# Almost thermal operations: inhomogeneous reservoirs


Angeline Shu,[1] Yu Cai,[2] Stella Seah,[1] Stefan Nimmrichter,[3] and Valerio Scarani[1,4]

[1]Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542.
[2]Département de Physique Appliquée, Université de Genéve, 1221 Genéve, Switzerland.
[3]Max Planck Institute for the Science of Light, Staudtstraße 2, 91058 Erlangen.
[4]Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543.



The resource theory of thermal operations explains the state transformations that are possible in a very specific thermodynamic setting: there is only one thermal bath, auxiliary systems can only be in the corresponding thermal state (free states), and the interaction must commute with the free Hamiltonian (free operation). In this paper we study the mildest deviation: the reservoir particles are subject to inhomogeneities, either in the local temperature (introducing resource states) or in the local Hamiltonian (generating a resource operation). For small inhomogeneities, the two models generate the same channel and thus the same state transformations. However, their thermodynamics is significantly different when it comes to work generation or to the interpretation of the "second laws of thermal operations".



arXiv:1904.08736v1 [quant-ph] 18 Apr 2019


## I. INTRODUCTION

Foundationally, thermodynamics is a theory of states and their transformations. In quantum information science, the same can be said for entanglement theory. This analogy was discussed very early [1, 2], and has later resulted in the development of the broad framework of resource theories. Among those, the *resource theory of thermal operations* is a formalisation of the thermodynamics of systems in contact with thermal baths [3–5]. The lack of resources is described by what can be achieved with a single thermal bath at temperature $T$ (because with two different temperatures one can run an engine). Specifically, the *free states* are the thermal states $\tau$ at temperature $T$, and the *free operations* $U$ are those that conserve the total energy. Both notions are defined with respect to a reference Hamiltonian, usually taken as $H = H_S + H_R$ where $S$ indicate the system and $R$ a reservoir of auxiliary systems. Then, thermal states read $\tau = \tau_S \otimes \tau_R$ where $\tau_X = e^{-\beta H_X}/Z_X$, $Z_X = \mathrm{tr}(e^{-\beta H_X})$ and $\beta = 1/k_\mathrm{B}T$. An operation represented by the unitary $U$ is a free operation if

$$[H, U] = 0\,. \tag{1}$$

If the system is prepared in the state $\rho$, a *free evolution* (i.e. one that can be achieved without resources) is then of the form

$$\mathcal{E}[\rho] = \mathrm{tr}_R\big[U(\rho\otimes\tau_R)U^\dagger\big]\,. \tag{2}$$

The robustness of the framework under modifications of the assumptions has been the object of recent studies [6–9]. In this paper, we look at what is arguably the mildest form of deviation: an *inhomogeneous reservoir*. This is a reservoir made of a large number $N$ of systems, whose local parameters deviate randomly from those that would define an exact thermal operation. For this first study, we shall focus on inhomogeneities either in local temperature or in the local Hamiltonian.

## II. THE MODEL

### A. Introducing inhomogeneities

The system is a qudit, and the reservoir consists of $N$ qudits labelled by $r \in \{1, 2, ..., N\}$. We work with a Hamiltonian of non-interacting systems

$$H = H_S + H_R = g_0 s_z^{(S)} + \sum_{r=1}^{N} g_r s_z^{(r)} \tag{3}$$

where $g_0 > 0$, the $g_r$ will be discussed later, and $s_z$ is the operator representing the spin in the direction $z$. For every qudit, the eigenstates of $s_z$ for the eigenvalue $\hbar\big(j - \frac{d-1}{2}\big)$ is denoted by $|j\rangle$ with $j \in \{0, 1, ..., d-1\}$ — in particular, $|0\rangle$ is the ground state of $gs_z$ whenever $g > 0$.

For simplicity, throughout this work we consider input states of the system $\rho = \sum_j p_j |j\rangle\langle j|$ that are *diagonal* in the eigenbasis of $H_S$. The qudits of the reservoir are prepared in the thermal state at the local temperature: $\tau_R = \bigotimes_r \tau_r$ with $\tau_r = e^{-\beta_r g_r s_z}/Z_r$.

The inhomogeneous reservoir is described by a configuration $\underline{\delta}_N = (\delta_1, ..., \delta_N)$, where $\delta_r$ is the inhomogeneity perceived by the $r$-th qudit of the reservoir. As random variables, we assume that the $\delta_r$ are independent and identically distributed (i.i.d.) with a distribution $G(\delta)$ centered at $\overline{\delta} = 0$. We consider two cases: that of *inhomogeneous temperature* defined by

$$\beta_r = \beta(1 + \delta_r) \text{ and } g_r = g_0 \;\forall r\,; \tag{4}$$

and that of *inhomogeneous Hamiltonian* defined by

$$g_r = g_0(1 + \delta_r) \text{ and } \beta_r = \beta \;\forall r\,. \tag{5}$$

In the language of resource theories, (4) amounts at considering *resource states* arising from having multiple temperatures; while (5) allows for *resource operations* which violate (1). We also note that both inhomogeneities have a clear physical flavor. For instance, if the qudits are magnetic moments, conditions (5) may describe the inhomogeneity of the intensity of the local magnetic field, or of the gyromagnetic factor (e.g. through the chemical environment).

