

# SCIENTIFIC REPORTS



OPEN

## Correlations in quantum thermodynamics: Heat, work, and entropy production

S. Alipour<sup>1,2</sup>, F. Benatti<sup>3,4</sup>, F. Bakhshinezhad<sup>2,\*</sup>, M. Afsary<sup>2,\*</sup>, S. Marcantoni<sup>3,4,\*</sup> & A. T. Rezakhani<sup>2</sup>

Received: 16 August 2016

Accepted: 03 October 2016

Published: 21 October 2016

**We provide a characterization of energy in the form of exchanged heat and work between two interacting constituents of a closed, bipartite, correlated quantum system. By defining a binding energy we derive a consistent quantum formulation of the first law of thermodynamics, in which the role of correlations becomes evident, and this formulation reduces to the standard classical picture in relevant systems. We next discuss the emergence of the second law of thermodynamics under certain—but fairly general—conditions such as the Markovian assumption. We illustrate the role of correlations and interactions in thermodynamics through two examples.**

Investigating the consistency of thermodynamics, as a successful classical theory of macroscopic physical systems, with quantum mechanics, as a fundamental theory of the underlying microscopic systems, is still an open problem and the subject of extensive recent research<sup>1</sup>. Defining quantum mechanical counterparts of the classical concepts of “heat” and “work” and describing the mechanisms underlying their exchange among microscopic quantum systems are the first steps towards this goal. From this point of view, studying the thermodynamic role of quantum mechanical features such as nonclassical correlations is of paramount importance.

There exist various approaches to defining heat and work *microscopically*. One widely-used definition has been proposed in ref. 2, where work exchange is due to an external periodic driving incorporated in a time-dependent Hamiltonian, while heat is absorbed or released because of interaction with an ambient environment. In order to show that also in time-independent Hamiltonian systems work can be extracted, other approaches have been proposed, such as using another quantum system as a “work storage”<sup>3</sup>. These proposed mechanisms, however, are not entirely consistent with each other. Indeed, the role of correlations is usually not considered in these approaches since the focus is mainly on one of the two parties of compound quantum systems, whereas a comprehensive study of thermodynamic behaviors should consider all effects concerning systems involved in nonequilibrium thermodynamic processes. Besides, a consistent definition of heat has not yet been known when a system is coupled strongly to its environment<sup>4</sup>.

Here we revisit the first and the second laws of thermodynamics in a closed (conservative) quantum system comprising two interacting parties, one denoted by  $S$  as the thermodynamic system of interest and the other denoted by  $B$  playing the role of its environment, and reformulate these laws in a way that clearly exposes the role played by  $SB$  correlations. In order to do so we introduce the notion of “binding energy” of two interacting quantum systems which, together with the internal energies of the two parties, provides the internal energy of the whole system.

The definitions of heat and work that we use are similar to those in ref. 2; however, unlike there, we show that, in general, nonequilibrium thermodynamic processes affecting a system  $S$  involve work exchange with  $B$  without the need for an external driving represented by a time-dependent parameter in the system Hamiltonian, but merely because of the interactions between  $S$  and  $B$ . Besides, we explicitly show that correlations do not play any role in work exchange, while they do play an unavoidable role in heat transfer between  $S$  and  $B$ . Furthermore, this latter process does not necessarily need  $B$  to be treated as an environment weakly coupled to  $S$ , thus  $S$  need not be expected to evolve in time according to a dissipative Markovian (Lindbladian) dynamics.

<sup>1</sup>School of Nano Science, Institute for Research in Fundamental Sciences (IPM), Tehran 19538, Iran. <sup>2</sup>Department of Physics, Sharif University of Technology, Tehran 14588, Iran. <sup>3</sup>Department of Physics, University of Trieste, I-34151 Trieste, Italy. <sup>4</sup>National Institute for Nuclear Physics (INFN), Trieste Section, I-34151 Trieste, Italy. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.T.R. (email: rezakhani@sharif.edu)

As a preliminary and necessary step toward investigating heat and work exchanges between two interacting systems  $S$  and  $B$ , one needs to unambiguously assign to the two parties a percentage of the interaction energy depending on the state of the compound system. However, due to  $SB$  correlations, there will always be part of the interaction energy that belongs to both  $S$  and  $B$  together. In thermodynamic terms, extracting this part of the energy would require accessibility of the total system. Thus, we distinguish three contributions to the total internal energy of  $SB$ : one accessible only through  $S$ , the other one only through  $B$ , and the last one only through  $SB$  (as a whole) via the  $SB$  correlations. We call this latter contribution to the internal energy the *binding energy*. Certainly, although (in the case of time-independent total Hamiltonian) the total internal energy remains constant in time, that of either  $S$  or  $B$  varies because they interact and thus exchange work and heat.

In a recent publication<sup>5</sup>, the internal energy of an open quantum system has been defined as the energy which is accessible through *measurements* in a fixed “local effective measurement basis”<sup>6</sup>, and the definitions of work and heat suggested by considering the ability of the energy changes in altering the von Neumann entropy; heat is the energy flux that may change the entropy whereas work is the part of the energy change that keeps entropy intact. In contrast, in our formalism the internal energy associated with each subsystem is defined as the energy which is locally accessible in each individual subsystem by means of arbitrary local measurements. Although it has been known in the literature that correlations play a role in heat exchange, this fact has not been shown explicitly thus far. In the following, we provide explicit relations that exhibit the role played by correlations in heat, work, and entropy exchange between constituents of a bipartite system.

The structure of this paper is as follows. First we lay out the framework to define basic thermodynamic properties such as heat and work, and show a first law governing their mutual transformations. The next section deals with finite and infinitesimal versions of a possible formulation of the second law of thermodynamics. We illustrate our formalism through two examples. The paper is concluded by a summary.

### First law of thermodynamics in the presence of interactions and correlations

We consider a closed quantum system  $SB$  consisting of two interacting quantum systems: the system of interest  $S$  and its *bath* or *environment*  $B$ —with no restrictions on the dimensionality of  $S$  and  $B$ . The state of  $SB$  is described by the density matrix  $\rho_{SB}(\tau)$ , evolving under a total time-independent Hamiltonian

$$H_{\text{tot}} = H_S + H_B + H_{\text{int}}. \quad (1)$$

The *internal energy* of the total system is the mean value of the total Hamiltonian with respect to the time-evolving state, namely  $\mathbb{U}_{\text{tot}} = \text{Tr}[\rho_{SB}(\tau)H_{\text{tot}}]$ , and is thus constant in time since the dynamics of the total system is governed by the Schrödinger equation

$$d\rho_{SB}(\tau) = -i[H_{\text{tot}}, \rho_{SB}(\tau)] d\tau. \quad (2)$$

We assume  $\hbar \equiv 1$  throughout the paper.

In thermodynamics, infinitesimal variations of the internal energy of a system occur because of infinitesimal exchanges of *heat*  $d\mathbb{Q}$  and/or *work*  $d\mathbb{W}$  between the system and the environment. The quantum mechanical counterparts of infinitesimal heat and work exchanges in a system with state  $\rho(\tau)$  and time-dependent Hamiltonian  $H(\tau)$  are given by<sup>7</sup>

$$d\mathbb{Q}(\tau) = \text{Tr}[d\rho(\tau)H(\tau)], \quad (3)$$

$$d\mathbb{W}(\tau) = \text{Tr}[\rho(\tau)dH(\tau)], \quad (4)$$

where “ $d$ ” denotes a time differential, while  $d\mathbb{Q}$  and  $d\mathbb{W}$  are in general inexact differentials. With these definitions, we have the following quantum version of the first law of thermodynamics for the internal energy  $\mathbb{U}(\tau) = \text{Tr}[H(\tau)\rho(\tau)]$ :

$$d\mathbb{U}(\tau) = d\mathbb{Q}(\tau) + d\mathbb{W}(\tau). \quad (5)$$

In the case of a compound, isolated system  $SB$ , the states of the constituent subsystems are obtained by partial tracing  $\rho_{S,B}(\tau) = \text{Tr}_{B,S}[\rho_{SB}(\tau)]$ ; thus from equation (2) we have

$$d\rho_S(\tau) = -i[H_S, \rho_S(\tau)]d\tau - i \text{Tr}_B[H_{\text{int}}, \rho_{SB}(\tau)]d\tau, \quad (6)$$

$$d\rho_B(\tau) = -i[H_B, \rho_B(\tau)]d\tau - i \text{Tr}_S[H_{\text{int}}, \rho_{SB}(\tau)]d\tau. \quad (7)$$

When *there are correlations* between the system and the environment, we can write

$$\rho_{SB}(\tau) = \rho_S(\tau) \otimes \rho_B(\tau) + \chi(\tau), \quad (8)$$

where  $\chi$  measures all correlations (classical or quantum). Replacing this decomposition into equations (6) and (7) yields

$$d\rho_S(\tau) = -i[H'_S(\tau), \rho_S(\tau)]d\tau - i \text{Tr}_B[H_{\text{int}}, \chi(\tau)]d\tau, \quad (9)$$

$$d\rho_B(\tau) = -i[H'_B(\tau), \rho_B(\tau)]d\tau - i \text{Tr}_S[H_{\text{int}}, \chi(\tau)]d\tau, \quad (10)$$

where we have modified the Hamiltonians as

$$H'_{S,B}(\tau) = H_{S,B} + \text{Tr}_{B,S}[\rho_{B,S}(\tau)H_{\text{int}}]. \quad (11)$$

The last term might have contributions to *Lamb shift* corrections in Markovian dynamics<sup>8</sup>. We can rewrite equation (1) as

$$H_{\text{tot}} = H'_S(\tau) + H'_B(\tau) + H'_{\text{int}}(\tau), \quad (12)$$

where

$$H'_{\text{int}}(\tau) = H_{\text{int}} - \text{Tr}_S[\rho_S(\tau)H_{\text{int}}] - \text{Tr}_B[\rho_B(\tau)H_{\text{int}}]. \quad (13)$$

Unlike  $H'_S(\tau)$  and  $H'_B(\tau)$ , the interaction Hamiltonian  $H'_{\text{int}}(\tau)$  is a nonlocal operator, that is, it cannot be assigned exclusively to either of subsystems  $S$  and  $B$  alone. It thus may seem physically reasonable (and rewarding) to modify this Hamiltonian such that it becomes locally inaccessible by  $S$  and  $B$ . An evident advantage of this redefinition is that it allows us to divide the total internal energy of a composite system into local parts, which are associated to and accessible by each system, and a nonlocal part, which is attributed only to these systems as a whole—whereas it does not contribute to the (local) energies of each system.

