs})$, we arrive at the desired result:

$$\begin{aligned}&\frac{(1-\varepsilon')p_{\rm w}^{\rm G}-(1-q)u(1-u)}{q-\varepsilon'}\sigma_{\rm Gibbs}-\frac{(1-q)(1-u)^2}{q-u}\sigma'\\ =&\frac{(1-\varepsilon')p_{\rm w}^{\rm G}-(1-q)u(1-u)}{q-\varepsilon'}\left(\sigma_{\rm Gibbs}-\frac{q-\varepsilon'}{q-u}\frac{(1-q)(1-u)^2}{(1-\varepsilon')p_{\rm w}^{\rm G}-(1-q)u(1-u)}\sigma'\right)\\ \geq&\frac{(1-\varepsilon')p_{\rm w}^{\rm G}-(1-q)u(1-u)}{q-\varepsilon'}(\sigma_{\rm Gibbs}-e^{-s'}\sigma')\\ \geq&0.\end{aligned} \tag{S.72}$$

□

D. Details of the toy example in the main text

We here prove

$$\rho_{\rm Gibbs}^{\otimes 8}-\lambda\left|+\right\rangle\left\langle +\right|^{\otimes 8}\geq 0 \tag{S.73}$$

for $\lambda\leq\left(\frac{3}{8}\right)^8$. We first decompose the energy eigenstates of the Hamiltonian $H=\left(\frac{3}{4}\left|0\right\rangle\left\langle 0\right|+\frac{1}{4}\left|1\right\rangle\left\langle 1\right|\right)^{\otimes 8}$ as

$$\left|E_i\right\rangle=a_i\left|+\right\rangle^{\otimes 8}+b_i\left|X\right\rangle, \tag{S.74}$$

where $\left|X\right\rangle$ is some state of 8 two-level systems orthogonal to $\left|+\right\rangle^{\otimes 8}$. Since all of 2^8 computational basis states $\left|00\cdots 00\right\rangle$, $\left|00\cdots 01\right\rangle$, ..., $\left|11\cdots 11\right\rangle$ are energy eigenstates of H, we find that $a_i=\frac{1}{2^{8/2}}=\frac{1}{16}$ for any i.

Note that maximum of λ satisfying Eq. (S.73), which we denote by λ^*, is expressed as

$$\lambda^*=\min_{\left|\phi\right\rangle}\frac{\left\langle\phi\right|\rho_{\rm Gibbs}^{\otimes 8}\left|\phi\right\rangle}{\left\langle\phi\right|\left(\left|+\right\rangle\left\langle +\right|^{\otimes 8}\right)\left|\phi\right\rangle}=\min_{c_i}\frac{\sum_i\left|c_i\right|^2p_i^{\rm G}}{(\sum_i\left|c_ia_i\right|)^2}=\frac{\sum_i\left|c_i^*\right|^2p_i^{\rm G}}{(\sum_i\left|c_i^*a_i\right|)^2}, \tag{S.75}$$

where $\left|\phi\right\rangle=\sum_i c_i\left|E_i\right\rangle$ runs all possible states spanned by $\{\left|0\right\rangle,\left|1\right\rangle\}^{\otimes 8}$, and c_i^* is the optimal choice of this minimization. In addition, $p_i^{\rm G}$ is the Boltzmann weight of the energy eigenstate $\left|E_i\right\rangle$ given by $p_i^{\rm G}=3^{N_i}/4^8$, where N_i is the number of $\left|0\right\rangle$'s in the energy eigenstate $\left|E_i\right\rangle$. Applying the Schwarz inequality to the right-hand side of Eq. (S.75), we have

$$\lambda^*=\frac{\sum_i\left|c_i^*\right|^2p_i^{\rm G}}{(\sum_i\left|c_i^*a_i\right|)^2}\geq\frac{1}{\sum_i\frac{\left|a_i\right|^2}{p_i^{\rm G}}}=\frac{1}{2^8\sum_i 3^{-N_i}}=\left(\frac{3}{8}\right)^8. \tag{S.76}$$

This relation directly implies that $\rho_{\rm Gibbs}^{\otimes 8}-\lambda\left|+\right\rangle\left\langle +\right|^{\otimes 8}\geq 0$ for $\lambda\leq\left(\frac{3}{8}\right)^8$.