

---

# Noncanonical Equilibrium and Finite-Bath Corrections in Quantum Thermodynamics

---

In our previous discussion, I was given a brief introduction to the main research problem. In the last paper we studied, it was assumed that a small system weakly coupled to a large heat bath is described at equilibrium by a canonical Gibbs state with a well-defined temperature. However, this picture essentially relies on the idealization of an infinite bath. When the system-bath coupling is non-negligible and the bath has finite size, this description breaks down, and we need to reconsider what *equilibrium* means for our system.

In this report, two key papers will be reviewed: *Noncanonical statistics of a finite quantum system with non-negligible system-bath coupling* by Xu and Sun, and *Finite-bath corrections to the second law of thermodynamics* by Richens, Alhambra, and Masanes. The first paper examines how system-bath interactions deform the equilibrium state away from the standard Gibbs form, while the second explores how finite baths modify the second law and the free-energy bounds on work extraction. After understanding these results, the goal is to identify conditions under which interacting quantum subsystems can be in global equilibrium yet exhibit different effective temperatures, and to see how these effective temperatures depend on the underlying interaction parameters.

## 1 Noncanonical statistics of a finite quantum system with non-negligible system-bath coupling

### 1.1 Introduction

Statistical mechanics describes the average properties of a system without referring to all its microscopic states. In most situations, the validity of the canonical statistical description is guaranteed in the thermodynamic limit, which requires that, while the degrees of freedom of the heat bath is infinite, the system-bath coupling approaches infinitesimal. However, if the system only interacts with a small heat bath with finite degrees of freedom, the system-bath interaction cannot be ignored compared with the energy of the heat bath.

Consider a closed universe consisting of a small system  $S$  and a large bath  $B$ , described by a pure state  $|\Psi\rangle$  constrained to a narrow energy shell. Results on canonical typicality show that for almost all such pure states, the reduced state  $\rho_S = \text{Tr}_B(|\Psi\rangle\langle\Psi|)$  is extremely close to the canonical Gibbs state  $e^{-\beta H_S}/Z$ . In other words, thermal behavior of the subsystem does not require a mixed global ensemble; it arises generically from typical pure states of the total system. This is proven using Levys lemma (see [Popescu, Short, and Winter, Section IV and V](#)), which guarantees that local observables take values very close to their microcanonical averages for the overwhelming majority of pure global states. Therefore, for almost all pure

states of the universe, the reduced density matrix of the small system is (approximately) canonical.

Nevertheless, the conventional canonical statistics still cannot describe well the thermodynamics behavior of the finite system when the sufficiently large system-bath interaction is taken into consideration. In this paper, an effective system-bath coupling is generally considered by assuming the bath possesses a much more dense spectrum than that of the system, and then the system-bath interaction energy can be treated as the deformation of the energy shell for the total system. In this case, the system-bath interaction simply deforms the microcanonical energy shell of the total system, leading to a modified state that counts  $\Omega_B(E - E_n - \Delta_n)$  for each system level and, consequently, to noncanonical equilibrium statistics for the system. This modified distribution obviously implies that corrections are necessary for the finite-system thermodynamic quantities in canonical statistics, such as average internal energy and its fluctuation.

The structure of this section is as follows: in Sec. 1.2, we outline how Xu and Sun derive an effective Hamiltonian for the total system using perturbation theory, and how this leads to a noncanonical statistical distribution that does not rely on any particular microscopic model. In Sec. 1.3, we then discuss their concrete example of coupled harmonic oscillators, where they use this framework to compute thermodynamic quantities such as internal energy, energy fluctuations, and the mutual information between two subsystems. Finally, Sec. 1.4 briefly summarizes the main results and their implications.

## 1.2 Finite system-bath coupling

In this paper, a composite coupled system is considered, consisting of a system  $S$  with Hamiltonian  $H_S$  and a heat bath  $B$  with Hamiltonian  $H_B$ . The interaction between the system and the bath is described by  $H_I$ , so that the total Hamiltonian is

$$H = H_S + H_B + H_I.$$

The heat bath is composed of  $N$  non-interacting particles. The system and bath Hamiltonians have the following spectral decompositions:

$$H_S = \sum_n E_n |n\rangle \langle n| \tag{1}$$

$$H_B = \sum_{j=1}^N \sum_{k_j} \epsilon_{k_j} |k_j\rangle \langle k_j| \otimes \prod_{i \neq j} I_i. \tag{2}$$

Here,  $|n\rangle$  is the eigenstate of the system with corresponding eigenvalue  $E_n$ . For the bath,  $|k_j\rangle$  denotes the eigenstate of the  $j$ -th particle with eigenenergy  $\epsilon_{k_j}$ , and  $I_i$  is the identity

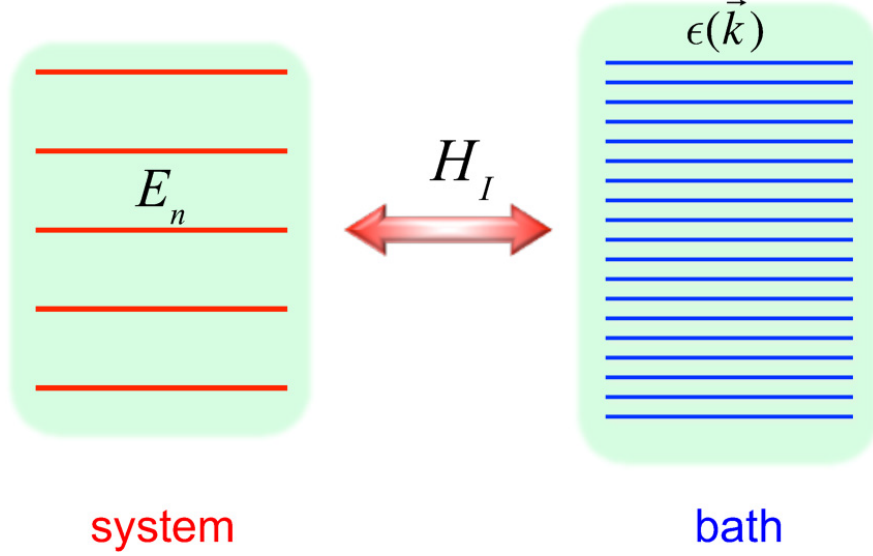


Figure 1: The energy spectra of the system and the heat bath.

operator on the Hilbert space of the  $i$ -th particle. For convenience, we introduce the notation

$$|\vec{k}\rangle \equiv \prod_{j=1}^N |k_j\rangle, \quad \vec{k} = (k_1, k_2, \dots, k_N),$$

to denote the eigenstates of  $H_B$ , and define their corresponding eigenenergies as

$$\epsilon(\vec{k}) = \sum_{j=1}^N \epsilon_{k_j}.$$

With these definitions, Eq. (2) can be rewritten in the usual compact form

$$H_B = \sum_{\vec{k}} \epsilon(\vec{k}) |\vec{k}\rangle \langle \vec{k}|.$$

Usually, the energy spectrum of the system is much sparser than the heat bath, i.e.,

$$\min |E_n - E_m| \gg \max |\epsilon(\vec{k}) - \epsilon(\vec{l})|, \quad (3)$$

where  $\min |E_n - E_m|$  denotes the minimal spacing between neighboring energy levels of the system, and  $\max |\epsilon(\vec{k}) - \epsilon(\vec{l})|$  is the maximal spacing between neighboring energy levels of the bath. This is illustrated in Fig. 1, taken from the original paper.

We assume that the interaction between the system and the bath can be written as a

sum of terms, each coupling the system to a single bath particle at a time:

$$H_I = \sum_{j=1}^N V_j,$$

where each  $V_j$  acts nontrivially only on  $\mathcal{H}_S \otimes \mathcal{H}_j$  and as the identity on all other bath degrees of freedom. Any operator on  $\mathbb{H}_S \otimes \mathbb{H}_j$  can be expanded in the product basis  $\{|n\rangle \otimes |k_j\rangle\}$ .

$$V_j = \sum_{n,n',k_j,k'_j} g_{nk_j,n'k'_j} |n, k_j\rangle \langle n', k'_j|,$$

with  $|n, k_j\rangle \equiv |n\rangle \otimes |k_j\rangle$  and  $g_{nk_j,n'k'_j} = \langle n, k_j | V_j | n', k'_j \rangle$ . Then embed with identities on the rest and we have

$$H_I = \sum_{j=1}^N \sum_{n,n',k_j,k'_j} g_{nk_j,n'k'_j} |n, k_j\rangle \langle n', k'_j| \otimes \prod_{i \neq j}^N I_i. \quad (4)$$

The system-bath interaction  $H_I$  is weak compared to  $H_0$ , with

$$H_0 = H_S + H_B.$$

To obtain the canonical density matrix  $\rho_S$  of the system, we start from the microcanonical ensemble of the composite system (system+bath), neglecting  $H_I$ . The microcanonical state is

$$\rho_{\text{mic}} = \frac{P_{V(E,\delta)}}{\dim V(E,\delta)}, \quad (5)$$

where  $V(E,\delta)$  is the subspace of  $H_S + H_B$  associated with the energy shell  $[E, E + \delta]$  (with  $\delta \ll E$ ),

$$V(E,\delta) = \text{Span}\{|n, \vec{k}\rangle : E \leq E_n + \epsilon(\vec{k}) \leq E + \delta\}, \quad (6)$$

and the corresponding projector onto  $V(E,\delta)$  has the explicit form

$$P_{V(E,\delta)} = \sum_{\substack{n, \vec{k} \\ E \leq E_n + \epsilon(\vec{k}) \leq E + \delta}} |n, \vec{k}\rangle \langle n, \vec{k}|. \quad (7)$$