The local inaccessibility requirement implies redefinition of  $H'_{\text{int}}(\tau)$  (of course consistently) such that the mean value of  $H'_{\text{int}}(\tau)$  with respect to the local states vanishes:  $\text{Tr}_i[\rho_i(\tau)H'_{\text{int}}(\tau)] = 0$ ,  $i \in \{S, B\}$ . However, with definition (13) one can see that

$$\begin{aligned} \text{Tr}_S[\rho_S(\tau)H'_{\text{int}}(\tau)] &= \text{Tr}_B[\rho_B(\tau)H'_{\text{int}}(\tau)] \\ &= -\text{Tr}[\rho_S(\tau) \otimes \rho_B(\tau)H_{\text{int}}], \end{aligned} \quad (14)$$

which is in general a nonzero *scalar*. One may remedy this in various ways. For example, a straightforward fix is to compensate for the nonzero scalar contribution of equation (14) by distributing it over the system and environment Hamiltonians through the real (but not necessarily positive) auxiliary parameters  $\alpha_S$  and  $\alpha_B = 1 - \alpha_S$  and hence defining the *effective* Hamiltonians

$$H_S^{(\text{eff})}(\tau) = H'_S(\tau) - \alpha_S \text{Tr}[\rho_S(\tau) \otimes \rho_B(\tau)H_{\text{int}}], \quad (15)$$

$$H_B^{(\text{eff})}(\tau) = H'_B(\tau) - \alpha_B \text{Tr}[\rho_S(\tau) \otimes \rho_B(\tau)H_{\text{int}}]. \quad (16)$$

Accordingly,  $H'_{\text{int}}(\tau)$  in equation (12) can be replaced with an effective interaction Hamiltonian

$$H_{\text{int}}^{(\text{eff})}(\tau) = H_{\text{tot}} - H_S^{(\text{eff})}(\tau) - H_B^{(\text{eff})}(\tau). \quad (17)$$

Note that the dynamical equations (9) and (10) remain valid where  $H'_{S,B}(\tau)$  are now replaced with  $H_{S,B}^{(\text{eff})}(\tau)$ , and we have

$$\text{Tr}_S[\rho_S(\tau)H_{\text{int}}^{(\text{eff})}(\tau)] = \text{Tr}_B[\rho_B(\tau)H_{\text{int}}^{(\text{eff})}(\tau)] = 0. \quad (18)$$

**Remark 1.** Note that the parameter  $\alpha_S$  (or identically,  $\alpha_B$ ) is still free. In general, there is no model-independent condition that fixes this value. We shall discuss later how this free parameter enters our thermodynamic expressions, and in the examples comment on how one may fix the value of this parameter in special cases and under some specific conditions. This parameter seems like another thermodynamic degree of freedom, whose role should be decided on the basis of the specific physical conditions of the nonequilibrium thermodynamic systems in question.

By defining the *internal energies* of the constituent systems through the effective Hamiltonians

$$\mathbb{U}_{S,B}(\tau) = \text{Tr}[\rho_{S,B}(\tau)H_{S,B}^{(\text{eff})}(\tau)], \quad (19)$$

an energy contribution remains, called *binding energy*, which can be naturally attributed to correlations  $\chi$  as

$$\mathbb{U}_\chi(\tau) = \text{Tr}[\chi(\tau)H_{\text{int}}^{(\text{eff})}(\tau)], \quad (20)$$

such that

$$\mathbb{U}_{\text{tot}} = \mathbb{U}_S(\tau) + \mathbb{U}_B(\tau) + \mathbb{U}_\chi(\tau). \quad (21)$$

Note that only if the interaction and correlations between the two systems were negligible (which is usually assumed in classical thermodynamics)—namely only if  $H_{\text{int}} \approx 0$  and  $\rho_{SB}(\tau) \approx \rho_S(\tau) \otimes \rho_B(\tau)$  at all times  $\tau$ —would have  $\mathbb{U}_{\text{tot}} \approx \mathbb{U}_S(\tau) + \mathbb{U}_B(\tau)$  and the internal energy be additive.

### Exchange of heat and work between two interacting systems

When two thermodynamical systems are independent, namely uncorrelated and noninteracting, the amount of the transferred heat/work from one system is equivalent to the heat/work received by the other system. In our context, the existence of interactions and correlations between  $S$  and  $B$  alters this picture for heat exchange, but

not for the exchanged work. Indeed, inserting equation (15) into equation (4), the infinitesimal works performed by  $S$  and  $B$  are obtained as

$$d\mathbb{W}_S(\tau) = \alpha_B \text{Tr}[\rho_S(\tau) \otimes d\rho_B(\tau) H_{\text{int}}] - \alpha_S \text{Tr}[d\rho_S(\tau) \otimes \rho_B(\tau) H_{\text{int}}] = -d\mathbb{W}_B(\tau). \quad (22)$$

The work absorbed or released by system  $S$ , respectively  $B$ , each depends on the scalars  $\alpha_{S,B}$ , which is reminiscent to the non-gauge-invariance feature of work<sup>9</sup>. Nevertheless, we always have

$$d\mathbb{W}_S(\tau) + d\mathbb{W}_B(\tau) = 0, \quad (23)$$

which is independent of  $\alpha_{S,B}$ , and looks like a form of the familiar classical law of action-reaction. Note that, unlike in refs 2 and 10 where it is the time dependence of the Hamiltonian through an external driving parameter that leads to work exchange, here the work exchange follows from the time dependence of the effective Hamiltonians which include time-dependent Lamb-shift-like corrections in equation (17). As a consequence, our formalism features that work exchange between two interacting subsystems is allowed even without an external driving.

Unlike the infinitesimal work  $d\mathbb{W}_S(\tau)$ , the scalar parameters  $\alpha_{S,B}$  do not contribute to the infinitesimal heat exchanges because  $\text{Tr}[d\rho(\tau)] = 0$ . In particular, inserting equations (15) and (16) into equation (3) yields

$$d\mathbb{Q}_S(\tau) = -i \text{Tr}[\chi(\tau)[H_S^{(\text{eff})}(\tau), H_{\text{int}}]] d\tau, \quad (24)$$

$$d\mathbb{Q}_B(\tau) = -i \text{Tr}[\chi(\tau)[H_B^{(\text{eff})}(\tau), H_{\text{int}}]] d\tau. \quad (25)$$

In addition, differentiating equation (21) and using equation (5) together with equation (23), one obtains that

$$d\mathbb{Q}_S(\tau) + d\mathbb{Q}_B(\tau) = -d\mathbb{U}_\chi(\tau). \quad (26)$$

Hence, in our setting the binding energy is completely of the heat type. This can also be seen from the relations  $d\mathbb{U}_\chi(\tau) = \text{Tr}[d\chi(\tau) H_{\text{int}}^{(\text{eff})}(\tau)]$  and  $\text{Tr}[\chi(\tau) dH_{\text{int}}^{(\text{eff})}(\tau)] = 0$ , which show that changes in the binding energy can only come from changes in the state correlations. Indeed, the heat balance equation (26) shows that (i) heat transfer is only due to interactions and correlations within the total system, in agreement with the result of refs 6 and 11, and (ii) that heat passing from one system to the other is paid for by varying the  $SB$  correlations that thus behave like a *heat storage*. That is, if correlations do not change and  $d\chi(\tau) = 0$ , then  $d\mathbb{U}_\chi = 0$ ; whence  $d\mathbb{Q}_S = -d\mathbb{Q}_B$ , in agreement with the standard textbook definition of “heat” in classical systems wherein no or only a negligibly-weak interaction between the system and the environment is assumed<sup>12</sup>.