Either way, the thermal state $\tau_r$ of each reservoir qudit is

$$\tau(\delta_r) = \frac{e^{-\beta g_0\,(1+\delta_r)s_z}}{\mathrm{Tr}(e^{-\beta g_0\,(1+\delta_r)s_z})} = \sum_j q_j(\delta_r)\,|j\rangle\,\langle j| \quad (6)$$

where $q_j(\delta) = \frac{1-a(\delta)}{1-a(\delta)^d}\,a(\delta)^j$ with $a(\delta) = e^{-\beta\hbar g_0(1+\delta)}$. Clearly $\tau(\delta=0) = \frac{e^{-\beta g_0\,s_z}}{\mathrm{Tr}(e^{-\beta g_0\,s_z})} \equiv \tau_S$ the thermal state of the system for $\beta$.

### B. Interaction: collisional model

Now we have to discuss the interaction $U$. With the aim of bringing out local inhomogeneities, it is convenient to have the system interact sequentially with each reservoir qudit. In other words, $U = U_{S,N}U_{S,N-1}...U_{S,1}$ is going to be the product of successive two-body interactions, each between the system and one of the reservoir qudits. Such *collisional models* have been used as toy models in several studies of quantum dynamics and thermodynamics, see e.g. [9–12]. In this paper we assume that all two-body interactions $U_{S,r}$ are given by the partial swap with mixing angle $\theta$:

$$U = \cos\theta\,\mathbb{I} + i\sin\theta\,\mathcal{S} \quad (7)$$

with $\mathcal{S}$ the swap operator for 2 qudits. If $g_r = g_0$ for all $r$, then $U$ couples only degenerate eigenstates of $H$ and (1) holds. If $g_r \neq g_0$, $U$ can be taken effectively independent of the inhomogeneities provided $\sqrt{\overline{\delta^2}} \ll g_{int}/g_0 d$, where $g_{int}$ is the strength of the interaction (Appendix A 1).

### C. Dynamics of the system

In the absence of inhomogeneities ($\tau_r = \tau_S$ for all $r$), the dynamics (2) can be solved analytically for our model. For diagonal input states, the state of the system after interaction with the first $r$ qudits of the reservoir is given by (Appendix A 2)

$$\begin{aligned}\rho_{S|r} &= \rho_{S|r-1}\cos^2\theta + \tau_S\sin^2\theta \\ &= \tau_S - (\tau_S - \rho_0)\cos^{2r}\theta \quad [\underline{\delta}_N = 0]\,. \end{aligned} \quad (8)$$

In particular, the state remains diagonal and converges to the thermal state $\tau_S$ in the limit $N \to \infty$.

Each configuration $\underline{\delta}_N$ of the inhomogeneities induces a new map on the system. If the inhomogeneities are frozen, the dynamics (2) defines a contractive map whose fixed point $\rho_{S|\infty}$ is determined by the specific $\underline{\delta}_N$, and there is little more to say. The model is more interesting if $\underline{\delta}_N$ is drawn independently for each use of the channel: then we can study the *ensemble average* over $G(\delta)$. The dynamics commutes with this average: for i.i.d. inhomogeneities, the reservoir qudits are all prepared in the ensemble-averaged thermal state

$$\overline{\tau} = \int_{-\infty}^{\infty} G(\delta)\tau(\delta)\,\mathrm{d}\delta\,. \quad (9)$$

Thus the similarly defined ensemble-averaged state of the system at step $r$ is

$$\overline{\rho}_{S|r} = \overline{\tau} - (\overline{\tau} - \rho_0)\cos^{2r}\theta\,. \quad (10)$$

For qubits, $\overline{\tau}$ is more mixed that $\tau_S$ and can be seen as a thermal state for an effective temperature larger than $T$ [13]; for $d > 2$, $\overline{\tau}$ won't be thermal in general.

## III. WORK

The change of $\mathrm{Tr}[(H_S + H_R)\rho_{SR}]$ during the dynamics can be identified with *work* [14]. We focus on the case (5) of inhomogeneous Hamiltonian, because in the case (4) it holds $[H_S + H_R, U] = 0$ and no work is generated during any collision.