The reduced state of the system is then

$$\rho_S = \text{Tr}_B (\rho_{\text{mic}}), \quad (8)$$

and the probability of finding the system in level  $|n\rangle$  is

$$P_n(E) = \langle n | \rho_S | n \rangle = \frac{\Omega(E - E_n, \delta)}{\dim V(E, \delta)}, \quad (9)$$

where  $\Omega(E - E_n, \delta) = \dim V_n(E - E_n, \delta)$  and

$$V_n(E - E_n, \delta) = \text{Span}\{|n, \vec{k}\rangle : E - E_n \leq \epsilon(\vec{k}) \leq E - E_n + \delta\}, \quad (10)$$

represents the subspace of  $H_S + H_B$  associated with the subenergy shell  $[E - E_n, E - E_n + \delta]$  when the system state is fixed at  $|n\rangle$ . Clearly,

$$V(E, \delta) = \sum_n \bigoplus V_n(E - E_n, \delta).$$

When the bath is finite and the system-bath interaction is small but not negligible, the reduced state of the system can be obtained by treating  $H_I$  as a perturbation to  $H_0$ , i.e. by writing  $H = H_0 + H_I$ . In the basis  $\{|n, k_j\rangle\}$ , the off-diagonal matrix elements of  $H_I$  are  $\langle n, k_j | H_I | n', k'_j \rangle = g_{nk_j, n'k'_j}$  with  $n \neq n'$ . From time-independent perturbation theory, the first-order correction to the eigenstate  $|n, k_j\rangle$  contains contributions

$$c_{(n'k'_j)(nk_j)} = \frac{\langle n', k'_j | H_I | n, k_j \rangle}{(E_n + \epsilon_{k_j}) - (E_{n'} + \epsilon_{k'_j})} = \frac{g_{n'k'_j, nk_j}}{E_n - E_{n'} + \epsilon_{k_j} - \epsilon_{k'_j}}, \quad (11)$$

so that

$$|c_{(n'k'_j)(nk_j)}| = \left| \frac{g_{nk_j, n'k'_j}}{E_{n'} - E_n + \epsilon_{k'_j} - \epsilon_{k_j}} \right|. \quad (12)$$

The first-order effect of these off-diagonal terms can be neglected provided

$$\left| \frac{g_{nk_j, n'k'_j}}{E_{n'} - E_n + \epsilon_{k'_j} - \epsilon_{k_j}} \right| \ll 1, \quad (13)$$

which is satisfied for  $n \neq n'$  because the system level spacings  $|E_{n'} - E_n|$  are large (Eq. (3)). However, for  $n = n'$  the denominator reduces to  $\epsilon_{k'_j} - \epsilon_{k_j}$ . Since the bath has a very dense spectrum, the difference  $|\epsilon_{k'_j} - \epsilon_{k_j}|$  can be extremely small. In this case the ratio

$$\left| \frac{g_{nk_j, nk'_j}}{\epsilon_{k'_j} - \epsilon_{k_j}} \right|$$

is not necessarily  $\ll 1$ , and the diagonal part of  $H_I$  cannot be ignored. Retaining only these diagonal terms in  $n$  leads to the approximate form

$$H_I \approx \sum_{j=1}^N \sum_{n, k_j, k'_j} g_{nk_j, nk'_j} |n, k_j\rangle \langle n, k'_j| \otimes \prod_{i \neq j} I_i, \quad (14)$$

which is the starting point for constructing the effective Hamiltonian used in the noncanonical statistics.

We now express the effective Hamiltonian in the system eigenbasis  $\{|n\rangle\}$ , since we are primarily interested in how the bath shifts the system energy levels. The bath influences the system through these level shifts, so it is sufficient to track how they modify the system states, rather than the individual bath states. For each bath particle  $j$ , the Hamiltonian  $H_B$  acts only on the bath part of the systembath state and leaves the system part unchanged. In the basis  $\{|n\rangle\}$  we can therefore write  $H_B$  as

$$\begin{aligned} H_B &= \mathbb{I}_S \otimes \left( \sum_{j=1}^N \sum_{k_j} \epsilon_{k_j} |k_j\rangle\langle k_j| \otimes \prod_{i \neq j} I_i \right) \\ &= \sum_n |n\rangle\langle n| \otimes \left( \sum_{j=1}^N \sum_{k_j} \epsilon_{k_j} |k_j\rangle\langle k_j| \otimes \prod_{i \neq j} I_i \right). \end{aligned} \quad (15)$$

Moreover,

$$|n, k_j\rangle = |n\rangle \otimes |k_j\rangle \implies |n, k_j\rangle\langle n, k'_j| = (|n\rangle\langle n|) \otimes (|k_j\rangle\langle k'_j|).$$

Using this, Eq. (14) can be written as

$$H_I \approx \sum_n |n\rangle\langle n| \otimes \left( \sum_{j=1}^N \sum_{k_j, k'_j} g_{nk_j, nk'_j} |k_j\rangle\langle k'_j| \otimes \prod_{i \neq j} I_i \right). \quad (16)$$

Adding these contributions, the effective Hamiltonian becomes

$$\begin{aligned} H_{\text{eff}} &= H_0 + H_I = H_S + H_B + H_I \\ &= \sum_n |n\rangle\langle n| \otimes \left( E_n + \sum_{j=1}^N \sum_{k_j} \epsilon_{k_j} |k_j\rangle\langle k_j| + \sum_{j=1}^N \sum_{k_j, k'_j} g_{nk_j, nk'_j} |k_j\rangle\langle k'_j| \right) \otimes \prod_{i \neq j} I_i \\ &= \sum_n |n\rangle\langle n| \otimes \left( E_n + \sum_{j=1}^N \sum_{k_j, k'_j} \left[ \epsilon_{k_j} \delta_{k_j, k'_j} + g_{nk_j, nk'_j} \right] |k_j\rangle\langle k'_j| \otimes \prod_{i \neq j} I_i \right). \end{aligned} \quad (17)$$

Thus the total effective Hamiltonian is diagonal with respect to the system eigenstates,

$$H_{\text{eff}} = \sum_n (E_n + h(n)) |n\rangle\langle n|, \quad (18)$$

where the operator  $h(n)$  acts on the bath Hilbert space and is given by

$$h(n) = \sum_{j=1}^N \sum_{k_j, k'_j} \left[ \epsilon_{k_j} \delta_{k_j, k'_j} + g_{nk_j, nk'_j} \right] |k_j\rangle\langle k'_j| \otimes \prod_{i \neq j} I_i. \quad (19)$$

Here  $h(n)$  describes the effective bath Hamiltonian conditioned on the system being in the state  $|n\rangle$ . It can be further diagonalised as

$$h(n) = \sum_{j=1}^N \sum_{\alpha_j} \varepsilon_{\alpha_j}(n) |\alpha_j\rangle \langle \alpha_j| \otimes \prod_{i \neq j}^N I_i, \quad (20)$$

with the eigenstate

$$|\vec{\alpha}\rangle = \prod_{j=1}^N |\alpha_j\rangle \quad (21)$$

and the corresponding eigenenergy

$$\varepsilon(\vec{\alpha}, n) = \sum_{j=1}^N \varepsilon_{\alpha_j}(n). \quad (22)$$

The new bath spectrum  $\varepsilon(\vec{\alpha}, n)$  is therefore modified relative to the original spectrum  $\epsilon(\vec{k})$  due to the systembath coupling.

This system-bath coupling is usually negligible when we study the thermalized state of the system in a large heat bath. However, in the situation of finite system statics, it is crucial to consider the system-bath interaction energy for the relatively small heat bath. The effective Hamiltonian Eq. (18) is already diagonalized with diagonal elements  $E(n, \vec{\alpha}) = E_n + \varepsilon(\vec{\alpha}, n)$ . Then it is straightforward to define the deformed subspace

$$\tilde{V}(E, \delta) = \text{Span}\{|n, \vec{\alpha}(n)\rangle : E \leq E_n + \varepsilon(\vec{\alpha}, n) \leq E + \delta\}, \quad (23)$$

which plays the same role as  $V(E, \delta)$ , with  $|n, \vec{k}\rangle$  replaced by  $|n, \vec{\alpha}(n)\rangle$  and  $\epsilon(\vec{k})$  replaced by  $\varepsilon(\vec{\alpha}, n)$ . In the same way, we define

$$\tilde{V}(E - E_n, \delta) = \text{Span}\{|n, \vec{\alpha}(n) : E - E_n \leq \varepsilon(\vec{\alpha}, n) \leq E - E_n + \delta\}, \quad (24)$$

and  $\tilde{\Omega}(E - E_n, \delta) = \dim \tilde{V}_n(E - E_n, \delta)$ . The authors schematically illustrate the deformation of the subenergy shell in Fig. 2 (inserted from the original paper). By analogy with the canonical construction, the deformed microcanonical state is

$$\tilde{\rho}_{\text{mic}} = \frac{P_{\tilde{V}(E, \delta)}}{\dim \tilde{V}(E, \delta)}, \quad \tilde{\rho}_S = \text{Tr}_B(\tilde{\rho}_{\text{mic}}). \quad (25)$$

where  $P_{\tilde{V}(E, \delta)}$  is the projector onto the deformed energy shell,

$$P_{\tilde{V}(E, \delta)} = \sum_{\substack{n, \vec{\alpha}(n) \\ E \leq E_n + \varepsilon(\vec{\alpha}, n) \leq E + \delta}} |n, \vec{\alpha}(n)\rangle \langle n, \vec{\alpha}(n)|. \quad (26)$$