**Remark 2.** In our derivations thus far we have assumed that the Hamiltonian of the total system  $SB$  is time-independent. If we relax this condition and allow a time dependence (e.g., due to the action of some external agent on the total system), part of our relations will be modified as follows:

$$d\mathbb{U}_{\text{tot}}(\tau) = d\mathbb{U}_S(\tau) + d\mathbb{U}_B(\tau) + d\mathbb{U}_\chi(\tau), \quad (27)$$

$$d\mathbb{Q}_S(\tau) + d\mathbb{Q}_B(\tau) = -d\mathbb{Q}_\chi(\tau), \quad (28)$$

$$d\mathbb{W}_S(\tau) + d\mathbb{W}_B(\tau) = \text{Tr}[\rho_S(\tau) \otimes \rho_B(\tau) dH_{\text{int}}(\tau)], \quad (29)$$

where  $d\mathbb{Q}_\chi(\tau) = \text{Tr}[d\chi(\tau) H_{\text{int}}(\tau)]$ . Equation (29) indicates that even in the time-dependent case correlations do not contribute in the exchange of work between the system and the environment. Since the dynamics of the compound system  $SB$  is generated by  $H_{\text{tot}}(\tau)$ , it follows that  $\text{Tr}[d\rho_{SB}(\tau) H_{\text{tot}}(\tau)] = 0$  (i.e., the total system is thermally isolated), so that there are no heat exchanges and the only possibility for  $SB$  is to perform work because of the external driving due to the rest of the universe.

## Second law of thermodynamics

According to the second law of thermodynamics, the entropy of a macroscopic closed system which is *thermally isolated* (in thermodynamics terminology) can only remain constant or increase in time<sup>12,13</sup>. However, the second law is not necessarily valid in nonequilibrium microscopic or even macroscopic systems<sup>14–19</sup>.

In the following we demonstrate the possible emergence of the second law of thermodynamics and the important role of system-bath correlations in this microscopic context.

In the case of a compound system  $SB$ , the subadditivity of the von Neumann entropy<sup>20</sup> (we set  $\kappa_B \equiv 1$  for the Boltzmann constant throughout the paper)

$$\mathbb{S}(\tau) = -\text{Tr}[\rho(\tau) \log \rho(\tau)], \quad (30)$$

implies that the *mutual information*

$$\mathbb{S}_\chi(\tau) = \mathbb{S}_S(\tau) + \mathbb{S}_B(\tau) - \mathbb{S}_{SB}(\tau) \quad (31)$$

is always nonnegative. Mutual information characterizes the amount of total correlations (both classical and quantum) shared by the two subsystems  $S$  and  $B$ <sup>21,22</sup>. Intuitively, if the correlations between  $S$  and  $B$  increases,  $\mathbb{S}_\chi$  becomes larger.

Since we have assumed that the total system  $SB$  is closed, it evolves unitarily and its von Neumann entropy  $\mathbb{S}_{SB}(\tau)$  does not change in time (even if its Hamiltonian depends on time). Hence, differentiating equation (31) yields

$$d\mathbb{S}_S(\tau) + d\mathbb{S}_B(\tau) = d\mathbb{S}_\chi(\tau). \quad (32)$$

Integrating both sides of this equation in the time interval  $[0, \tau]$ , with the assumption that the initial state of  $SB$  is uncorrelated (i.e.,  $\mathbb{S}_\chi(0) = 0$ ), leads to

$$\Delta\mathbb{S}_S + \Delta\mathbb{S}_B = \mathbb{S}_\chi(\tau) \geq 0, \quad (33)$$

as obtained in ref. 23. This relation states that, as long as one observes subsystems  $S$  and  $B$  locally and their initial state is without any correlations, the sum of the total variations of the entropies of  $S$  and  $B$  is always nonnegative. One can consider this property as a form of the second law of thermodynamics for the compound system  $SB$ .

Unlike in equilibrium thermodynamics, in a general nonequilibrium system “temperature” is not a well-defined quantity (see, e.g., refs 24 and 25 for some recent discussions). However, at fixed “volume” ( $V$ ) and “number of particles” ( $N$ ), one can introduce a time-dependent *pseudo-temperature* by means of the internal energy and the von Neumann entropy through

$$\frac{1}{T(\tau)} := \frac{d\mathbb{S}(\tau)}{dU(\tau)}, \quad (34)$$

which is somewhat reminiscent of the standard, equilibrium definition  $1/T = (\partial\mathbb{S}/\partial U)_{N,V}$ .

**Remark 3.** *In generic quantum systems, it is not always clear how to define  $V$  and  $N$  (or other relevant thermodynamic properties). Additionally, in thermodynamic equilibrium we deal with the partial derivative  $(\partial\mathbb{S}/\partial U)_{N,V}$  rather than the ratio of two total derivatives  $((d\mathbb{S}/d\tau)/(dU/d\tau))$ , which can be different quantities. Noting equations (15) and (16), the free parameter  $\alpha_S$  (and  $\alpha_B$ ) would also appear in the pseudo-temperature. In general then, one should not expect that the pseudo-temperature necessarily have definite relation with the equilibrium temperature, unless under certain conditions. Later in the examples we show explicitly how in special cases the pseudo-temperature may relate to the equilibrium temperature by appropriately fixing the scalar  $\alpha_S$  through thermodynamic properties of the system in question.*

Adopting the concept of pseudo-temperature, one can associate (time-dependent) pseudo-temperatures  $T_{S,B}(\tau)$  with subsystems  $S$  and  $B$  and also a pseudo-temperature  $T_\chi(\tau)$  with the binding energy. As a result, inserting equation (34) into equation (32) gives

$$\frac{dU_S(\tau)}{T_S(\tau)} + \frac{dU_B(\tau)}{T_B(\tau)} = \frac{dU_\chi(\tau)}{T_\chi(\tau)}. \quad (35)$$

It then follows that, when  $dU_\chi(\tau) = 0$  but  $T_\chi(\tau) \neq 0$ , since  $dU_S(\tau) = -dU_B(\tau)$  [equation (21)], the two subsystems must have the same instantaneous pseudo-temperature:  $T_S(\tau) = T_B(\tau)$ . Another possibility is when  $dU_\chi(\tau) = 0$  and  $T_\chi(\tau) = 0$  such that  $d\mathbb{S}_\chi(\tau) \neq 0$ , whence

$$dU_B(\tau) = \frac{T_S(\tau)T_B(\tau)}{T_S(\tau) - T_B(\tau)} d\mathbb{S}_\chi(\tau). \quad (36)$$

We remark that a somewhat similar result in ref. 26 is akin to our general expression in equation (35) (of course here with pseudo-temperature instead of equilibrium temperature). This interesting relation yet again indicates the role of correlations; for energy transport to the bath (and similarly to the system) correlations are necessary, where in turn development of correlations ensues from interaction.

**Remark 4.** *Under the same conditions, when the total Hamiltonian is time-dependent, equation (36) is modified to*

$$dU_B(\tau) = \frac{T_S(\tau)T_B(\tau)}{T_S(\tau) - T_B(\tau)} \left( d\mathbb{S}_\chi(\tau) - \frac{dU_{\text{tot}}(\tau)}{T_S(\tau)} \right). \quad (37)$$

Furthermore, from equations (5) and (34), it follows that

$$d\mathbb{S}_{S,B}(\tau) = \frac{dQ_{S,B}(\tau)}{T_{S,B}(\tau)} + \frac{dW_{S,B}(\tau)}{T_{S,B}(\tau)}. \quad (38)$$

Formally, the difference between the total variations of the entropies and the contributions coming from the heat exchanges are given by the contributions due to the work exchanges,

$$d\Sigma_{S,B}(\tau) := d\mathbb{S}_{S,B}(\tau) - \frac{dQ_{S,B}(\tau)}{T_{S,B}(\tau)} = \frac{dW_{S,B}(\tau)}{T_{S,B}(\tau)}. \quad (39)$$

The quantity  $d\tilde{\Sigma}_{S,B}$  resembles the infinitesimal *internal entropy production* as defined in ref. 2, where the case of an externally driven system  $S$  has been discussed which is weakly coupled to a conservative heat bath  $B$  inducing a dissipative dynamics<sup>8,27–29</sup> (in general explicitly time-dependent) of the Lindblad type. In this particular context, the infinitesimal entropy production is modified as the difference between the variation of the entropy  $\mathbb{S}_S(\tau)$  and the entropy flux into or out of the system associated to the heat flux  $d\mathbb{Q}_S(\tau)$  divided by the (initial) temperature of the bath  $T$  (rather than  $T_S(\tau)$  as in equation (38)),

$$d\Sigma_S(\tau) := d\mathbb{S}_S(\tau) - \frac{d\mathbb{Q}_S(\tau)}{T}. \tag{40}$$

This expression can be interpreted as an internal entropy production for system  $S$  and its nonnegativity,  $d\Sigma(\tau) \geq 0$ , can be considered as an infinitesimal expression of the second law of thermodynamics. If the time evolution  $\rho_S(0) \mapsto \rho_S(\tau)$  is given by a Lindblad-type generator  $\mathcal{L}_\tau$  that preserves the instantaneous Gibbs state  $\rho_S^\beta(\tau) = \exp(-\beta H_S(\tau))/Z_\beta(\tau)$ , with  $\beta = 1/T$  and  $Z_\beta(\tau) = \text{Tr}[e^{-\beta H_S(\tau)}]$ , namely  $\mathcal{L}_\tau[\rho_S^\beta(\tau)] = 0$ , then one can recast the infinitesimal entropy production (40) as

$$d\Sigma_S(\tau) = -\text{Tr}[\mathcal{L}_\tau[\rho_S(\tau)](\log \rho_S(\tau) - \log \rho_S^\beta(\tau))]d\tau. \tag{41}$$