### A. Work generated in a single collision

We consider first the collision between the system and the $r$-th reservoir qudit. The work generated during this collision is

$$W_r = \mathrm{Tr}\left[\rho_{S|r-1}\otimes\tau(\delta_r)(U^\dagger H_{\delta_r}U - H_{\delta_r})\right] \quad (11)$$

where $H_\delta = g_0[s_z^{(S)} + (1+\delta)s_z^{(r)}]$. The calculation eventually yields (see Appendix B 1)

$$W_r = \delta_r\,\hbar g_0\sin^2\theta\sum_{j,j'}(j-j')\,p_j(\underline{\delta}_{r-1})\,q_{j'}(\delta_r) \quad (12)$$

As expected by the structure of the partial swap, only the transitions between levels with different values of $j$ generate work. For qubits, Eq. (12) becomes

$$W_r = \delta_r\,\hbar g_0\sin^2\theta\,[q_0(\delta_r) - p_0(\underline{\delta}_{r-1})] \quad (13)$$

since $p_0(\underline{\delta}_{r-1}) + p_1(\underline{\delta}_{r-1}) = 1$ and $q_0(\delta_r) + q_1(\delta_r) = 1$ for every $r$. Figure 1(a) shows how $W_r$ varies with

$\delta_r$ for various values of $p_0$. From (13) and the knowledge of $G(\delta)$, we can find the *statistical distribution* of single-collision work (see Appendix B 2).Figure 1(b) illustrates this distribution for a Gaussian distribution of inhomogeneities $G(\delta)$ and a few values of $p_0$. For $p_0 = 1/(1+a)$, that is for $\rho_{S|r-1} = \tau_S$, the distribution is the narrowest and diverges as $1/\sqrt{W}$ at $W = 0$.

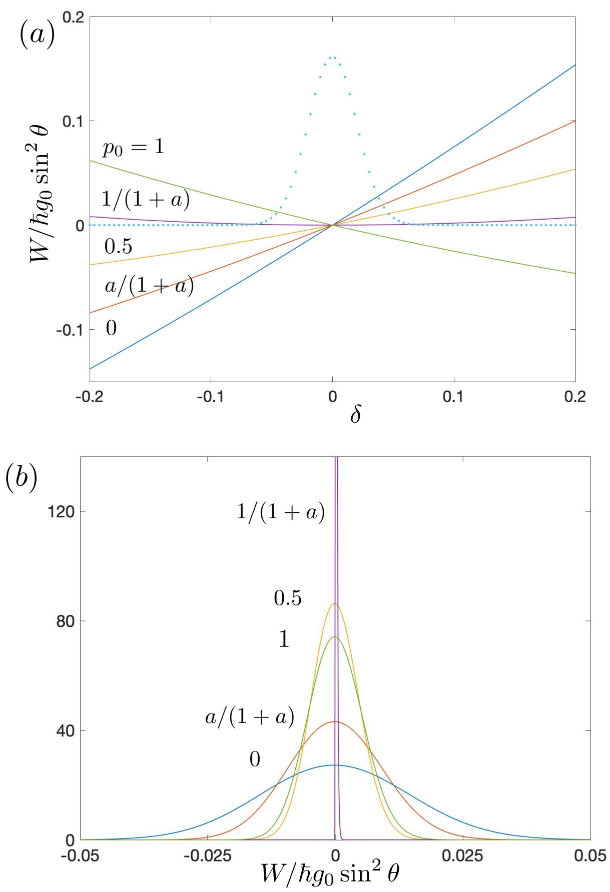


FIG. 1. Single-interaction work and its distribution for qubits, assuming a diagonal input state for the system. (a) The solid lines are single-interaction work $W/\hbar g_0 \sin^2\theta$ according to Eq. (13) as a function of the inhomogeneity $\delta$, for various values of $p_0$ (indicated next to each curve). The dotted line shows the distribution $G(\delta)$ used for the lower plot, multiplied by 2 for visibility. (b) Distribution of work $G_W(y)$, with $y = W/\hbar g_0 \sin^2\theta$, normalised according to $\int G_W(y)dy = 1$, for the same values of $p_0$. Both plots use $\beta\hbar g_0 = 1$, (whence $a = e^{-1}$ i.e. $\frac{1}{1+a} \approx 0.73$) and $\sqrt{\overline{\delta^2}} = 0.02$.

### B. Ensemble average of single-collision work

Now we compute the ensemble average $\overline{W_r}$ of (12). One could think that $[H_S + \overline{H}_R, U] = 0$ implies $\overline{W_r} = 0$. But this is not the case because the reservoir states also depend on $\underline{\delta}_N$. The actual expression is $\overline{W_r} = \hbar g_0 \sin^2\theta \sum_{j,j'}(j-j')\overline{p_j(\underline{\delta}_{r-1})}\;\overline{\delta_r\, q_{j'}(\delta_r)}$, having noticed that $\underline{\delta}_{r-1}$ and $\delta_r$ are not correlated. The value of $\overline{p_j(\underline{\delta}_{r-1})}$ can be read from (10).