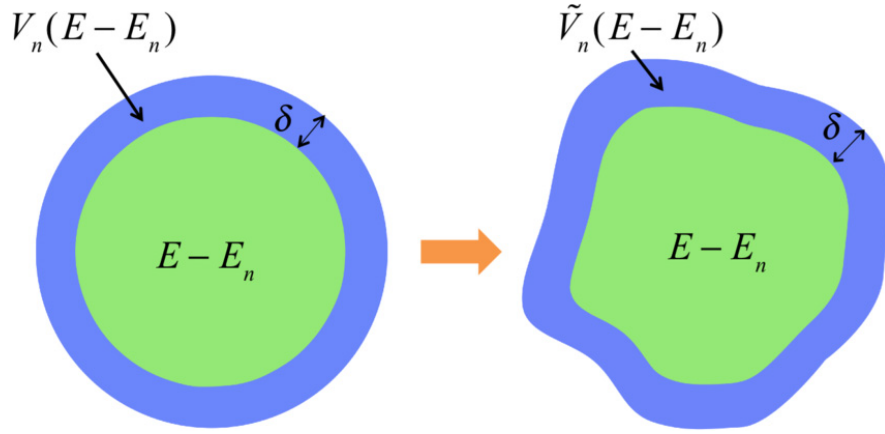


Figure 2: The subenergy shell without and with the consideration of the system-bath coupling are schematically shown in a manner analogous to classical phase space. This image was taken from the original paper.

Substituting this into Eq. (25) yields

$$\begin{aligned}
 \tilde{\rho}_S &= \text{Tr}_B(\tilde{\rho}_{\text{mic}}) \\
 &= \frac{1}{\dim \tilde{V}(E, \delta)} \sum_{\substack{n, \vec{\alpha}(n) \\ E \leq E_n + \varepsilon(\vec{\alpha}, n) \leq E + \delta}} \text{Tr}_B(|n, \vec{\alpha}(n)\rangle \langle n, \vec{\alpha}(n)|) \\
 &= \sum_n \frac{\tilde{\Omega}(E - E_n, \delta)}{\dim \tilde{V}(E, \delta)} |n\rangle \langle n|.
 \end{aligned} \tag{27}$$

Thus, the probability of finding the system in state  $|n\rangle$  is

$$\tilde{P}_n(E) = \langle n | \tilde{\rho}_S | n \rangle = \frac{\tilde{\Omega}(E - E_n, \delta)}{\dim \tilde{V}(E, \delta)}. \tag{28}$$

The matrix  $\tilde{\rho}_S$  is therefore the noncanonical reduced density matrix of the system, and  $\tilde{P}_n(E)$  is the corresponding noncanonical distribution that incorporates finite system-bath coupling.

To further obtain an explicit expression of  $\tilde{P}_n(E)$ , we introduce the entropy defined by  $S(E) = k_B \ln(\Omega(E, s))$ , where  $k_B$  is the Boltzmann constant. Note that the functional form of  $\Omega(E, \delta)$  depends on the underlying Hamiltonian  $H$ . In particular, including or excluding the interaction  $H_I$  leads to different densities of states, so that the undeformed and deformed energy shells correspond to  $\Omega(E, \delta)$  and  $\tilde{\Omega}(E, \delta)$ , respectively. Therefore, it is useful to regard  $S(E)$  as a functional of the chosen energy shell  $V(E, \delta)$ , and write  $S(E) \equiv S(V(E, \delta))$ . When



the system is in the state  $|n\rangle$ , the entropy can be expanded as,

$$\tilde{S}(E - E_n) = S(E - E_n) + \delta_n S(E - E_n) \quad (29)$$

$$\begin{aligned} &\approx S(E) - \frac{\partial S(E)}{\partial E} E_n + \delta_n S(E) - \frac{\partial[\delta_n S(E)]}{\partial E} E_n \\ &\approx S(E) - \frac{\partial S(E)}{\partial E} E_n + \phi_n, \end{aligned} \quad (30)$$

where

$$\phi_n = \delta_n S(E) - \frac{\partial[\delta_n S(E)]}{\partial E} E_n. \quad (31)$$

The term  $\delta_n S(E - E_n)$  is the functional variation of  $S(E - E_n)$ , where the subscript  $n$  means that the functional variation also depends on the state of the system. In the above derivation, we implicitly assume that the deformation of the energy shell is small; that is, for large  $N$  the difference between  $\Omega(E - E_n, \delta)$  and  $\tilde{\Omega}(E - E_n, \delta)$  is taken to be small. To see why this is a reasonable assumption, it is helpful to sketch how this works in a concrete model. Xu and Sun point to the example of Dong *et al.*, “Quantum thermalization with couplings”<sup>1</sup>, where the difference between the deformed and undeformed densities of states vanishes in the thermodynamic limit.

In the model of Dong *et al.*, the number of bath microstates compatible with the system level  $|n\rangle$  scales as

$$\Omega_N(n) \propto [E - \epsilon_n(\kappa)]^{N-1}, \quad (32)$$

where

$$\epsilon_n(\kappa) = \epsilon_n - \kappa \lambda_n^2 \quad (33)$$

is the interaction-shifted system energy. This can be written as

$$\Omega_N(n) \propto \exp\left[(N-1) \ln(E - \epsilon_n(\kappa))\right]. \quad (34)$$

Substituting  $\epsilon_n(\kappa)$  and rearranging gives

$$E - \epsilon_n(\kappa) = E + \kappa \lambda_n^2 - \epsilon_n, \quad (35)$$

so that

$$\ln(E - \epsilon_n(\kappa)) = \ln(E + \kappa \lambda_n^2 - \epsilon_n). \quad (36)$$

For a large bath with weak coupling, we have

$$\frac{\epsilon_n}{E + \kappa \lambda_n^2} \ll 1, \quad (37)$$

---

<sup>1</sup>H. Dong, S. Yang, X. F. Liu, and C. P. Sun, [Phys. Rev. A \*\*76\*\*, 044104 \(2007\)](#).

and we may expand the logarithm to first order:

$$\ln(E + \kappa\lambda_n^2 - \epsilon_n) \simeq \ln(E + \kappa\lambda_n^2) - \frac{\epsilon_n}{E + \kappa\lambda_n^2}. \quad (38)$$

Inserting this back into  $\Omega_N(n)$  yields

$$\begin{aligned} \Omega_N(n) &\propto \exp\left[(N-1)\ln(E + \kappa\lambda_n^2)\right] \exp\left[-(N-1)\frac{\epsilon_n}{E + \kappa\lambda_n^2}\right] \\ &\equiv (\text{slowly varying prefactor}) \times \exp(-\beta_n \epsilon_n), \end{aligned} \quad (39)$$

with

$$\beta_n = \frac{N-1}{E + \kappa\lambda_n^2}. \quad (40)$$

In the large-bath, weak-coupling regime ( $N$  large,  $E \sim N$ ,  $\kappa\lambda_n^2 \ll E$ ), the quantity  $\beta_n$  becomes effectively independent of both the level index  $n$  and the coupling, and the prefactor depends only weakly on  $n$ . The resulting diagonal distribution  $\tilde{P}_n$  is then approximately canonical and essentially insensitive to the shell deformation. Xu and Sun summarise this robustness by stating that the difference between  $\Omega$  and  $\tilde{\Omega}$  vanishes in the thermodynamic limit, even though Dong *et al.* do not phrase their result explicitly in these terms.

The derivative terms of  $S(E)$  in Eq. (30) can be evaluated by the idea from classical phase case: since  $\Omega(E)$  is the dimension of the subspace  $V(E, \delta)$ , it is proportional to the volume of the energy shell in phase space. In an  $N$ -dimensional space, the volume confined by an isoenergetic surface of energy  $E$  can be considered as the volume of an  $N$ -dimensional polyhedron with effective radius  $R(E)$ , which is proportional to  $E^\zeta$ , so that the volume is proportional to  $E^{\zeta N}$ , where  $\zeta$  is a dimensionless real number independent of  $N$  and is usually related to the degeneracy of the system. The above analysis is from the original paper,<sup>2</sup> and is quite hand-wavy, so we will try to make sense of this. The phase space is  $N$ -dimensional. We pick the bath coordinates collectively as  $\mathbf{X} \in \mathbb{R}^N$ . The Hamiltonian  $H(\mathbf{X})$  is an energy function on this space. In many physical cases (free particles, oscillators, polynomial potentials), the dominant part of  $H$  is homogeneous of some degree  $\alpha$ , meaning

$$H(\lambda\mathbf{X}) = \lambda^\alpha H(\mathbf{X}).$$

Given this homogeneity, we look at the region

$$\mathcal{D}(E) = \{\mathbf{X} : H(\mathbf{X}) \leq E\}.$$

Now define a scaled coordinate

$$\mathbf{X} = E^{1/\alpha} \mathbf{Y}.$$

---

<sup>2</sup>Here the "original paper" refers to "[Noncanonical statistics of a finite quantum system with non-negligible system-bath coupling](#)" by Xu and Sun.