If the generator  $\mathcal{L}_\tau$  is of the Lindblad form, for each fixed  $\tau \geq 0$ , the maps  $E_s = e^{s\mathcal{L}_\tau}$  form—with respect to the nonnegative parameter  $s$ —a semigroup of completely-positive and trace preserving family of maps<sup>20</sup>. Since  $E_s[\rho_S^\beta(\tau)] = \rho_S^\beta(\tau)$  and the relative entropy  $D(E_s[\rho_S(\tau)]\|\rho_S^\beta(\tau))$  is a monotonically decreasing function of  $s$ <sup>30,31</sup>, one obtains the following infinitesimal quantum version of the second law of thermodynamics:

$$d\Sigma_S(\tau) = -\frac{d}{ds}D(E_s[\rho_S(\tau)]\|\rho_S^\beta(\tau))\Big|_{s=0} d\tau \geq 0. \tag{42}$$

**Remark 5.** A simpler physical context is provided when there is no external driving for system  $S$ , namely its Hamiltonian  $H_S(\tau) = H_S$  is time-independent, and so are the Lindblad generator  $\mathcal{L}_\tau = \mathcal{L}$  of its dissipative dynamics  $\rho_S(0) \mapsto \rho_S(\tau)$  and the Gibbs state  $\rho_S^\beta$  (such that  $\mathcal{L}[\rho_S^\beta] = 0$ ). In such a case, the proof of the positivity of the entropy production follows from equation (40) becoming

$$d\Sigma_S(\tau) = -dD(\rho_S(\tau)\|\rho_S^\beta) \tag{43}$$

and from the monotonicity of the relative entropy under completely-positive, trace preserving dynamics<sup>30</sup>.

In the finite expression of the second law of thermodynamics (which follows from equation (33) in the absence of initial correlations between  $S$  and  $B$ ), the heat bath  $B$  is taken explicitly and directly into account (though the term  $\Delta\mathbb{S}_B$ ). Rather, in the infinitesimal expression (42), the heat bath is indirectly accounted for by the fact that (i) the heat exchange occurs at the bath temperature, and (ii) that the dissipative reduced dynamics of system  $S$  is determined by the bath in the weak-coupling limit.

Notwithstanding these fundamental physical differences, it is still interesting to study to which extent the thermodynamical inequality  $d\Sigma_S(\tau) \geq 0$  can be related to the behavior of  $d\tilde{\Sigma}_{S,B}(\tau)$  in equation (39). It is evident that  $d\tilde{\Sigma}_{S,B}(\tau)$  cannot be both strictly positive in general. For example, in the case of the same instantaneous pseudo-temperatures, as when  $dU_\chi(\tau) = 0$  and  $T_\chi(\tau) \neq 0$ , from

$$d\tilde{\Sigma}_S(\tau) + d\tilde{\Sigma}_B(\tau) = \frac{d\mathbb{W}_S(\tau)}{T_S(\tau)} + \frac{d\mathbb{W}_B(\tau)}{T_B(\tau)} \tag{44}$$

we obtain  $d\tilde{\Sigma}_S(\tau) = -d\tilde{\Sigma}_B(\tau)$  for  $d\mathbb{W}_S(\tau) = -d\mathbb{W}_B(\tau)$ . In general, it is not true that the finite variation

$$\Delta\tilde{\Sigma}_S(\tau) + \Delta\tilde{\Sigma}_B(\tau) = \int_0^\tau ds \frac{d\mathbb{W}_S(s)}{ds} \left( \frac{1}{T_S(s)} - \frac{1}{T_B(s)} \right) \tag{45}$$

becomes nonnegative in the absence of initial correlations between  $S$  and  $B$ —unlike the case for the finite variations of the von Neumann entropies of the reduced states  $\rho_{S,B}(\tau)$ .

One can argue that the infinitesimal quantities  $d\tilde{\Sigma}_{S,B}(\tau)$  do not generically behave as expected from true thermodynamic quantities because the instantaneous pseudo-temperatures do not behave themselves as thermodynamic temperatures. This, however, does not exclude that, under certain conditions, proper thermodynamic patterns might emerge.

To alleviate the above situation, we can discern a better motivated notion of temperature by appealing to analogy with standard thermodynamics. In *classical* thermodynamics the relation

$$d\mathbb{S} = \frac{1}{T}d\mathbb{Q} \tag{46}$$

holds for a system undergoing a quasistatic reversible transformation, whereas for a nonequilibrium process there is an extra term corresponding to the internal entropy production  $\Sigma$ ,

$$d\mathbb{S} = \frac{1}{T}d\mathbb{Q} + d\Sigma. \tag{47}$$



In this case the “temperature” is fixed by the external environment (bath) which is supposed to exchange heat always quasistatically (because of its short relaxation times), *without changing its temperature*. In our formalism, however, we treat the system and bath similarly. Thus we can extend equation (47) and identify an *extended* temperature and an entropy production for both system and bath and see how they compare at long times with expected thermodynamic temperatures. One way to do so is to explicitly compute  $dS$  and  $dQ$  and next compare them to derive a physically meaningful extended temperature  $T(\tau)$  as

$$dS(\tau) = \frac{1}{T(\tau)}dQ(\tau) + d\Sigma(\tau). \quad (48)$$

**Remark 6.** Note that equation (48) defines both the extended temperature  $T(\tau)$  and the generalized entropy production  $d\Sigma(\tau)$ . Moreover, unlike the pseudo-temperature  $T(\tau)$ ,  $T(\tau)$  is by construction  $\alpha_{S,B}$ -independent because neither heat nor entropy depends on  $\alpha_{S,B}$ . In the following examples, we discuss both nonequilibrium temperatures  $T(\tau)$  and  $T(\tau)$  by comparing them with the equilibrium temperature  $T$  (of the bath).

## Examples

Here we study in detail two examples, in one of which *thermalization* occurs, whereas the other one does not exhibit this feature.

**Example I: Thermalizing qubit.** Consider a two-state system (e.g., a spin-1/2 particle or a two-level atom) interacting with a thermal environment, comprised of infinitely many modes at (initial) temperature  $T = 1/\beta$ , through the Jaynes-Cummings total Hamiltonian  $H = H_0 + H_{\text{int}}^{(\lambda)}$ , where

$$H_0 = \frac{1}{2}\omega_0\sigma_z + \sum_{k=1}^{\infty}\omega_k\mathbf{a}_k^\dagger\mathbf{a}_k, \quad (49)$$

$$H_{\text{int}}^{(\lambda)} = \lambda\sum_k(f_k^*\sigma_+ \otimes \mathbf{a}_k + f_k\sigma_- \otimes \mathbf{a}_k^\dagger). \quad (50)$$

Here  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z = \text{diag}(1, -1)$  are the Pauli operators,  $\sigma_\pm = \sigma_x \pm i\sigma_y$ , and  $\mathbf{a}_k$  is the bosonic annihilation operator for mode  $k$ . Although this model is not exactly solvable, we can find the exact states of the system and bath up to any order in  $\lambda$ ; see supplementary information for details of  $O(\lambda^3)$  calculations. A separate study of the dynamics of the subsystems in the damped Jaynes-Cummings model can be found in ref. 32.

In the weak-coupling, long-time,  $\omega$ -continuum, Markovian limit (where  $\lambda \rightarrow 0$  and  $\tau \rightarrow \infty$  such that  $\lambda^2\tau = \text{const.}$  and  $\sum_k \rightarrow \int_0^\infty d\omega$ ), we can find the following Lindblad-type dynamical equation:

$$\begin{aligned} \frac{d}{d\tau}\rho_S^{(\lambda)}(\tau) = & -i[H_S + H_{\text{LS}}, \rho_S^{(\lambda)}(\tau)] + \frac{\gamma}{2}(\bar{n}(\omega_0, \beta) + 1)(2\sigma_- \rho_S^{(\lambda)}(\tau)\sigma_+ \\ & - \{\sigma_+\sigma_-, \rho_S^{(\lambda)}(\tau)\}) + \frac{\gamma}{2}\bar{n}(\omega_0, \beta) \times (2\sigma_+ \rho_S^{(\lambda)}(\tau)\sigma_- \\ & - \{\sigma_-\sigma_+, \rho_S^{(\lambda)}(\tau)\}), \end{aligned} \quad (51)$$

where

$$H_{\text{LS}} = 2\lambda^2\mathbb{P}\int_0^\infty d\omega \frac{|f(\omega)|^2}{\omega_0 - \omega}(2\bar{n}(\omega, \beta) + 1)\sigma_z \quad (52)$$

$$=: (1/2)\Omega(\lambda^2, \omega_0, \beta)\sigma_z, \quad (53)$$

is the Lamb-shift Hamiltonian,  $\mathbb{P}$  denotes the Cauchy principal value,  $\beta$  is the inverse temperature of the bath,

$$\bar{n}(\omega_0, \beta) = (e^{\beta\omega_0} - 1)^{-1} \quad (54)$$

is the Planck distribution or the mean quanta number in a mode with frequency  $\omega_0$ , and

$$\gamma = 2\pi\lambda^2|f(\omega_0)|^2 \quad (55)$$

is the spontaneous emission rate. This evolution agrees with the Markovian master equation derived in ref. 29. The solution to equation (51) is given by

$$\rho_S^{(\lambda)}(\tau) = \frac{1}{2}\left[ \begin{array}{cc} 1 + z(0)e^{-\tilde{\gamma}\tau} + \tanh(\beta\omega_0/2)(e^{-\tilde{\gamma}\tau} - 1) & (x(0) - iy(0))e^{-i(\omega_0+\Omega)\tau - \tilde{\gamma}\tau/2} \\ (x(0) + iy(0))e^{i(\omega_0+\Omega)\tau - \tilde{\gamma}\tau/2} & 1 - z(0)e^{-\tilde{\gamma}\tau} - \tanh(\beta\omega_0/2)(e^{-\tilde{\gamma}\tau} - 1) \end{array} \right], \quad (56)$$

where  $\tilde{\gamma} = \gamma \coth(\beta\omega_0/2)$ . It is evident from this solution that system S eventually thermalizes,

$$\lim_{\tau \rightarrow \infty} \rho_S^{(\lambda)}(\tau) = \rho_S^\beta, \quad (57)$$

where  $\rho_S^\beta = (1/Z_\beta)e^{-\beta\omega_0\sigma_z/2}$  is a thermal state in the Gibbs form, in which  $Z_\beta = \text{Tr}[e^{-\beta\omega_0\sigma_z/2}]$  is the partition function.