To get a more compact expression, we use the Taylor expansion $q_{j'}(\delta) = q_{j'}(0) + q'_{j'}(0)\delta + O(\delta^2)$ to find

$$\overline{W_r} = \hbar g_0 \Sigma_r\, \overline{\delta^2}\, \sin^2\theta + O(\overline{\delta^3})\,, \qquad (14)$$

with $\Sigma_r \equiv \sum_{j,j'}(j-j')\overline{p_j(\underline{\delta}_{r-1})}q'_{j'}(0)$ and $q'_j(0) = -\frac{\beta\hbar g_0 a^j}{(1-a^d)^2}\left[j-(1+j)a+(d-j)a^d-(d-j-1)a^{d+1}\right]$ with $a = e^{-\beta\hbar g_0}$ [15]. For qubits, $q'_0(0) = -q'_1(0)$ and so $\Sigma_r = q'_0(0) = \beta\hbar g_0\frac{a}{(1+a)^2}$, independent of $r$ and positive.

### C. Accumulated work and dynamics

Assuming $\Sigma_r \approx \Sigma$ for all $r$ and independent of $\theta$, the work accumulated during the $N$ collisions is

$$\overline{W}_N = \sum_r \overline{W_r} \approx \hbar g_0 \Sigma\, \overline{\delta^2}\, N \sin^2\theta\,. \qquad (15)$$

This may be kept bounded for all $N$ by choosing a suitable scaling of $\theta$ with $N$. However, the value of $\theta$ affects also the dynamics (10): in particular,

$$\mathcal{D}(\overline{\rho}_{S|N}, \overline{\tau}) = \cos^{2N}\theta\, \mathcal{D}(\rho_0, \overline{\tau})\,. \qquad (16)$$

where $\mathcal{D}(\rho, \rho') = \frac{1}{2}\mathrm{Tr}(|\rho - \rho'|)$ is the trace distance.

Let's then look at the scaling $\sin^2\theta = cN^{-\xi}$. If $\xi > 1$, in the limit of large $N$ one has $\overline{W}_N \to 0$, but also $\mathcal{D}(\overline{\rho}_N, \overline{\tau}) \approx \mathcal{D}(\rho_0, \overline{\tau})$: no work is produced because the dynamics is frozen. If $\xi < 1$, then in the limit of large $N$ one has $\mathcal{D}(\overline{\rho}_N, \overline{\tau}) \to 0$ but $\overline{W}_N \to \infty$. A good compromise is

$$\sin^2\theta = \frac{c}{N} \Longrightarrow \begin{cases} \overline{W}_N \approx \hbar g_0 \Sigma\, c\, \overline{\delta^2} \\ \mathcal{D}(\overline{\rho}_N, \overline{\tau}) \approx e^{-c}\, \mathcal{D}(\rho_0, \overline{\tau}) \end{cases} \qquad (17)$$

The trace distance with the steady state decreases exponentially with $c$, while the total accumulated work increases linearly with $c$ but remains bounded.

## IV. THE "SECOND LAWS OF THERMAL OPERATIONS" AND INHOMOGENEOUS RESERVOIRS

The set of criteria under which a target state $\rho'$ can be obtained from $\rho$ by free evolution can be seen as the analog of the second law of thermodynamics. The transformation $\rho \longrightarrow \rho'$ under free operation does not define a total order: as a result, it cannot be characterised by a single criterion [4]. Brandao and coworkers [16] wrote the *second laws of thermal operations* as the monotonical decrease

$$\Delta F_\alpha = F_\alpha(\mathcal{E}[\rho]||\tau_S) - F_\alpha(\rho||\tau_S) \leq 0, \alpha \in \mathbb{R} \quad (18)$$

of a continuous family of *generalised free energies*

$$F_\alpha(\rho||\tau_S) = k_{\mathrm{B}}T\, [D_\alpha(\rho||\tau_S) - \log Z_S] \quad (19)$$

defined from the $\alpha$-Rényi divergence $D_\alpha(\rho||\tau_S)$. If $\rho$ and $\tau_S$ are diagonal in the same basis, as we are assuming since the beginning, it holds

$$D_\alpha(\rho||\tau_S) = \frac{\mathrm{sgn}(\alpha)}{\alpha - 1} \log \sum_j p_j^\alpha q_j^{1-\alpha} \quad (20)$$

with $q_j = e^{-\beta E_j}/Z_S$ the eigenvalues of $\tau_S$.