Then

$$H(\mathbf{X}) = H(E^{1/\alpha} \mathbf{Y}),$$

so the condition  $H(\mathbf{X}) \leq E$  becomes  $H(\mathbf{Y}) \leq 1$ . That means the region  $\mathcal{D}(E)$  is just the region  $\mathcal{D}(1)$  scaled by a linear factor  $E^{1/\alpha}$  in all directions. Therefore the “radius” of the region with energy  $E$  is

$$R(E) \propto E^{1/\alpha}.$$

For convenience, define  $\zeta := 1/\alpha$ , so that

$$R(E) \propto E^\zeta.$$

The corresponding density of states then behaves as

$$g(E) \sim \frac{d}{dE} D(E) \propto E^{\zeta N - 1}, \quad D(E) = \text{Vol}\{\mathbf{X} : H(\mathbf{X}) \leq E\}.$$

This scaling makes the connection between the exponent  $\zeta$  and the growth of the degeneracy with energy more transparent.

We also know that in  $N$  dimensions, if we scale all the coordinates by a factor  $R$ , the volume is scaled by  $R^N$ . So the volume of the region below energy  $E$  is

$$D(E) = \int_{\mathcal{D}(E)} d^N \mathbf{X} = \int_{\mathcal{D}(1)} d^N (R(E) \mathbf{Y}) = R(E)^N \int_{\mathcal{D}(1)} d^N \mathbf{Y} = R(E)^N D(1).$$

Therefore, the volume of the energy shell defined for energies in  $[E, E + \delta]$  is given by

$$\Omega(E) \propto (E + \delta)^{\zeta N} - E^{\zeta N}. \quad (41)$$

Now we assume  $\delta \ll E$  (thin shell compared to the total energy, but still containing many states). Then we expand

$$\begin{aligned} f(E) = E^{\zeta N} &\rightarrow f(E + \delta) \approx f(E) + f'(E) \delta = E^{\zeta N} + (\zeta N) E^{\zeta N - 1} \delta \\ \implies (E + \delta)^{\zeta N} - E^{\zeta N} &\approx \zeta N \delta E^{\zeta N - 1}. \end{aligned} \quad (42)$$

Since  $\zeta N$  and  $\delta$  are treated as constants, we can write

$$\Omega(E) \propto E^{\zeta N - 1}. \quad (43)$$

Using  $S(E) = k_B \ln \Omega(E)$ , this gives

$$\beta(E) = \frac{\partial S(E)}{\partial E} = \frac{\partial}{\partial E} [k_B (\zeta N - 1) \ln E] \approx \frac{k_B (\zeta N - 1)}{E}, \quad (44)$$

where we identify  $\beta = (k_B T)^{-1}$  via the thermodynamic relation between entropy and temperature, with  $T$  the equilibrium temperature of the system. Obviously, Eq. (44) recovers the equipartition theorem  $E \approx \zeta N k_B T$ , i.e., each degree of freedom contributes  $\zeta k_B T$  to the total

energy. Using Eq. (30) we write

$$\frac{\tilde{S}(E - E_n)}{k_B} \simeq \frac{S(E)}{k_B} - \beta E_n + \frac{\phi_n}{k_B}.$$

Substituting this into Eq. (28) gives

$$\tilde{P}_n(E) = \frac{e^{S(E)/k_B} e^{-\beta E_n + \phi_n/k_B}}{\sum_m e^{S(E)/k_B} e^{-\beta E_m + \phi_m/k_B}}.$$

The common factor  $e^{S(E)/k_B}$  cancels between numerator and denominator, so the noncanonical distribution takes the form

$$\tilde{P}_n(E) = \frac{1}{\tilde{\mathcal{Z}}} e^{-\beta E_n + \phi_n/k_B}, \quad (45)$$

where  $\tilde{\mathcal{Z}} = \sum \exp(-\beta E_n + \phi_n/k_B)$  is the partition function.

If the interaction energy could be neglected compared with the total energy  $E$ , the variation terms  $\phi_n$  containing  $\delta_n S(E)$  in Eq. (45) can be dropped and thus naturally lead to the usual canonical form,

$$P_n(E) = \frac{1}{\mathcal{Z}} e^{-\beta E_n}. \quad (46)$$

Otherwise, the deformation of the energy shell which leads to  $\phi_n$  will shift the system energy level and thus modify  $P_n(E)$ . Usually  $\phi_n$  cannot be calculated easily; however, for some special models (one of which will be discussed in the next section), we can find a simple expression for  $\phi_n$ .

### 1.3 Illustration with harmonic oscillators system

In this section the authors consider a coupled harmonic oscillators system as an example to illustrate the statistical thermodynamic properties of a finite system. In what follows, we set  $\hbar = 1$ . The system  $S$  is a single harmonic oscillator with eigenfrequency  $\omega$  described by the Hamiltonian

$$H_S = \omega a^\dagger a, \quad (47)$$

where  $a^\dagger$  and  $a$  are the usual creation and annihilation operators. The zero-point energy term  $1/2\omega$  has been omitted since it corresponds to an overall constant shift of the spectrum and does not affect the microcanonical state or the normalized probability distribution  $\tilde{P}_n(E)$ . The heat bath is generally modeled as a collection of harmonic oscillators with Hamiltonian

$$H_B = \sum_{j=1}^N \omega_j b^\dagger b_j. \quad (48)$$

Here  $b_j^\dagger$  ( $b_j$ ) is the creation (annihilation) operator of the oscillator with frequency  $\omega_j$ . The zero-point term  $1/2\omega_j$  has been omitted again for the same reason. We work in the weak-coupling regime, where the system-bath coupling constants  $\lambda_j$  are small compared to the characteristic energy scales of the uncoupled system and bath. Concretely, we assume  $|\lambda_j| \ll \omega$  and  $|\lambda_j| \ll \omega_j$ , so that the interaction energy is much smaller than the level spacings of both  $H_S$  and  $H_B$ . In this regime the coupling primarily induces small level-dependent energy shifts and deformations of the energy shell, without strongly mixing different system eigenstates. The authors assume the effective system-bath interaction

$$H_I = \sum_{j=1}^N \lambda_j a^\dagger a (b_j^\dagger + b_j), \quad (49)$$

which satisfies  $[H_S, H_I] = 0$ . Thus the system number states  $|n\rangle$  remain eigenstates of the full Hamiltonian, and the interaction does not induce transitions between different system levels; it only shifts the energy of each level via the bath. This structure allows the full Hamiltonian to be diagonalized by a displacement of the bath modes.

In the product basis  $|n\rangle \otimes |k_1, \dots, k_N\rangle$ , the free Hamiltonians  $H_S$  and  $H_B$  are diagonal, whereas

$$H_I = \sum_{j=1}^N \lambda_j a^\dagger a (b_j^\dagger + b_j)$$

is off-diagonal in the bath occupation numbers  $k_j$ . Since  $a^\dagger a |n\rangle = n |n\rangle$ , we first fix the system in  $|n\rangle$  and take the matrix element of  $H_B + H_I$ :

$$\begin{aligned} \langle n | H_I | n \rangle &= \sum_j \lambda_j \langle n | a^\dagger a (b_j^\dagger + b_j) | n \rangle \\ &= \sum_j \lambda_j n (b_j^\dagger + b_j), \end{aligned}$$

so the effective bath Hamiltonian in sector  $n$  is

$$H_{\text{bath}}^{(n)} = \langle n | (H_B + H_I) | n \rangle = \sum_{j=1}^N \left[ \omega_j b_j^\dagger b_j + n \lambda_j (b_j^\dagger + b_j) \right], \quad (50)$$

where  $a^\dagger a$  has been replaced by its eigenvalue  $n$ .