We can explicitly compute  $d\mathbb{S}_S^{(\lambda)}(\tau)$  as

$$d\mathbb{S}_S^{(\lambda)}(\tau) = -\frac{1}{2} \log \frac{1 + r_S^{(\lambda)}(\tau)}{1 - r_S^{(\lambda)}(\tau)} dr_S^{(\lambda)}(\tau), \tag{58}$$

where  $r_S^{(\lambda)}(\tau)$  is the norm of the Bloch vector  $\mathbf{r}_S^{(\lambda)} = (x, y, z)$  associated with  $\rho_S^{(\lambda)}(\tau)$  as  $\rho_S = (1/2)(I + \mathbf{r} \cdot \boldsymbol{\sigma})$  (here  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ ), and from equation (51) we have

$$dr_S(\tau) = \frac{\tilde{\gamma}[x^2(\tau) + y^2(\tau)] - 2\gamma z(\tau) - 2\tilde{\gamma}z^2(\tau)}{2\sqrt{x^2(\tau) + y^2(\tau) + z^2(\tau)}} d\tau. \tag{59}$$

Additionally, in the Markovian limit, the energy of this system is obtained as

$$d\mathbb{U}_S^{(\lambda)}(\tau) = -\frac{\omega_0}{2}\gamma e^{-\tilde{\gamma}\tau} (\coth(\beta\omega_0/2)z(0) + 1) d\tau. \tag{60}$$

As a result,

$$\lim_{\tau \rightarrow \infty} \frac{1}{T_S^{(\lambda)}(\tau)} = \beta \left[ 1 - \frac{(x^2(0) + y^2(0)) \coth(\beta\omega_0/2)}{2(z(0) + \tanh(\beta\omega_0/2))} \right]. \tag{61}$$

This pseudo-temperature behaves well, i.e., exhibits thermalization, if there is no initial coherence ( $\rho_{10} = 0$ , or equivalently,  $x(0) = y(0) = 0$ ).

In the Markovian regime we consider the thermal bath always in equilibrium (namely,  $\rho_B(\tau) \approx \rho_B^\beta$ ), and as a consequence the effective energy of S reduces to (see equation (8))

$$\mathbb{U}_S(\tau) = \text{Tr}[\rho_S(\tau)H_S], \tag{62}$$

and the heat flux reads

$$d\mathbb{Q}_S(\tau) = \text{Tr}[d\rho_S(\tau)H_S] = \frac{\omega_0}{2} dz(\tau). \tag{63}$$

Comparing equations (58) and (63) yields

$$\frac{1}{T_S(\tau)} = -\frac{1}{\omega_0} \frac{z(\tau)}{r_S(\tau)} \log \frac{1 + r_S(\tau)}{1 - r_S(\tau)}, \tag{64}$$

$$d\Sigma_S(\tau) = -\frac{1}{2} \frac{dr_S(\tau)}{r_S(\tau)} \log \frac{1 + r_S(\tau)}{1 - r_S(\tau)}, \tag{65}$$

which are both  $\alpha_{S,B}$ -independent. By substituting the Bloch vector components of the Gibbs state  $\rho_S^\beta$ , ( $x=0, y=0, z = -\tanh(\beta\omega_0/2)$ ), in equation (64), we also see that  $\lim_{\tau \rightarrow \infty} T_S(\tau) = T$ , which gives the expected equilibrium temperature.

For the bath thermodynamics, after some algebra we find that when  $\tau \rightarrow \infty$  (up to  $O(\lambda^3)$ )

$$d\mathbb{Q}_B^{(\lambda)}(\tau) = 4\omega_0\gamma [(\bar{n}(\omega_0, \beta) + 1)\rho_{00} - \bar{n}(\omega_0, \beta)\rho_{11}] - |\rho_{10}|^2 d\tau, \tag{66}$$

$$d\mathbb{U}_B^{(\lambda)}(\tau) = 4\gamma\omega_0 [(\bar{n}(\omega_0, \beta) + 1)\rho_{00} - \bar{n}(\omega_0, \beta)\rho_{11}] d\tau, \tag{67}$$

$$d\mathbb{S}_B^{(\lambda)}(\tau) = 4\beta\gamma\omega_0 [(\bar{n}(\omega_0, \beta) + 1)\rho_{00} - \bar{n}(\omega_0, \beta)\rho_{11} - |\rho_{10}|^2] d\tau, \tag{68}$$

whence

$$\lim_{\tau \rightarrow \infty} T_B^{(\lambda)}(\tau) = \frac{1}{\beta} \left[ 1 + \frac{|\rho_{10}|^2}{\bar{n}(\omega_0, \beta)(\rho_{00} - \rho_{11}) + \rho_{00} - |\rho_{10}|^2} \right], \tag{69}$$

$$d\mathbb{S}_B^{(\lambda)}(\tau) = \beta d\mathbb{Q}_B^{(\lambda)}(\tau). \tag{70}$$

Note that the limit (69) is independent of  $\alpha_B$  but it depends on the initial state of the system. However, if  $\rho_S(0)$  does not have any coherence, i.e.,  $\rho_{10} = 0$ , one retrieves the expected value  $1/\beta$  for the pseudo-temperature  $T_B^{(\lambda)}$ . But regardless of the initial state of the system, from equation (70), we see that the extended temperature behaves as expected,  $\lim_{\tau \rightarrow \infty} T_B^{(\lambda)}(\tau) = T$ . Besides, the internal entropy production of the bath up to  $O(\lambda^3)$  vanishes,



$$d\Sigma_B^{(\lambda)} = 0. \tag{71}$$

**Remark 7.** Following the discussion in Remark 3, the reason for the difference between the pseudo-temperature and the standard thermodynamic temperature lies in the definition of the former. The entropy of the qubit in this example can be computed using its eigenvalues, which in general depend on the  $(x, y, z)$  components of the Bloch vector. From equation (62), we can identify the  $z$  component with  $\mathbb{U}$  (assuming  $\omega_0/2 = 1$  for simplicity). Thus we can say that  $S$  is a function of  $(x, y, \mathbb{U})$ , and we can compute the partial derivative of  $S$  with respect to  $\mathbb{U}$  (while keeping  $x$  and  $y$  fixed),

$$\left(\frac{\partial S(x, y, \mathbb{U})}{\partial \mathbb{U}}\right)_{x,y} = -\frac{1}{2} \frac{\mathbb{U}}{\sqrt{x^2 + y^2 + \mathbb{U}^2}} \log \frac{1 + \sqrt{x^2 + y^2 + \mathbb{U}^2}}{1 - \sqrt{x^2 + y^2 + \mathbb{U}^2}}. \tag{72}$$

If we now consider  $x(\tau), y(\tau)$  and  $\mathbb{U}(\tau)$  evolving according to the dissipative thermalizing dynamics (56), we obtain

$$\lim_{\tau \rightarrow \infty} \left(\frac{\partial S(x, y, \mathbb{U})}{\partial \mathbb{U}}\right)_{x,y} = \beta, \tag{73}$$

which agrees with the standard definition of the equilibrium temperature. Rather, the inverse pseudo-temperature  $1/T_S^{(\lambda)}(\tau)$  reads as

$$\frac{dS(x(\tau), y(\tau), z(\tau))/d\tau}{d\mathbb{U}/d\tau}, \tag{74}$$

which corresponds to inverting the function  $\mathbb{U}(\tau)$ , finding  $\tau(\mathbb{U})$ , and computing the total derivative with respect to  $\mathbb{U}$ ,

$$\frac{dS}{d\mathbb{U}}(x(\mathbb{U}), y(\mathbb{U}), \mathbb{U}) = -\frac{1}{2} \frac{\mathbb{U} + x(\mathbb{U}) \frac{dx}{d\mathbb{U}} + y(\mathbb{U}) \frac{dy}{d\mathbb{U}}}{\sqrt{x^2(\mathbb{U}) + y^2(\mathbb{U}) + \mathbb{U}^2}} \log \frac{1 + \sqrt{x^2(\mathbb{U}) + y^2(\mathbb{U}) + \mathbb{U}^2}}{1 - \sqrt{x^2(\mathbb{U}) + y^2(\mathbb{U}) + \mathbb{U}^2}}. \tag{75}$$

In the  $\tau \rightarrow \infty$  limit (or  $\mathbb{U} \rightarrow U_{\text{thermal}}$ ) this derivative is different from  $\beta$  because in general

$$x(\mathbb{U}) \frac{dx}{d\mathbb{U}} + y(\mathbb{U}) \frac{dy}{d\mathbb{U}} = \frac{x^2(0) + y^2(0)}{2(z(0) + \tanh(\beta))} \neq 0. \tag{76}$$

The two derivatives coincide only if  $x$  and  $y$  are fixed during the dynamics, which is the case of vanishing initial coherence.