The conditions (18) are necessary and sufficient for free evolution. Since inhomogeneous reservoirs deviate from free dynamics, they should violate these conditions in some cases. The following protocol leads to a violation for *all* $\alpha$: prepare the system in the state $\tau_S$ and let it evolve to $\bar{\tau}$ according to (10). In this case, $\beta\Delta F_\alpha = D_\alpha(\bar{\rho}_{S|N}||\tau_S) - D_\alpha(\tau_S||\tau_S)$ is strictly positive, since $D_\alpha(\rho||\tau_S) \geq 0$ with equality if and only if $\rho = \tau_S$.

Updating the laws (18) to take into account any deviation from free evolution is an open challenge. Our study of inhomogeneous reservoirs may serve as starting point for this task. We first stress that, in our model, the possible state transformations are given by (10) for both inhomogeneous temperature and Hamiltonian. The generalised laws that single out these transformations must therefore be independent of the type of inhomogeneity [17].

However, their thermodynamical meaning will have to be different. When work is generated and $\beta$ is unique, thermodynamics requires $\Delta F_1 \leq W$, which was indeed proved for collisional models [12]. Our model of inhomogeneous Hamiltonian (5) shows that the generalisation $\Delta F_\alpha \leq W$ won't hold for $\alpha > 1$ [18], see Figure 2. In the case of inhomogeneous temperature (4), work is not generated; and in fact, in this narrative, the laws should not even involve free energies, since the second law of thermodynamics can be cast in terms of free energy only if the system is in contact with a bath at a single temperature. One could opt for reading (4) in the narrative of resource theories, where there is still a single reference temperature $\beta$, the $\tau(\delta_r)$ playing the role of non-thermal (i.e. resource) states. In this context, Ref. [7] defined approximate second laws with free energies $F_\alpha^\varepsilon$ where $\varepsilon$ is the maximal distance between a target state reachable with free operation and one reachable with the resource operation. In our case $\varepsilon = \mathcal{D}(\tau_S, \overline{\tau})$. For an analytical estimate for qubits, we compute the upper bound $\varepsilon \lesssim \overline{\mathcal{D}(\tau_S, \tau(\delta))} = \sqrt{2/\pi}\beta g_0 \hbar \frac{a}{(1+a)^2} \sqrt{\overline{\delta^2}} + O(\delta^2)$.

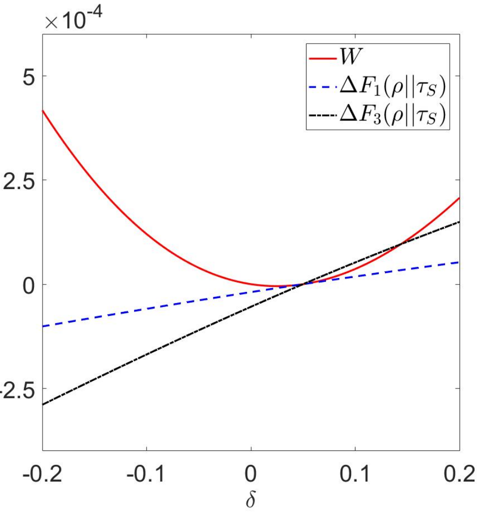


FIG. 2. Comparison of $W$ and $\Delta F_\alpha$ as a function of $\delta$, for $\beta\hbar g_0 = 1$, $\sqrt{\overline{\delta^2}} = 0.2$ and input state characterised by $p_0 = 0.75$. The expected violation of conditions (18) happens for $\delta > 0$. $\Delta F_1$ is upper bounded by $W$, as it should; but for larger values of $\alpha$, this upper bound is also violated (plotted for $\alpha = 3$).

## V. CONCLUSION

Extending the resource theory of thermal operations to non-ideal reservoirs is not trivial [6]. In this paper, we have introduced the notion of inhomogeneous reservoirs. Using the most standard collisional model, which fits well the definition of free dynamics in the absence of inhomogeneity, we have studied the two simplest cases of i.i.d. inhomogeneities: either in local temperature (which can be interpreted as having "resource states") or in the local Hamiltonian (which is an instance of "resource operations").

There are clearly many ways in which this study can be extended. Here we have restricted our attention to states of the system that are diagonal in the energy eigenbasis, and it would be worth considering general states of the systems and the role of coherence. Also, even staying within the family of collisional models, one can study different parameters. A standing open problem is the formulation of the rules for state transformation ("second laws") for inhomogeneous reservoirs: this paper has provided only an initial insight on this question.

## ACKNOWLEDGMENTS

We acknowledge illuminating discussions with and useful feedback from Philippe Faist, Rodrigo Gallego, Gabriel Landi, Kavan Modi, Markus Müller, Nelly Ng and Henrik Wilming.