Each mode  $j$  is thus a harmonic oscillator with a linear term in  $b_j^\dagger + b_j$ . We diagonalize it by a displacement transformation. For each mode, define

$$c_j := b_j + \alpha_j, \quad c_j^\dagger := b_j^\dagger + \alpha_j^*. \quad (51)$$

Substituting  $b_j = c_j - \alpha_j$  into  $H_{\text{bath}}^{(n)}$ , we get

$$\begin{aligned}\omega_j b_j^\dagger b_j &= \omega_j (c_j^\dagger - \alpha_j^*) (c_j - \alpha_j) = \omega_j c_j^\dagger c_j - \omega_j \alpha_j c_j^\dagger - \omega_j \alpha_j^* c_j + \omega_j |\alpha_j|^2, \\ n\lambda_j (b_j^\dagger + b_j) &= n\lambda_j (c_j^\dagger + c_j) - n\lambda_j (\alpha_j + \alpha_j^*).\end{aligned}$$

Therefore,

$$H_{\text{bath}}^{(n)} = \sum_{j=1}^N \left[ \omega_j c_j^\dagger c_j + (-\omega_j \alpha_j + n\lambda_j) c_j^\dagger + (-\omega_j \alpha_j^* + n\lambda_j) c_j + \omega_j |\alpha_j|^2 - n\lambda_j (\alpha_j + \alpha_j^*) \right]. \quad (52)$$

We choose  $\alpha_j$  so that the linear terms vanish:

$$-\omega_j \alpha_j + n\lambda_j = 0 \quad \Rightarrow \quad \alpha_j = \frac{n\lambda_j}{\omega_j}.$$

The second condition  $-\omega_j \alpha_j^* + n\lambda_j = 0$  is then automatically satisfied, showing that  $\alpha_j$  is real in this model. The constant term becomes

$$\omega_j |\alpha_j|^2 - n\lambda_j (\alpha_j + \alpha_j^*) = \frac{n^2 \lambda_j^2}{\omega_j} - \frac{2n^2 \lambda_j^2}{\omega_j} = -\frac{n^2 \lambda_j^2}{\omega_j}.$$

Hence

$$H_{\text{bath}}^{(n)} = \sum_{j=1}^N \left( \omega_j c_j^\dagger c_j - \frac{n^2 \lambda_j^2}{\omega_j} \right), \quad (53)$$

which is diagonal in the number basis  $\{|k_j\rangle_c\}$  of the displaced oscillators.

Adding back the system Hamiltonian  $H_S = \omega a^\dagger a$ , we obtain

$$H^{(n)} = \omega n + \sum_{j=1}^N \left( \omega_j c_j^\dagger c_j - \frac{n^2 \lambda_j^2}{\omega_j} \right), \quad (54)$$

so the eigenenergies are

$$E_{n, \{k_j\}} = \omega n + \sum_{j=1}^N \omega_j k_j - \sum_{j=1}^N \frac{n^2 \lambda_j^2}{\omega_j}. \quad (55)$$

This result can be directly matched to the notation in the original paper. There, the bath modes are diagonalized by the displacement operator

$$D(\alpha_j) = \exp(\alpha_j b_j^\dagger - \alpha_j^* b_j), \quad (56)$$

leading to displaced Fock states

$$|m_j(n)\rangle = D\left(-\frac{\lambda_j n}{\omega_j}\right)|m_j\rangle. \quad (57)$$

This is equivalent to our approach; instead of introducing displaced operators  $c_j = b_j + \alpha_j$ , one can keep  $b_j$  fixed and act with the unitary displacement operator  $D(-n\lambda_j/\omega_j)$  on the bath Fock states. Displacing the operators and displacing the states are two aspects of the same unitary transformation.

Matching our notation to that of the paper, Eq. (55) for the eigenvalues of the total Hamiltonian becomes

$$E(n, \vec{m}(n)) = n\omega + \Delta_n + \sum_{j=1}^N m_j \omega_j, \quad (58)$$

corresponding to the eigenstates

$$|n, \vec{m}(n)\rangle = |n\rangle \otimes \prod_{j=1}^N |m_j(n)\rangle. \quad (59)$$

Thus the level-dependent energy shift obtained above,

$$\Delta_n = -\sum_{j=1}^N \frac{n^2 \lambda_j^2}{\omega_j}, \quad (60)$$

coincides with Eq. (24) of the paper, where they write

$$\Delta_n = -\kappa n^2, \quad \kappa := \sum_{j=1}^N \frac{\lambda_j^2}{\omega_j}. \quad (61)$$

Therefore, in this model the deformation of each subenergy shell is only related to  $\Delta_n$ , which is independent of the specific state of the heat bath. Then the deformed energy shell is simply related with the original energy shell without considering  $H_I$  by

$$\tilde{V}_n(E - E_n, \delta) = V_n(E - \Delta_n - E_n, \delta). \quad (62)$$

Furthermore, the functional form of  $\tilde{S}(E - E_n)$  is the same as that of  $S(E - E_n)$ , with the replacement  $E \rightarrow E - \Delta_n$ :

$$\tilde{S}(E - E_n) = S(E - \Delta_n - E_n). \quad (63)$$

For small  $|\Delta_n|$ , a first-order Taylor expansion in  $\Delta_n$  gives

$$S(E - \Delta_n - E_n) \approx S(E - E_n) - \left. \frac{\partial S}{\partial E} \right|_{E=E_n} \Delta_n. \quad (64)$$

Using Eq. (29), we have

$$\delta_n S(E - E_n) = S(E - \Delta_n - E_n) - S(E - E_n), \quad (65)$$

and therefore obtain

$$\delta_n S(E - E_n) \approx - \left. \frac{\partial S}{\partial E} \right|_{E-E_n} \Delta_n. \quad (66)$$

Since the bath is large, the entropy gradient varies slowly with energy, and one may approximate  $\left. \frac{\partial S}{\partial E} \right|_{E-E_n} \approx \left. \frac{\partial S}{\partial E} \right|_E \equiv \beta(E)$ . Thus

$$\delta_n S(E) \approx - \left. \frac{\partial S}{\partial E} \right|_E \Delta_n. \quad (67)$$

(Note that Eq. (27) in the original paper appears to miss a minus sign.)

Differentiating  $\delta_n S(E)$  with respect to  $E$  gives

$$\frac{\partial \delta_n S(E)}{\partial E} \approx - \frac{\partial^2 S(E)}{\partial E^2} \Delta_n.$$

Plugging this into  $\phi_n(E)$ ,

$$\begin{aligned} \phi_n &= \delta_n S(E) - \frac{\partial \delta_n S(E)}{\partial E} E_n \\ &\approx - \frac{\partial S(E)}{\partial E} \Delta_n + \frac{\partial^2 S(E)}{\partial E^2} \Delta_n E_n = -\beta(E) \Delta_n + \xi \Delta_n E_n, \end{aligned} \quad (68)$$

where, using Eq. (44),

$$\xi = \frac{\partial^2 S(E)}{\partial E^2} = - \frac{k_B(\zeta N - 1)}{E^2}. \quad (69)$$

For the harmonic oscillator model we have

$$\Delta_n = -\kappa n^2, \quad E_n = n\omega,$$

so that

$$\phi_n = \beta(E) \kappa n^2 - \xi \omega \kappa n^3.$$

Substituting this into Eq. (45) for  $\tilde{P}_n(E)$  yields

$$\tilde{P}_n(E) = \frac{1}{\mathcal{Z}} e^{-\beta n \omega + \beta \kappa n^2 - \xi \omega \kappa n^3} = \frac{1}{\mathcal{Z}} e^{-\beta(n\omega - n^2 \kappa) - \xi \omega \kappa n^3}. \quad (70)$$

The square and cubic terms in  $n$  in the exponent of  $\tilde{P}_n(E)$  strongly modify the distribution compared with the canonical form  $P_n(E) = \exp(-\beta n \omega) / \mathcal{Z}$ , especially for large  $n$ . However, Eq. (70) does not apply to arbitrarily large  $N$  for the following reason: the total energy  $E$



of the system plus bath is conserved, and the bath, being a collection of harmonic modes, has energy  $\sum_{\{m_j\}} m_j \omega_j \geq 0$ , so its energy must be non-negative,  $E_{\text{bath}} \geq 0$ . On the deformed energy shell for level  $n$ , the bath energy is

$$E_{\text{bath}} = E - E_n - \Delta_n = E - n\omega + \kappa n^2,$$

where  $E_n = n\omega$  and  $\Delta_n = -\kappa n^2$ . Thus

$$E - n\omega + \kappa n^2 \geq 0 \implies \kappa n^2 - \omega n + E \geq 0. \quad (71)$$

Solving the corresponding quadratic equation

$$\kappa n^2 - \omega n + E = 0$$

gives

$$n_{\pm} = \frac{\omega \pm \sqrt{\omega^2 - 4\kappa E}}{2\kappa}. \quad (72)$$

For an upward parabola ( $\kappa > 0$ ), the inequality is satisfied outside the interval  $[n_-, n_+]$ . In the regime where the noncanonical distribution is only a small deformation of the canonical one, physically relevant levels are low-lying, so we restrict to  $n \leq n_-$ . This yields an upper bound on the occupied system levels,

$$n < \frac{\omega - \sqrt{\omega^2 - 4\kappa E}}{2\kappa}. \quad (73)$$

Therefore, for such a coupled system, the maximum system level is

$$n_{\max} = \left\lfloor \frac{\omega - \sqrt{\omega^2 - 4\kappa E}}{2\kappa} \right\rfloor, \quad (74)$$

where  $\lfloor x \rfloor$  denotes the largest integer not exceeding  $x$ . Hence, Eq. (70) cannot be extended to arbitrarily large  $n$  beyond this constraint.