**Example II: Dephasing qubit.** We apply the previous considerations to the exactly solvable model of a qubit in interaction with a thermal bosonic bath<sup>33</sup>. The total Hamiltonian is given by  $H_{\text{tot}} = H_0 + H_{\text{int}}^{(\lambda)}$  with

$$H_0 = \frac{\omega_0}{2} \sigma_z + \sum_{k=1}^{\infty} \omega_k a_k^\dagger a_k, \quad H_{\text{int}}^{(\lambda)} = \lambda \sigma_z \otimes (a(f) + a^\dagger(f)),$$

where  $a_k$  is the bosonic annihilation operator of mode  $k$ , satisfying the commutation relations  $[a_k, a_l^\dagger] = \delta_{kl}$ , and

$$a(f) = \sum_k f_k^* a_k, \tag{77}$$

with complex quantities  $f_k$  forming a square-summable vector  $f = \{f_k\} \in L^2(-\infty, +\infty)$  such that

$$[a(f), a^\dagger(g)] = \langle f|g \rangle. \tag{78}$$

In the interaction picture, the Hamiltonian becomes

$$\widetilde{H}_{\text{int}}^{(\lambda)}(\tau) = U_0^\dagger(\tau) H_{\text{int}}^{(\lambda)} U_0(\tau) = \lambda \sigma_z \otimes (a(f_\tau) + a^\dagger(f_\tau)), \tag{79}$$

where  $U_0(\tau) = e^{-iH_0\tau}$  and  $f_\tau$  is the vector with components  $f_k^* e^{-i\omega_k\tau}$ . The time-ordered exponentiation of  $\widetilde{H}_{\text{int}}^{(\lambda)}(\tau)$  yields

$$\begin{aligned} \widetilde{U}_\lambda(\tau) &= \mathbb{T} e^{-i\lambda\sigma_z \otimes \int_0^\tau ds (a(f_s) + a^\dagger(f_s))} \\ &= e^{-i\lambda^2 \varphi(\tau)} e^{-i\lambda\sigma_z \otimes \int_0^\tau ds (a(f_s) + a^\dagger(f_s))}, \end{aligned}$$

where the pure phase factor

$$\varphi(\tau) = \sum_k \frac{|f_k|^2}{\omega_k^2} (\omega_k \tau - \sin(\omega_k \tau)) \tag{80}$$

does not contribute to the evolution

$$\rho_{SB}^{(\lambda)}(\tau) = U_0(\tau) \tilde{U}_\lambda(\tau) \rho_{SB}(0) \tilde{U}_\lambda^\dagger(\tau) U_0^\dagger(\tau). \tag{81}$$

Let us assume that the initial state of the total system is factorized and given by  $\rho_{SB}(0) = \rho_S(0) \otimes \rho_B^\beta$ , where

$$\rho_S(0) = \sum_{\ell, \ell'=0}^1 \rho_{\ell\ell'} |\ell\rangle \langle \ell'|, \sigma_z |\ell\rangle = (-)^\ell |\ell\rangle, \tag{82}$$

is the initial state of the qubit, and  $\rho_B^\beta$  is the Gibbs (thermal) state of the bosonic bath with the inverse temperature  $\beta$ ,

$$\rho_B^\beta = e^{-\beta \sum_k \omega_k a_k^\dagger a_k} / Z_\beta, \tag{83}$$

and  $Z_\beta = \text{Tr}[e^{-\beta \sum_k \omega_k a_k^\dagger a_k}]$  is the associated partition function. One can see that

$$\rho_{SB}^{(\lambda)}(\tau) = \sum_{\ell, \ell'=0}^1 \rho_{\ell\ell'} e^{-i\omega_0 \zeta_{\ell\ell'} \tau / 2} |\ell\rangle \langle \ell'| \otimes D_\ell(g_\tau) \rho_B^\beta D_\ell^\dagger(g_\tau), \tag{84}$$

where  $\zeta_{\ell\ell'} = (-)^\ell - (-)^{\ell'}$ ,  $g_\tau$  is the vector with components

$$g_k^*(\tau) = f_k^* (e^{-i\omega_k \tau} - 1) / \omega_k, \tag{85}$$

and  $D_\alpha(g_\tau)$  is the displacement operator

$$D_\ell(g_\tau) = e^{(-)^\ell \lambda [a^\dagger(g_\tau) - a(g_\tau)]}, \tag{86}$$

whose action can be derived from equation (78) as

$$D_\ell(g_\tau) a_k D_\ell^\dagger(g_\tau) = a_k - (-)^\ell \lambda g_k(\tau) \tag{87}$$

$$=: A_k(\ell, \lambda, \tau). \tag{88}$$

From here the reduced density matrices of the two subsystems read as

$$\rho_S^{(\lambda)}(\tau) = \rho_{00} |0\rangle \langle 0| + \rho_{11} |1\rangle \langle 1| + e^{-8\lambda^2 \Gamma(\tau)} (\rho_{10} e^{i\omega_0 \tau} |1\rangle \langle 0| + \rho_{01} e^{-i\omega_0 \tau} |0\rangle \langle 1|), \tag{89}$$

$$\rho_B^{(\lambda)}(\tau) = \sum_{\ell=0}^1 \rho_{\ell\ell} D_\ell(g_\tau) \rho_B^\beta D_\ell^\dagger(g_\tau) = \sum_{\ell=0}^1 \frac{\rho_{\ell\ell}}{Z_\beta} e^{-\beta \sum_k \omega_k A_k^\dagger(\ell, \lambda, \tau) A_k(\ell, \lambda, \tau)}, \tag{90}$$

where  $\text{Tr}[D_\ell(g_\tau) \rho_B^\beta D_\ell^\dagger(g_\tau)] = e^{-8\lambda^2 \Gamma(\tau)}$  for  $\alpha \neq \delta$ , with

$$\Gamma(\tau) = \sum_k \frac{|f_k|^2}{\omega_k^2} \coth(\beta \omega_k / 2) \sin^2(\omega_k \tau / 2). \tag{91}$$

Further, denoting the qubit polarization at time  $\tau=0$  by  $\langle \sigma_z \rangle_S$ , the effective qubit Hamiltonian takes the form

$$H_S^{(\text{eff})}(\tau) = \left( \frac{\omega_0}{2} - 4\lambda^2 \langle \sigma_z \rangle_S \Delta(\tau) \right) \sigma_z + 4\lambda^2 \alpha_S \langle \sigma_z \rangle_S^2 \Delta(\tau), \tag{92}$$

where the explicit time dependence is provided by

$$\Delta(\tau) = -\frac{1}{4\lambda \langle \sigma_z \rangle_S} \text{Tr}[\rho_B^{(\lambda)}(\tau) (a(f) + a^\dagger(f))] = \sum_k \frac{|f_k|^2}{\omega_k} \sin^2(\omega_k \tau / 2). \tag{93}$$

Similarly, the bath effective Hamiltonian reads as

$$H_B^{(\text{eff})}(\tau) = \sum_k \omega_k a_k^\dagger a_k + \lambda \langle \sigma_z \rangle_S (a(f) + a^\dagger(f)) + 4\lambda^2 \alpha_B \langle \sigma_z \rangle_S^2 \Delta(\tau), \tag{94}$$

where the time-dependent appears only in the scalar term. From the above relations, the exchanged works between system  $S$  and bath  $B$  are calculated by using equation (4),

$$d\mathbb{W}_B^{(\lambda)}(\tau) = 4\lambda^2 \alpha_B \langle \sigma_z \rangle_S^2 d\Delta(\tau) = -d\mathbb{W}_S^{(\lambda)}(\tau), \tag{95}$$

where the last equality verifies equation (23). In addition, using equation (24) and the fact that

$$[H_S^{(\text{eff})}(\tau), H_{\text{int}}] = 0, [H_B^{(\text{eff})}(\tau), H_{\text{int}}] = \lambda \sigma_z \otimes \sum_k \omega_k (f_k a_k^\dagger - f_k^* a_k),$$

the infinitesimal heat exchanges are given by

$$dQ_S^{(\lambda)}(\tau) = 0, \quad (96)$$

$$dQ_B^{(\lambda)}(\tau) = 4\lambda^2(1 - \langle \sigma_z \rangle_S^2) d\Delta(\tau). \quad (97)$$

The binding energy also becomes

$$\mathbb{U}_\chi^{(\lambda)}(\tau) = -4\lambda^2(1 - \langle \sigma_z \rangle_S^2) \Delta(\tau), \quad (98)$$

whence  $dQ_B^{(\lambda)}(\tau) = -d\mathbb{U}_\chi^{(\lambda)}(\tau)$ , in agreement with equation (26).