This research is supported by the National Research Fund and the Ministry of Education, Singapore, under the Research Centres of Excellence programme.


---


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

[2] M. Horodecki, J. Oppenheim, and R. Horodecki, Phys. Rev. Lett. **89**, 240403 (2002).

[3] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Phys. Rev. Lett. **111**, 250404 (2013).

[4] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Y. Halpern, Phys. Rep. **583**, 1 (2015).

[5] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, J. Phys. A: Math. Theor. **49**, 143001 (2016).

[6] C. Sparaciari, D. Jennings, and J. Oppenheim, Nat. Commun. **8**, 1895 (2017).

[7] R. van der Meer, N. H. Y. Ng, and S. Wehner, Phys. Rev. A **96**, 062135 (2017).

[8] M. P. Müller, Phys. Rev. X **8**, 041051 (2018).

[9] E. Bäumer, M. Perarnau-Llobet, P. Kammerlander, and R. Renner, "Partial thermalizations allow for optimal thermodynamic processes," (2017), preprint arXiv:1712.07128.

[10] V. Scarani, M. Ziman, P. Štelmachovič, N. Gisin, and V. Bužek, Phys. Rev. Lett. **88**, 097905 (2002).

[11] S. Lorenzo, R. McCloskey, F. Ciccarello, M. Paternostro, and G. M. Palma, Phys. Rev. Lett. **115**, 120403 (2015).

[12] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, Phys. Rev. X **7**, 021003 (2017).

[13] For $d = 2$, the ground state occupation of $\tau_S$ is $q_0 = (1 + e^{-\beta\hbar g_0})^{-1}$; that of $\overline{\tau}$ is $\overline{q}_0 = q_0 \int G(\delta) \frac{1+e^{-\beta\hbar g_0}}{1+e^{-\beta\hbar g_0(1+\delta)}}\, d\delta$. If $G(\delta) = G(-\delta)$, then $\int G(\delta) f(\delta)\, d\delta = \int G(\delta) f_{\text{even}}(\delta)\, d\delta$ with $f_{\text{even}}(\delta) = \frac{1}{2}(f(\delta) + f(-\delta))$. For our case, given $e^{-\beta\hbar g_0} \leq 1$, it is elementary to prove that the maximum of $f_{\text{even}}(\delta)$ is achieved at $\delta = 0$ and is $f_{\text{even}}(0) = 1$; then the function decreases monotonically towards $\lim_{\delta\to\pm\infty} f_{\text{even}}(\delta) = \frac{1}{2}(1 + e^{-\beta\hbar g_0})$. Thus $\overline{q}_0 \leq q_0$, i.e. $\overline{\tau}$ is more mixed that $\tau_S$.

[14] When the Hamiltonian is time-dependent, work is usually defined as $W(t_2, t_1) = \int_{t_1}^{t_2} \text{Tr}(\rho(t)\dot{H}(t))dt$. In our model, a collision can be seen in this framework: the interaction Hamiltonian $H_{int}$ that generates $U$ is switched on abruptly at $t_1$, then switched off abruptly at $t_2 = t_1 + t_{int}$. Thus $\dot{H}_{int}(t) = H_{int}[\delta(t-t_1) - \delta(t-t_2)]$. Noticing that during the interaction $H_S + H_R + H_{int}$ is obviously conserved, we find indeed $W(t_2, t_1) = \text{Tr}[(\rho(t_2) - \rho(t_1))(H_S + H_R)]$.

[15] Notice that for a distribution such that $G(\delta) = G(-\delta)$ the expression (14) is exact up to order $O(\overline{\delta^4})$.

[16] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, Proc. Natl. Acad. Sci. **112**, 3275 (2015).

[17] The replacement of $\tau_S$ with $\overline{\tau}$ in (18) is a formal fix that ignores the physics of the problem. With such a fix, every contractive map would define a "second law", without any reference to thermodynamics.

[18] The bound $\Delta F_\alpha \leq W$ was obtained for all $\alpha$ in Section G.3 of the Supplementary Information of [16]. But their definition of work is different: they are looking at state transformations catalysed by a two-level battery (a "work bit") prepared in the thermal state, and $W$ is the value of the gap. In other words, $W$ is a parameter of the state, chosen so that the state transformation becomes possible, and is not related to the time-dependent dynamics (also, its value does not match the "change in energy" of the joint system).