We can check the weak-coupling limit to see that this bound is sensible. Expanding for small  $\kappa$ , we find

$$\sqrt{\omega^2 - 4\kappa E} = \omega \sqrt{1 - \frac{4\kappa E}{\omega^2}} \approx \omega \left( 1 - \frac{2\kappa E}{\omega^2} \right) = \omega - \frac{2\kappa E}{\omega}.$$

Then

$$n_{\max} \approx \frac{\omega - \left( \omega - \frac{2\kappa E}{\omega} \right)}{2\kappa} = \frac{E}{\omega}. \quad (75)$$

In the weak-coupling limit we thus recover the obvious bound that the system cannot carry

more energy than the total energy  $E$ :

$$E_n = n\omega \leq E \implies n \leq \frac{E}{\omega}.$$

For low energy levels ( $n \lesssim 20$ ), the noncanonical distribution  $\tilde{P}_n(E)$  does not differ much from the canonical one, as shown in Fig. 3, taken from the paper.

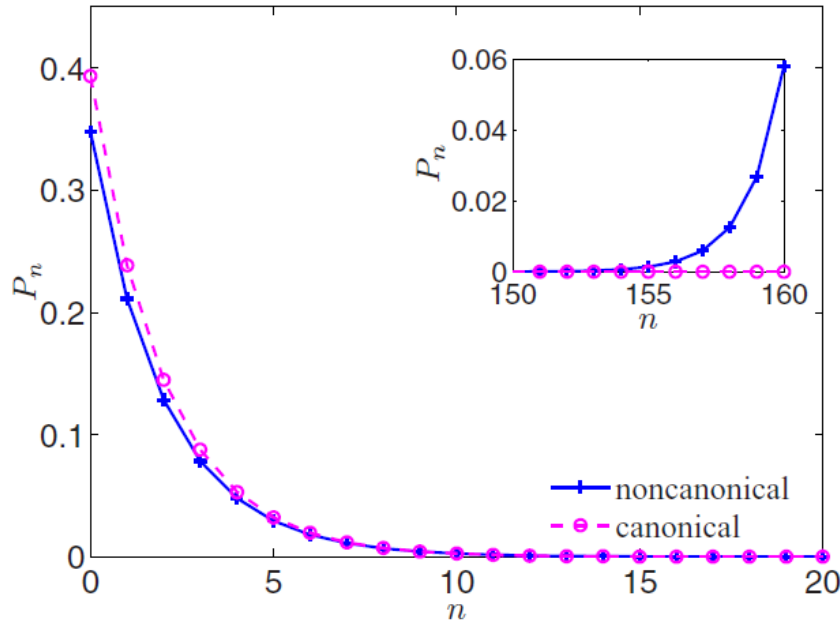


Figure 3: Noncanonical (blue solid line) and canonical (pink dashed line) distribution functions for the coupled oscillator system. The authors choose  $\omega = 1$ ,  $E = 100$ ,  $\beta = 0.5$ , and  $\kappa = 0.00235$ , for which the highest energy level is  $n_{\max} = 160$ . The noncanonical and canonical distributions are very similar for low energy levels, but differ significantly for high energy levels.

The figure, generated with  $\omega = 1$ , the total energy  $E = 100$ , and  $\kappa = 0.00235$ , shows a system where the highest energy level is  $n_{\max} = 160$  according to Eq. (74). The canonical distribution is plotted by setting  $\kappa = 0$  in  $\tilde{P}_n(E)$ .

From Eq. (43),

$$S(E) = \ln(\Omega(E)) = (\zeta N - 1) \ln(E), \quad (76)$$

it follows that

$$\beta = \frac{\partial S(E)}{\partial E} > 0 \quad \rightarrow \quad \xi = \frac{\partial^2 S(E)}{\partial E^2} < 0.$$

Consequently, the extra quadratic and cubic terms in the exponent of (70) become positive at large  $n$ . These noncanonical corrections therefore push the probability upward, partially canceling the usual  $-\beta\omega n$  decay. Since  $|\Delta_n|$  grows with  $n^2$ , the high energy levels gain more population than in the canonical ensemble, as visible in the inset of Fig. 3.

The internal energy of the system is defined as

$$U = \langle H_S \rangle = \text{Tr} (\tilde{\rho}_S H_S). \quad (77)$$

Putting the noncanonical density matrix from Eq. (29),

$$\tilde{\rho}_S = \sum_n \tilde{P}_n(E) |n\rangle \langle n|,$$

and the harmonic oscillator Hamiltonian

$$H_S = \sum_n E_n |n\rangle \langle n|, \quad E_n = n\omega,$$

together we obtain

$$\begin{aligned} U &= \text{Tr} \left( \sum_n \tilde{P}_n(E) |n\rangle \langle n| \cdot \sum_m E_m |m\rangle \langle m| \right) = \sum_{n,m} \tilde{P}_n(E) E_m \text{Tr} (|n\rangle \langle n| m\rangle \langle m|) \\ &= \sum_n \tilde{P}_n(E) E_n = \sum_n n\omega \tilde{P}_n(E). \end{aligned} \quad (78)$$

Because the high-energy states have relatively larger populations, the internal energy of the system under the noncanonical statistics is larger than that under the canonical one. This fact is illustrated in Fig. 4, where the internal energy  $U$  is plotted with respect to  $\beta$ . The distinction between noncanonical and canonical results for  $U$  evidently appears when the inverse temperature  $\beta$  decreases and the interaction energy strength  $\kappa$  grows. As  $\beta$  approaches zero, the terms proportional to  $\beta$  in (70) vanish. Additionally,  $\xi = \partial^2 S / \partial E^2 \propto 1/E^2$ , and since the total energy  $E$  of the system and the bath becomes large at high temperature,  $\xi$  tend to zero. Thus, the entire exponent approaches zero, so:

$$\tilde{P}_n(E) \propto 1. \quad (79)$$

The partition function  $\mathcal{Z}$  then simply counts the number of allowed states,

$$\begin{aligned} \mathcal{Z} &= \sum_{n=0}^{n_{\max}} 1 = n_{\max} + 1, \\ \implies \tilde{P}_n(E) &= \frac{1}{n_{\max} + 1}. \end{aligned} \quad (80)$$

Substituting this into Eq. (78) gives

$$U = \frac{\omega}{n_{\max} + 1} \sum_{n=0}^{n_{\max}} n = \frac{\omega}{n_{\max} + 1} \times \frac{n_{\max}(n_{\max} + 1)}{2} = \frac{\omega n_{\max}}{2}. \quad (81)$$

Therefore, the high-temperature limit of the internal energy  $U$  is  $n_{\max}\omega/2$ , which is finite as the total energy is upper-bounded by  $E$  for a small heat bath. This differs substantially from

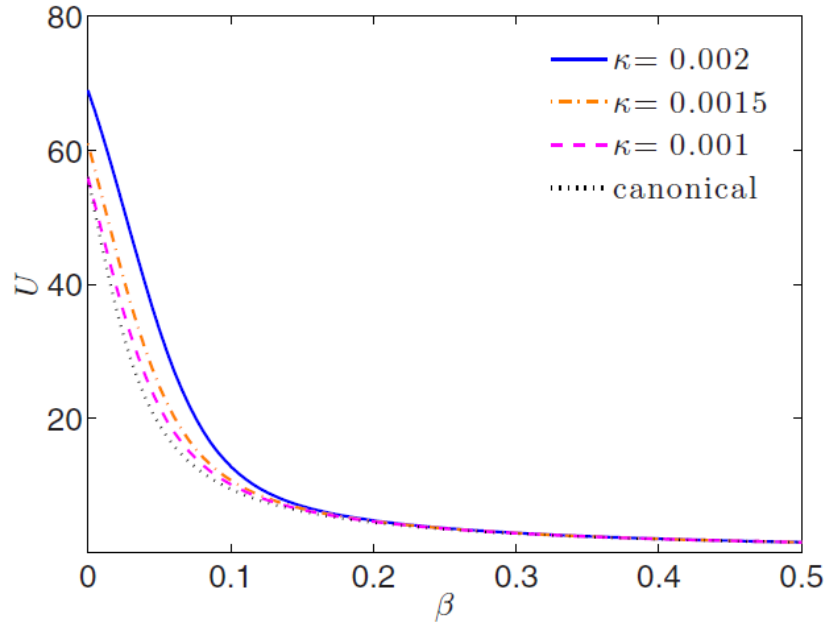


Figure 4: The authors plot the internal energy  $U$  versus  $\beta$  for both noncanonical and canonical statistics, with  $\omega = 1$  and  $E = 100$ . Noncanonical cases are shown for several coupling strengths:  $\kappa = 0.002$  ( $n_{\max} = 138$ ),  $\kappa = 0.0015$  ( $n_{\max} = 122$ ), and  $\kappa = 0.001$  ( $n_{\max} = 112$ ). These are compared to the canonical case ( $\kappa = 0$ ,  $n_{\max} = 122$ ). The plot reveals that for small  $\beta$ , the internal energy  $U$  in the noncanonical statistics is significantly larger than that predicted by the canonical distribution.

the case in the thermodynamics limit: the average energy of a harmonic oscillator which is in contact with an infinite heat bath will diverge when  $\beta$  decreases to zero.