Equation (96) is physically expected because, with our specific system Hamiltonian ( $H_S \propto \sigma_z$ ) and the interaction Hamiltonian ( $H_{\text{int}} \propto \sigma_z \otimes (a + a^\dagger)$ ), we have  $[H_S, H_{\text{int}}] = 0$ . That is, this interaction with the environment cannot excite or change the *populations* of  $\rho_S(0)$ ;  $\rho_{00}(\tau) = \rho_{00}$  [equation (89)]. Thus according to the definition of the heat exchange, we should have  $dQ_S^{(\lambda)}(\tau) = \text{Tr}[d\rho_S^{(\lambda)}(\tau)H_S^{\text{eff}}(\tau)] = \sum_{\ell=0}^1 d\rho_{\ell\ell}(\tau)q(\tau)\sigma_z$ ,  $\ell\ell = 0$ , where we have used the fact that  $H_S^{\text{eff}}(\tau) = q(\tau)\sigma_z$  (for some appropriate  $q$  read through equation (92)).

Furthermore, using equations (95) and (97), and the fact that  $\alpha_S + \alpha_B = 1$ , it turns out that, unlike the infinitesimal heat exchanges, the infinitesimal variation of the internal energy of  $B$  depends on  $\alpha_S$ ,

$$d\mathbb{U}_B^{(\lambda)}(\tau) = 4\lambda^2(1 - \alpha_S \langle \sigma_z \rangle_S^2) d\Delta(\tau). \quad (99)$$

One expects the final pseudo-temperature of  $T_B^{(\lambda)}(\infty)$ —as defined by equation (34)—to tend to the (initial) bath temperature  $T = 1/\beta$  in the limit  $\lambda \rightarrow 0$  of vanishing coupling between  $S$  and  $B$ . Indeed, if  $\lambda = 0$ , the thermal state (83) is time-invariant. Since  $\text{Tr}[d\rho_B^{(\lambda)}(\tau)] = 0$ , the infinitesimal variation of the von Neumann entropy of  $B$  is given by

$$dS_B^{(\lambda)}(\tau) = -\text{Tr}[d\rho_B^{(\lambda)}(\tau) \log \rho_B^{(\lambda)}(\tau)]. \quad (100)$$

By expanding equation (90) up to  $O(\lambda^3)$  one obtains

$$dS_B^{(\lambda)}(\tau) = 4\beta\lambda^2(1 - \langle \sigma_z \rangle_S^2) d\Delta(\tau). \quad (101)$$

Now if we use equations (95) and (97), together with the definition of the pseudo-temperature (34), we obtain

$$\lim_{\tau \rightarrow \infty} T_B^{(\lambda)}(\tau) = \frac{(1 - \alpha_S \langle \sigma_z \rangle_S^2)}{(1 - \langle \sigma_z \rangle_S^2)} T. \quad (102)$$

It is evident from this expression that in order to make the pseudo-temperature  $T_B^{(\lambda)}$  to be equal to  $T$  (in the weak-coupling limit) we need to set  $\alpha_S = 1$ .

Additionally, we note that by comparing equations (97) and (101), these quantities are related as

$$dS_B^{(\lambda)}(\tau) = \beta dQ_B^{(\lambda)}(\tau). \quad (103)$$

Hence, we have  $T_B(\tau) = T$  and the inverse temperature  $\beta = 1/T$  of the bath shows up as the prefactor of the heat flux, as expected in the standard equilibrium thermodynamics (equation (47)). Thus up to  $O(\lambda^3)$  the internal entropy production in the bath vanishes,

$$d\Sigma_B^{(\lambda)}(\tau) = 0. \quad (104)$$

This is consistent with the classical picture where the bath always exchanges heat quasistatically—see our discussion in the second law of thermodynamics.

**Remark 8.** We have verified in two different models that the internal entropy production in a thermal bath vanishes in the weak-coupling limit up to the leading order in  $\lambda$ . This seems to be a general result and is consistent with our expectation from standard, equilibrium thermodynamics.

Now we consider the pseudo-temperature  $T_S^{(\lambda)}(\tau)$ . We first note that, from equations (95) and (96) and after setting  $\alpha_S = 1$ , we have

$$dQ_S^{(\lambda)}(\tau) = d\mathbb{W}_S^{(\lambda)}(\tau) = 0, \quad (105)$$

and thus

$$d\mathbb{U}_S^{(\lambda)}(\tau) = 0. \quad (106)$$

That is, despite interacting with bath  $B$ , system  $S$  does not exchange any heat or work (and thus internal energy) with  $B$ . Hence intuitively we should not expect that its temperature  $T_S^{(\lambda)}$  to change; it should remain constant. This is explicitly seen by calculating

$$T_S^{(\lambda)}(\tau) = \frac{dU_S^{(\lambda)}(\tau)}{dS_S^{(\lambda)}(\tau)} = 0. \quad (107)$$

Note that if the system were initially prepared, e.g., in a thermal state with temperature  $T_S^{(0)} \neq 0$ , in principle its temperature should not change because this system does not thermalize [equation (89)]. This fact is captured by our pseudo-temperature as  $T_S^{(\lambda)} = 0$ . However, we note that  $T(\tau)$  is defined by the given *dynamics* of  $S$  and cannot therefore be related to an initial (dynamics-independent) temperature such as an equilibrium temperature assigned to the preparation of the state.

Having calculated the heat and work exchanges by the system, it is also important to see how the entropy of the system behaves. Using equation (89), the entropy of  $S$  can be explicitly calculated from the eigenvalues  $(1 \pm r_S^{(\lambda)}(\tau))/2$  of  $\rho_S^{(\lambda)}(\tau)$ , where

$$r_S^{(\lambda)}(\tau) = \sqrt{1 - 4(\rho_{00}\rho_{11} - e^{-16\lambda^2\Gamma(\tau)} |\rho_{01}|^2)}, \quad (108)$$

as well as its infinitesimal variation

$$dS_S^{(\lambda)}(\tau) = -\frac{1}{2} \log \frac{1 + r_S^{(\lambda)}(\tau)}{1 - r_S^{(\lambda)}(\tau)} dr_S^{(\lambda)}(\tau) = \lambda^2 b^{(\lambda)}(\tau) d\Gamma(\tau), \quad (109)$$

where

$$b^{(\lambda)}(\tau) = \frac{16 |\rho_{01}|^2}{r_S^{(\lambda)}(\tau)} e^{-16\lambda^2\Gamma(\tau)} \log \frac{1 + r_S^{(\lambda)}(\tau)}{1 - r_S^{(\lambda)}(\tau)}. \quad (110)$$

Note that the quantity  $b^{(\lambda)}(\tau)$  is nonnegative and has a well-defined time-independent limit,

$$b^{(0)} = \frac{16 |\rho_{01}|^2}{r_S^{(0)}(0)} \log \frac{1 + r_S^{(0)}(0)}{1 - r_S^{(0)}(0)}, \quad (111)$$

when  $\lambda \rightarrow 0$ . In order to study the time-derivatives  $d\Delta(\tau)/d\tau$  and  $d\Gamma(\tau)/d\tau$ , we consider an *infinite* thermal bath with a *continuum*  $\omega$  and a regularized Ohmic spectral density given by  $f_k \simeq \sqrt{\omega} e^{-\omega\epsilon/2}$  (in which  $\epsilon \geq 0$ ). Thus we substitute the discrete sums in equations (91) and (93) with the following integrals:

$$\Delta(\tau) = \int_0^\infty d\omega \sin^2(\omega\tau/2) e^{-\epsilon\omega} = \frac{\tau^2}{2\epsilon(\epsilon^2 + \tau^2)}, \quad (112)$$

$$\Gamma(\tau) = \int_0^\infty d\omega \frac{1}{\omega} \coth(\beta\omega/2) \sin^2(\omega\tau/2) e^{-\epsilon\omega}. \quad (113)$$

Hence,  $d\Delta(\tau)/d\tau = \epsilon\tau/(\tau^2 + \epsilon^2)^2$  as well as  $d\Gamma(\tau)/d\tau \geq 0$ , as one can check by changing the variable  $\omega\tau = \tilde{\omega}$  and taking explicitly the derivative with respect to  $\tau$ . As a result, we see that  $dS_S^{(\lambda)}(\tau) \geq 0$ . Furthermore, as a consequence of equation (96), in this regime, the internal entropy production relation (39) reduces to

$$d\Sigma_S(\tau) = dS_S^{(\lambda)}(\tau) \geq 0. \quad (114)$$

That is, the whole entropy change in the system is entirely due to the internal entropy production, whence the extended temperature  $\mathcal{T}(\tau)$  remains undefined because of Remark 6.