## Appendix A: The partial swap

In this appendix, we study the partial swap. We write $|jk\rangle \equiv |j\rangle_S |k\rangle_r$ where $r$ is a reservoir qudit. The partial swap is

$$U = \cos\theta\, \mathbb{I} + i \sin\theta\, \mathcal{S} = \sum_{j,k=1}^{d} \big( \cos\theta\, |jk\rangle + i \sin\theta\, |kj\rangle \big) \langle jk| \tag{A1}$$

$$= e^{i\theta} \Big( \sum_j |jj\rangle \langle jj| + \sum_{k>j} \Big|\Psi^+_{jk}\Big\rangle \Big\langle \Psi^+_{jk}\Big| \Big) + e^{-i\theta} \sum_k \sum_{k>j} \Big|\Psi^-_{jk}\Big\rangle \Big\langle \Psi^-_{jk}\Big| = e^{i\theta}\Pi_{\text{sym}} + e^{-i\theta}\Pi_{\text{asym}} \tag{A2}$$

where $\Big|\Psi^\pm_{jk}\Big\rangle = \frac{1}{\sqrt{2}}(|jk\rangle \pm |kj\rangle)$, and $\Pi_{\text{sym/asym}}$ are the projectors on the syymetric and antisymmetric space respectively.

Positing $U = e^{-iH_{int}t_{int}/\hbar}$, the generating Hamiltonian is found to be

$$H_{int} = -\hbar g_{int}(\Pi_{\rm sym} - \Pi_{\rm asym}) \; = \; -\hbar g_{int} \sum_{j=1}^{d}\Big[|jj\rangle\langle jj| + \sum_{k>j}\left(|jk\rangle\langle kj| + |kj\rangle\langle jk|\right)\Big] \tag{A3}$$

and $\theta = g_{int}t_{int}$. This can be expressed in terms of spin operators only for $d = 2$, in which case it is

$$H_{int} = -\frac{\hbar g_{int}}{2}\,\vec{\sigma}^{\,(S)}\cdot\vec{\sigma}^{\,(r)} \tag{A4}$$

up to a term proportional to $\mathbb{I}$. This Hamiltonian is frequently found, but is usually not fundamental: rather, it is the result of a resonant (a.k.a. rotating-wave) approximation. Thus, our toy model may match realistic systems only if $g_{int} \ll g_0$.

### 1. Justification for the fixed $U$

In the main text, the evolution $U$ in the presence of interaction was taken to be the partial swap regardless of $H_S + H_R$. Strictly speaking, the evolution is

$$U_{\rm exact} = e^{-iH_{tot}t_{int}/\hbar} \tag{A5}$$

where $H_{tot}$ is the *total* Hamiltonian acting during the time $t_{int}$. In a toy model, one could get away by assuming that $H_S$ and $H_R$ are switched off during the interaction; but in this case the expression of work would not be the change of $\mathrm{Tr}[(H_S + H_R)\rho_{SR}]$, as studied later in the text. Rather, we work with $H_{tot} = H_S + H_R + H_{int}$. In this case, $U_{\rm exact} = U$ holds exactly only if $[H_S + H_R, H_{int}] = 0$. To show the consistency of our study, we need to find that there exists a range of parameters in which $U_{\rm exact} \approx U$ holds even if $[H_S + H_R, H_{int}] \neq 0$.

In our case of inhomogeneous Hamiltonian (5), the total Hamiltonian is

$$\begin{aligned} H_{tot} &= g_0 s_z \otimes \mathbb{I} + g_0(1+\delta)\mathbb{I}\otimes s_z + H_{int} \\ &= \hbar g_0 \sum_{jk}[j + (1+\delta)k]\,|jk\rangle\langle jk| + H_{int} - \hbar g_0\frac{d-1}{2}(2+\delta)\mathbb{I}\,. \end{aligned}$$

This $H_{tot}$ can be diagonalised exactly since it's a sum of $2\times 2$ blocks; but the result we are looking for can already be guessed (and is of course confirmed by looking at the exact expressions): the contribution of the inhomogeneity will be negligible if $g_0|\delta|\frac{d-1}{2} \ll g_{int}$, i.e. if

$$\sqrt{\delta^2} \ll \frac{2g_{int}}{g_0(d-1)}\,. \tag{A6}$$

In the main text we used the simplified expression $\sqrt{\delta^2} \ll g_{int}/g_0 d$. In other words, $g_{int}$ should be moderately large. Notice that this does not mean that $\theta$ should be large: since $\theta = g_{int}t_{int}$, we can always make $t_{int}$ small. This is the same limit used for the Trotter-Suzuki decomposition: for very short interaction times, the non-commutation of the terms of the Hamiltonian does not really matter.