Fig. 4 illustrates this behavior for a fixed  $\omega = 1$  and  $E = 100$ . Different coupling strength  $\kappa$  yield different maximum levels  $n_{\max}$ :

- $\kappa = 0.002$ :

$$n_{\max} = \left\lfloor \frac{1 - \sqrt{1 - 4 \times 0.002 \times 100}}{2 \times 0.002} \right\rfloor = 138$$

$$\Rightarrow U_{\max} = \frac{1 \times 138}{2} = 69$$

- $\kappa = 0.0015$ :

$$n_{\max} = \left\lfloor \frac{1 - \sqrt{1 - 4 \times 0.0015 \times 100}}{2 \times 0.0015} \right\rfloor = 122$$

$$\Rightarrow U_{\max} = \frac{1 \times 122}{2} = 61$$

- $\kappa = 0.001$ :

$$n_{\max} = \left\lfloor \frac{1 - \sqrt{1 - 4 \times 0.001 \times 100}}{2 \times 0.001} \right\rfloor = 112$$

$$\Rightarrow U_{\max} = \frac{1 \times 112}{2} = 56$$

As shown, the noncanonical internal energy is significantly larger at small  $\beta$ .

Another feature reflecting the nonmonotony of the noncanonical distribution is the relative fluctuation of the system's internal energy. The variance of the system energy is

$$\text{Var}(U) = \langle (E_n - U)^2 \rangle = \sum_{n=0}^{n_{\max}} (E_n - U)^2 \tilde{P}_n(E) = \sum_{n=0}^{n_{\max}} (n\omega - U)^2 \tilde{P}_n(E).$$

and the relative fluctuation is defined as

$$\Delta U = \frac{\sqrt{\text{Var}(U)}}{U},$$

so that

$$(\Delta U)^2 = \frac{1}{U^2} \sum_{n=0}^{n_{\max}} (n\omega - U)^2 \tilde{P}_n(E). \quad (82)$$

As shown in Fig. 5, at both the low- and high-temperature limits, the noncanonical and canonical statistics of the system present similar fluctuation behavior. To characterize these two limits, we consider the system as a harmonic oscillator with truncated energy levels (the highest energy level is labeled by  $n_{\max}$ ) under canonical statistics ( $\kappa = 0$ ), whose energy fluctuation is denoted as  $(\Delta U_C)^2$ .

The limit  $n_{\max} \rightarrow \infty$  corresponds to the standard canonical ensemble. Comparing the truncated system (finite bath) with the infinite one, we can identify when finite-bath effects become relevant.

At very low temperature ( $\beta\omega \gg 1$ ), the Boltzmann factor  $e^{-\beta n\omega}$  decays rapidly. The ground state ( $n = 0$ ) dominates, the first excited state ( $n = 1$ ) is suppressed by a factor  $e^{-\beta\omega}$ , and higher levels are negligible. Thus, only the lowest one or two energy levels contribute significantly, and the system behaves like a two-level system. In this approximation, the mean energy is

$$U_C = \sum_{n=0}^1 E_n \cdot P_n = 0 \cdot P_0 + \omega \cdot P_1 = \omega P_1,$$

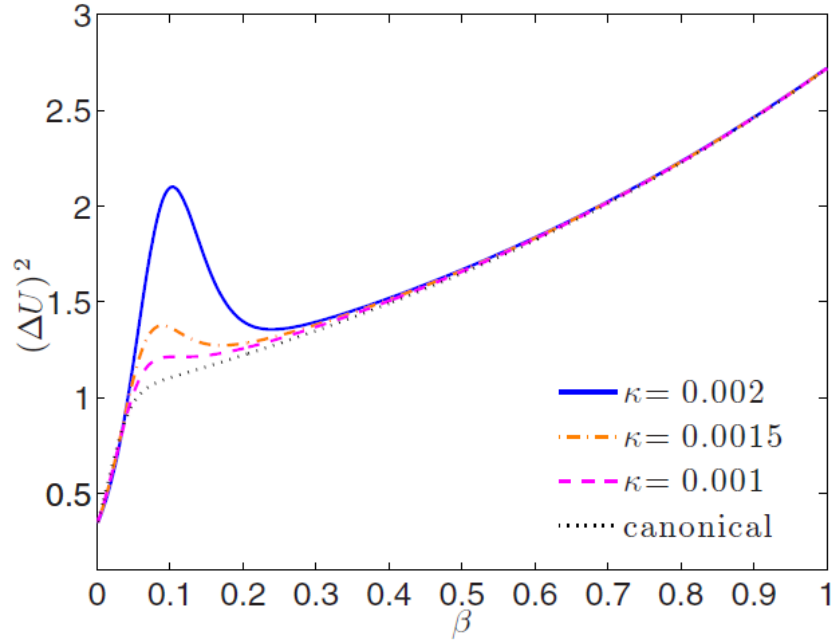


Figure 5: The internal energy relative fluctuation  $(\Delta U)^2$  of the system with respect to  $\beta$  for noncanonical and canonical statistics. The parameters are chosen the same as Fig. 4. A local maximum appears in the energy fluctuation for small  $\beta$  which results of the noncanonical statistics.

where the probabilities are

$$P_0 = \frac{1}{1 + e^{-\beta\omega}},$$

$$P_1 = \frac{e^{-\beta\omega}}{1 + e^{-\beta\omega}} \approx e^{-\beta\omega} \quad (\text{since } e^{-\beta\omega} \ll 1).$$

The variance is

$$\begin{aligned} \text{Var}(U_C) &= \sum_{n=0}^1 (E_n - U)^2 P_n = (0 - \omega P_1)^2 P_0 + (\omega - \omega P_1)^2 P_1 \\ &= \omega^2 [P_1^2 P_0 + (1 - P_1)^2 P_1] = \omega^2 [P_1^2 (1 - P_1) + (1 - P_1)^2 P_1] \\ &= \omega^2 P_1 (1 - P_1) [P_1 + (1 - P_1)] = \omega^2 P_1 (1 - P_1) \\ &= \omega^2 e^{-\beta\omega} (1 - e^{-\beta\omega}) \approx \omega^2 e^{-\beta\omega} \end{aligned}$$

The relative fluctuation then becomes

$$(\Delta U_C)^2 = \frac{\text{Var}(U_C)}{U_C^2} \approx \frac{\omega^2 P_1 (1 - P_1)}{\omega^2 P_1^2} \approx e^{\beta\omega}. \quad (83)$$

To confirm that this result aligns with the canonical low-temperature limit for  $n_{\max} \rightarrow \infty$ ,

we evaluate the full canonical partition function. Since  $\beta\omega > 0$ , then exponential satisfies

$$e^{\beta\omega} > 1 \implies e^{-\beta\omega} < 1.$$

$e^{-\beta\omega}$  is a positive real number, its absolute value is itself, so  $|e^{-\beta\omega}| = e^{-\beta\omega} < 1$ . With this condition we can use the geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad (84)$$

for  $n_{\max} \rightarrow \infty$ :

$$\mathcal{Z} = \sum_{n=0}^{\infty} e^{-\beta\omega n} = \frac{1}{1 - e^{-\beta\omega}}. \quad (85)$$

The mean energy is

$$\langle E \rangle = \frac{1}{\mathcal{Z}} \sum_{n=0}^{\infty} n\omega e^{-\beta\omega n}.$$

Using

$$\sum_{n=0}^{\infty} nx^n = \frac{x}{(1-x)^2}, \quad (86)$$

with  $x = e^{-\beta\omega}$ , we obtain

$$\begin{aligned} \sum_{n=0}^{\infty} ne^{-\beta\omega n} &= \frac{e^{-\beta\omega}}{(1 - e^{-\beta\omega})^2}, \\ \langle E \rangle &= \frac{\omega}{\mathcal{Z}} \times \frac{e^{-\beta\omega}}{(1 - e^{-\beta\omega})^2} = \frac{\omega e^{-\beta\omega}}{1 - e^{-\beta\omega}}. \end{aligned} \quad (87)$$

For the second moment,

$$\langle E^2 \rangle = \frac{1}{\mathcal{Z}} \sum_{n=0}^{\infty} (n\omega)^2 e^{-\beta\omega n} = \frac{\omega^2}{\mathcal{Z}} \sum_{n=0}^{\infty} n^2 e^{-\beta\omega n},$$

and

$$\sum_{n=0}^{\infty} n^2 x^n = \frac{x(1+x)}{(1-x)^3}, \quad (88)$$

gives

$$\langle E^2 \rangle = \frac{\omega^2}{\mathcal{Z}} \times \frac{e^{-\beta\omega}(1 + e^{-\beta\omega})}{(1 - e^{-\beta\omega})^3} = \frac{\omega^2 e^{-\beta\omega}(1 + e^{-\beta\omega})}{(1 - e^{-\beta\omega})^2}. \quad (89)$$

The variance is therefore

$$\begin{aligned}\text{Var}(E) &= \langle E^2 \rangle - \langle E \rangle^2 = \frac{\omega^2}{(1 - e^{-\beta\omega})^2} [e^{-\beta\omega}(1 + e^{-\beta\omega}) - e^{-2\beta\omega}] \\ &= \frac{\omega^2}{(1 - e^{-\beta\omega})^2} e^{-\beta\omega}\end{aligned}$$

Finally, the relative fluctuation is

$$(\Delta U_C)^2 = \frac{\text{Var}(E)}{\langle E \rangle^2} = \frac{\frac{\omega^2 e^{-\beta\omega}}{(1 - e^{-\beta\omega})^2}}{\frac{\omega^2 e^{-2\beta\omega}}{(1 - e^{-\beta\omega})^2}} = \frac{e^{-\beta\omega}}{e^{-2\beta\omega}} = e^{\beta\omega} \quad (90)$$

This analytic result confirms that in the low-temperature limit, the fluctuation  $(\Delta U_C)^2 \approx \exp(\beta\omega)$  is the same whether the spectrum is truncated or infinite. This is very important as it tells us that for sufficiently cold systems, finite baths behave exactly like infinite baths.