**Remark 9.** *It is an appealing feature of this model that the qubit does not exchange any energy with its environment ( $dU_S^{(\lambda)}(\tau) = 0$ ), whilst its (internal) entropy may change ( $dS_S^{(\lambda)}(\tau) \neq 0$ ) completely because of its (varying) correlations with the environment ( $dS_\chi(\tau) \neq 0$ ).*

It also may also be interesting to investigate the behavior of the various thermodynamic quantities in the Markovian regime for system  $S$ . This is determined by the condition  $\beta \ll \tau$  over the long timescale  $1/\lambda^2$  when  $\lambda \rightarrow 0$ . Under these conditions and after removal of the regularization parameter  $\epsilon$ , one obtains  $\Gamma(\tau) \simeq \pi\tau/(2\beta)$ . Thus, the dynamics of system  $S$  [equation (89)] reads as

$$\rho_S^{(\lambda)}(\tau) = \rho_{00}|0\rangle\langle 0| + \rho_{11}|1\rangle\langle 1| + e^{-\gamma\tau}(\rho_{10}e^{i\omega_0\tau}|1\rangle\langle 0| + \rho_{01}e^{-i\omega_0\tau}|0\rangle\langle 1|), \quad (115)$$

in which  $\gamma = 4\pi\lambda^2/\beta$ . This state solves the Lindblad-type master equation

$$\frac{d\rho_S^{(\lambda)}(\tau)}{d\tau} = -i \left[ \frac{1}{2} \omega_0 \sigma_z, \rho_S^{(\lambda)}(\tau) \right] + \frac{\gamma}{2} (\sigma_z \rho_S^{(\lambda)}(\tau) \sigma_z - \rho_S^{(\lambda)}(\tau)). \quad (116)$$

Note that this dynamics, similarly to dynamics generated by a Lindblad equation, has a fixed point as  $\lim_{\tau \rightarrow \infty} (1/\tau) \int_0^\tau ds \rho_S^{(\lambda)}(s) = \rho_{00}|0\rangle\langle 0| + \rho_{11}|1\rangle\langle 1|$ . Thus if we start with the system initially with no coherence

(i.e., vanishing off-diagonal elements,  $\rho_{10} = 0$ ), it will not evolve in time, and because of equation (110) the bath will not experience any entropy change either;  $\mathbb{S}_B^{(\lambda)}(\tau) = \text{const}$ .

## Summary and Outlook

This paper highlights the role of correlations in the nonequilibrium thermodynamic behavior of generic bipartite interacting quantum systems. In this formulation, interesting relations emerge between correlations, on the one hand, and heat, work exchanges, as well as possible definitions of nonequilibrium temperatures of each subsystem, on the other hand. These relations may enable the extraction of desired thermodynamic properties by partially controlling or manipulating the underlying dynamics of the system. A notion of binding energy has been introduced which only depends on the interaction Hamiltonian and correlations of the total system state, whose variation has been shown to be only of the heat type. In addition, this energy has been shown not to be locally accessible by the subsystems, but it provides a heat transmission channel between the parties. In this sense, correlations act as a resource or storage for heat. We have also defined two notions of nonequilibrium temperatures for the subsystems and discussed their relevance in the thermodynamic equilibrium. We have also associated a nonequilibrium temperature with correlations. This temperature may enable one to obtain conditions such that the two subsystems have same nonequilibrium temperatures, which are generically different exactly because of correlations. These results have been illustrated in detail through two examples: a qubit in interaction with a thermalizing bath and a qubit interacting with a dephasing environment.

Our methodology may provide techniques and tools for employing quantum resources, such as manybody correlations and memory, to engineer thermodynamic processes, for example, to build efficient quantum heat engines, or shed light on our understanding of the role of correlations in biological processes in relation to, e.g., the efficiency of photosynthetic light-harvesting complexes<sup>34</sup>.

## References

- Gemmer, J., Michel, M. & Mahler, G. *Quantum Thermodynamics* (Springer, Berlin, 2009).
- Alicki, R. The quantum open system as a model of the heat engine. *J. Phys. A: Math. Gen.* **12**, 103–107 (1979).
- Skrzypczyk, P., Short, A. J. & Popescu, S. Work extraction and thermodynamics for individual quantum systems. *Nature Commun.* **5**, 4185 (2014).
- Hänggi, P. & Talkner, P. The other QFT. *Nature Phys.* **11**, 108–110 (2015).
- Hossein-Nejad, H., O'Reilly, E. J. & Olaya-Castro, A. Work, heat and entropy production in bipartite quantum systems. *New J. Phys.* **17**, 075014 (2015).
- Weimer, H., Henrich, M. J., Rempp, F., Schröder, H. & Mahler, G. Local effective dynamics: A generalized approach to work and heat. *Europhys. Lett.* **83**, 30008 (2008).
- Balian, R. *From Microphysics to Macrophysics – Methods and Applications of Statistical Physics, Vol. I* (Springer, Berlin, 1991).
- Breuer, H.-P. & Petruccione, F. *The Theory of Open Quantum Systems* (Oxford University Press, New York, 2002).
- Campisi, M., Hänggi, P. & Talkner, P. Quantum fluctuation relations: Foundations and applications. *Rev. Mod. Phys.* **83**, 771–791 (2011).
- Gelbwaser-Klimovsky, D., Alicki, R. & Kurizki, G. Work and energy gain of heat-pumped quantized amplifiers. *Europhys. Lett.* **103**, 60005 (2013).
- Schröder, H. & Mahler, G. Work exchange between quantum systems: The spin-oscillator model. *Phys. Rev. E* **81**, 021118 (2010).
- Blundell, S. J. & Blundell, K. M. *Concepts in Thermal Physics* (Oxford University Press, New York, 2010).
- Landau, L. D. & Lifshitz, E. M. *Statistical Physics – Part I* (Pergamon Press, Oxford, 1980).
- Evans, D. J., Cohen, E. G. D. & Morriss, G. P. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **71**, 2401–2404 (1993).
- Evans, D. J. & Searles, D. J. Equilibrium microstates which generate second law violating steady states. *Phys. Rev. E* **40**, 1645–1648 (1994).
- Crooks, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **60**, 2721–2726 (1999).
- Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**, 2790–2693 (1997).
- Wang, G. M., Seville, E. M., Mittag, E., Searles, D. J. & Evans, D. J. Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys. Rev. Lett.* **89**, 050601 (2002).
- An, S., Zhang, J.-N., Um, M., Lv, D., Lu, Y., Zhang, J., Yin, Z.-Q., Quan, H. T. & Kim, K. Experimental test of the quantum Jarzynski equality with a trapped-ion system. *Nature Phys.* **11**, 193–199 (2015).
- Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).
- Wolf, M. M., Verstraete, F., Hastings, M. B. & Cirac, J. I. Area laws in quantum systems: mutual informations and correlations. *Phys. Rev. Lett.* **100**, 070502 (2008).
- Hall, M. J. W. *Entropy* **15**, 3698–3713 (2013).
- Reeb, D. & Wolf, M. M. An improved Landauer principle with finite-size corrections. *New J. Phys.* **16**, 103011 (2014).
- Casas-Vázquez, J. & Jou, D. Temperature in non-equilibrium states: a review of open problems and current proposals. *Rep. Prog. Phys.* **66**, 1937–2023 (2003).
- Martens, K., Bertin, E. & Droz, M. Dependence of the fluctuation-dissipation temperature on the choice of observable. *Phys. Rev. Lett.* **103**, 260602 (2009).
- Lloyd, S., Chiloyan, V., Hu, Y., Huberman, S., Liu, Z.-W. & Chen, G. No energy transport without discord. Preprint at <https://arxiv.org/abs/1510.05035> (2015).
- Lindblad, G. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **48**, 119–130 (1976).
- Gorini, V., Kossakowski, A. & Sudarshan, E. C. G. Completely positive dynamical semigroups of  $N$ -level systems. *J. Math. Phys.* **17**, 821–825 (1976).
- Carmichael, H. *An Open Systems Approach to Quantum Optics* (Springer, Berlin, 1993).
- Lindblad, G. Completely positive maps and entropy inequalities. *Commun. Math. Phys.* **40**, 147–151 (1975).
- Wehrli, A. General properties of entropy. *Rev. Mod. Phys.* **50**, 221–260 (1978).
- Rendell, R. W. & Rajagopal, A. K. Revivals and entanglement from initially entangled mixed states of a damped Jaynes-Cummings model. *Phys. Rev. A* **67**, 062110 (2003).
- Palma, G. M., Suominen, K.-A. & Ekert, A. Quantum computers and dissipation. *Proc. R. Soc. Lond. A* **452**, 567–584 (1996).
- Killoran, N., Huelga, S. F. & Plenio, M. B. Enhancing light-harvesting power with coherent vibrational interactions: A quantum heat engine picture. *J. Chem. Phys.* **143**, 155102 (2015).

## Acknowledgements

Hospitality of the Abdus Salam International Center for Theoretical Physics (ICTP) and the Institute for Research in Fundamental Sciences (IPM) are acknowledged, respectively, by S.A. and S.M., where parts of this work were completed. A.T.R. acknowledges financial support by Sharif University of Technology's Office of Vice President for Research, the Iran Science Elites Federation, and the IPM.

## Author Contributions

S.A., F. Benatti, F. Bakhshinezhad, M.A., S.M. and A.T.R. all contributed to the development and completion of the idea, performing the calculations, analyzing the results, discussions and writing the manuscript. F. Bakhshinezhad, M.A. and S.M. contributed equally to this work. The final draft of the manuscript was revised by A.T.R. and reviewed by all authors.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Alipour, S. *et al.* Correlations in quantum thermodynamics: Heat, work, and entropy production. *Sci. Rep.* **6**, 35568; doi: 10.1038/srep35568 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016