### 2. Evolution under $U$

Given an input product state $\rho\otimes\tau$, the evolved state under $U$ is

$$U(\rho\otimes\tau)U^\dagger = c^2\rho\otimes\tau + s^2\tau\otimes\rho + ics[S,\rho\otimes\tau]\,. \tag{A7}$$

where with $c \equiv \cos\theta$ and $s \equiv \sin\theta$. For a generic state of the system and for the reservoir qudit in state $\tau = \sum_l q_l\,|l\rangle\langle l|$, the commutator reads explicitly

$$[S,\rho\otimes\tau] = \sum_{jj'l}\rho_{jj'}q_l\,\left(|lj\rangle\langle j'l| - |jl\rangle\langle lj'|\right)\,.$$

If we further assume the state of the system to be diagonal, i.e. $\rho_{jj'} = p_j \delta jj'$, we find

$$[S, \rho \otimes \tau] = \sum_{jl} (p_j q_l - p_l q_j) \, |lj\rangle \langle jl| \; .$$

Both partial traces of this term are zero. Therefore indeed, for diagonal input states of the system, $\rho_{S|r} = c^2 \rho_{S|r-1} + s^2 \tau$, as written in (8).

### Appendix B: Calculations for single-collision work

#### 1. Calculations leading to (12)

The work generated during this interaction is defined as

$$\begin{aligned} W_r &= \mathrm{tr}\left[(U\rho_{S|r-1} \otimes \tau(\delta) U^\dagger) H_\delta\right] - \mathrm{tr}\left[\rho_{S|r-1} \otimes \tau(\delta) H_\delta\right] \\ &= \mathrm{tr}\left[\rho_{S|r-1} \otimes \tau(\delta)(U^\dagger H_\delta U - H_\delta)\right] \; . \end{aligned} \tag{B1}$$

With $\eta = (d-1)/2$, we have the explicit expressions

$$H_\delta = g_0[s_z^{(S)} + (1+\delta) s_z^{(r)}] \; = \; \hbar g_0 \sum_{j,j'=1}^{d} \left[ j + j' - 2\eta + \delta(j' - \eta) \right] |jj'\rangle \langle jj'| \; .$$

Therefore, again with $c \equiv \cos\theta$ and $s \equiv \sin\theta$:

$$\begin{aligned} U^\dagger H_\delta U - H_\delta \; &= \; \sin^2\theta (S H_\delta S - H_\delta) + i \, \sin\theta\cos\theta [H_\delta, S] \\ = \delta \, \hbar g_0 \; &\textstyle\sum_{j,j'} (j - j') \left[\sin^2\theta \, |jj'\rangle \langle jj'| + i \, \sin\theta\cos\theta \, |j'j\rangle \langle jj'|\right] \end{aligned} \tag{B2}$$

If we assume that the initial state of the system is diagonal, we have

$$\rho_{S|r-1} \otimes \tau(\delta) = \sum_{jj'} p_j(\underline{\delta}_{r-1}) q_{j'}(\delta) \, |jj'\rangle \langle jj'| \; . \tag{B3}$$

Inserting (B2) and (B3) into (B1), one obtains the expression for work (12) given in the main text.

#### 2. Statistical distribution of single-collision work

We derive the statistics of work for qubits prepared in a diagonal state, induced by the inhomogeneity (5) of the Hamiltonian. Let us start by rewriting (13) as

$$y = \delta \left[ q_0(\delta) - p_0 \right] \tag{B4}$$

with $y \equiv \frac{W}{\hbar g_0 \sin^2\theta}$. We need to invert this function in order to find the distribution of work $G_W(y)$ induced by the distribution of the inhomogeneity $G(\delta)$. It's clear that the function cannot be inverted analytically. However, we can resort to the Taylor expansion $q_0(\delta) = q_0(0) + q_0'(0)\delta + O(\delta^2)$ where

$$q_0(0) = \frac{1}{1+a} \, , \;\; q_0'(0) = \beta\hbar g_0 \tfrac{a}{(1+a)^2} \, , \tag{B5}$$

with $a = e^{-\beta\hbar g_0}$. This reduces (B4) to a quadratic equation whose solutions are

$$\delta_\pm(y) = \frac{p_0 a - p_1 \pm \sqrt{(p_0 a - p_1)^2 + 4\beta\hbar g_0 \, a \, y}}{2\beta\hbar g_0 \, a/(1+a)} \tag{B6}$$

We notice that the thermal state ($p_1/p_0 = a$) has a special role and that, for any choice of parameters, there is a lower bound for work:

$$W \geq -\frac{1}{\beta} \sin^2\theta \, \frac{(p_0 a - p_1)^2}{4a} \tag{B7}$$

The distribution of $W$ above this lower bound is given by

$$\begin{aligned} G_W(y) &= \sum_{s=\pm} G(x_s(y)) \left| \frac{d\delta_s}{dy'}(y'=y) \right| \\ &= \frac{1+a}{\sqrt{(p_0 a - p_1)^2 + 4\beta\hbar g_0\, a\, y}} \sum_{s=\pm} G(\delta_s(y))\,. \end{aligned} \tag{B8}$$

This is the distribution plotted in Fig. 1 (b).