Now consider the high-temperature limit. For a harmonic oscillator (setting  $\omega = 1$  for simplicity) truncated at  $n_{\max}$ , the finite geometric series

$$\sum_{k=0}^n x^k = \frac{x^{n+1} - 1}{x - 1}, \quad (91)$$

gives

$$\mathcal{Z} = \sum_{n=0}^{n_{\max}} e^{-\beta n} = \frac{1 - e^{-\beta(n_{\max}+1)}}{1 - e^{-\beta}}. \quad (92)$$

For small  $\beta$  we can expand the numerator and the denominator of  $\mathcal{Z}$ . Using

$$1 - e^{-ax} = ax - \frac{a^2 x^2}{2} + \frac{a^3 x^3}{6} + O(x^4), \quad (93)$$

$$\frac{1}{1 - e^{-x}} = \frac{1}{x} + \frac{1}{2} + \frac{x}{12} + O(x^3), \quad (94)$$

we obtain

$$1 - e^{-\beta(n_{\max}+1)} = \beta(n_{\max} + 1) - \frac{\beta^2(n_{\max} + 1)^2}{2} + \frac{\beta^3(n_{\max} + 1)^3}{6} + O(\beta^4), \quad (95)$$

$$\frac{1}{1 - e^{-\beta}} = \frac{1}{\beta} + \frac{1}{2} + \frac{\beta}{12} + O(\beta^3). \quad (96)$$



Multiplying,

$$\begin{aligned}
 \mathcal{Z} &= \left( \beta(n_{\max} + 1) - \frac{\beta^2(n_{\max} + 1)^2}{2} + \frac{\beta^3(n_{\max} + 1)^3}{6} \right) \left( \frac{1}{\beta} + \frac{1}{2} + \frac{\beta}{12} \right) \\
 &= (n_{\max} + 1) - \frac{n_{\max}(n_{\max} + 1)}{2}\beta + \frac{n_{\max}(n_{\max} + 1)(2n_{\max} + 1)}{12}\beta^2 + O(\beta^3) \\
 &= (n_{\max} + 1) \left( 1 - \frac{n_{\max}}{2}\beta + \frac{n_{\max}(2n_{\max} + 1)}{12}\beta^2 \right). \tag{97}
 \end{aligned}$$

Taking the logarithm and using  $\ln(1+x) = x - x^2/2 + O(x^3)$  with  $x = -\frac{n_{\max}}{2}\beta + \frac{n_{\max}(2n_{\max}+1)}{12}\beta^2$ , we find

$$\begin{aligned}
 \ln(\mathcal{Z}) &= \ln(n_{\max} + 1) + \ln \left( 1 - \frac{n_{\max}}{2}\beta + \frac{n_{\max}(2n_{\max} + 1)}{12}\beta^2 \right) \\
 &= \ln(n_{\max} + 1) - \frac{n_{\max}}{2}\beta + \frac{n_{\max}(n_{\max} + 2)}{24}\beta^2 + O(\beta^3). \tag{98}
 \end{aligned}$$

The mean energy  $\langle E \rangle = -\partial \ln(\mathcal{Z})/\partial \beta$  is

$$\langle E \rangle = \frac{n_{\max}}{2} - \frac{n_{\max}(n_{\max} + 2)}{12}\beta + O(\beta^2), \tag{99}$$

and the variance  $\text{Var}(E) = \partial^2 \ln(\mathcal{Z})/\partial \beta^2$  is

$$\text{Var}(E) = \frac{n_{\max}(n_{\max} + 2)}{12} + O(\beta^2). \tag{100}$$

Using  $1/(1-x)^2 \approx 1 + 2x + 3x^2$ , we find

$$\begin{aligned}
 (\Delta U_C)^2 &= \frac{\frac{n_{\max}(n_{\max} + 2)}{12}}{\left( \frac{n_{\max}}{2} - \frac{n_{\max}(n_{\max} + 2)}{12}\beta \right)^2} \approx \frac{n_{\max} + 2}{3n_{\max}} \left( 1 + \frac{n_{\max} + 2}{3}\beta + O(\beta^2) \right) \\
 &= \frac{1}{3} + \frac{2}{3n_{\max}} + \frac{(n_{\max} + 2)^2}{9n_{\max}}\beta + O(\beta^2). \tag{101}
 \end{aligned}$$

for large  $n_{\max}$ ,  $(n_{\max} + 2)^2 \approx n_{\max}^2$ . Thus, in the high-temperature limit, the energy fluctuation behaves as

$$(\Delta U_C)^2 \approx \frac{1}{3} + \frac{2}{3n_{\max}} + \frac{n_{\max}}{9}\beta, \tag{102}$$

which is a linear function of  $\beta$ . Note that the high-temperature limit and thermodynamic limit do not commute with each other, as

$$\lim_{n_{\max} \rightarrow \infty} \lim_{\beta \rightarrow 0} (\Delta U_C)^2 = \frac{1}{3}, \tag{103}$$

while

$$\lim_{\beta \rightarrow 0} \lim_{n_{\max} \rightarrow \infty} (\Delta U_C)^2 = \lim_{\beta \rightarrow 0} e^{\beta \omega} = 1. \quad (104)$$

Physically, at high temperature a finite bath cannot supply many excited levels; the system hits the ceiling at  $n_{\max}$ , and the fluctuations saturate differently than in the true infinite-bath limit.

However, in the intermediate range of  $\beta$ , a local maximum in energy fluctuation distinguishes the noncanonical distribution from the canonical one, especially for strong system-bath interaction  $\kappa$ . This maximum can be qualitatively understood as follows: we can write the noncanonical distribution as

$$\tilde{P}_n(E) = \mathcal{Z}^{-1} \exp(-\beta \eta_n n \omega), \quad (105)$$

where

$$\eta_n = 1 - \kappa n / \omega - \kappa n^2 / E, \quad \xi = -\frac{\beta}{E}, \quad (106)$$

is a positive factor for  $n \leq n_{\max}$ . As  $\eta_n \leq 1$ , the factor  $\eta_n$  effectively reduces the inverse temperature  $\beta$  appearing in the Boltzmann weight, so that the system behaves as if it were at a higher effective temperature,

$$\beta_{\text{eff}}(n) = \beta \eta_n \leq \beta. \quad (107)$$

thus the linear region for  $\beta$  is enlarged in the noncanonical statistics. In other words, the coupled system behaves as if it were effectively hotter than a canonical oscillator at the same  $\beta$ . As a result, the range of  $\beta$  over which thermodynamic quantities vary approximately linearly with  $\beta$  (the high-temperature linear regime) becomes larger in the noncanonical case. On the fluctuation curves in Fig. 5, this shows up as a broader interval of  $\beta$  over which the noncanonical curve remains almost linear, whereas the canonical curve already bends away. These effects are most pronounced when the interaction strength  $\kappa$  is large and the temperature is high, because then many highly excited levels with  $\eta_n \ll 1$  are populated and the deviations from canonical statistics become significant.

Besides the high distribution tail for a single system, the noncanonical statistics provides other new characters when the system is a composite of two subsystems  $l_1$  and  $l_2$ . Even if these two subsystems do not directly interact with each other, the deformation of the energy shell can effectively result in a correlation between them. Here we still use harmonic oscillator (HO) systems for illustration. The system consists of two single-mode HOs with Hamiltonian

$$H_S = \sum_{k=1,2} \omega_k a_k^\dagger a_k. \quad (108)$$

The system interacts with a common small heat bath, which can be modeled by the Hamil-

tonian

$$H_B = \sum_{j=1}^N \omega_j b_j^\dagger b_j \quad (109)$$

The interaction term reads

$$H_I = \sum_{k=1,2} \sum_{j=1}^N \lambda_{kj} a_k^\dagger a_k (b_j^\dagger + b_j). \quad (110)$$

Following the same discussion about the energy shell deformation for a single system, we can obtain the joint distribution of the composite system. Given that the system is in level  $(n, m)$ , its effective energy is

$$E_{\text{sys}}(n, m) = n\omega_1 + m\omega_2 + \Delta_{n,m} = E_{nm} + \Delta_{n,m}, \quad (111)$$

where  $E_{nm} = n\omega_1 + m\omega_2$ . Then the energy available to the bath is

$$E_B = E - E_{\text{sys}}(n, m) = E - E_{nm} - \Delta_{n,m}. \quad (112)$$

We now restrict the bath+interaction to the system state  $|n, m\rangle$ . Using

$$a_1^\dagger a_1 |n, m\rangle = n |n, m\rangle, \quad a_2^\dagger a_2 |n, m\rangle = m |n, m\rangle,$$

we obtain

$$\begin{aligned} H_B + H_I \Big|_{(n,m)} &= \sum_j \left[ \omega_j b_j^\dagger b_j + (\lambda_{1j}n + \lambda_{2j}m)(b_j^\dagger + b_j) \right] \\ &= \sum_j \left[ \omega_j b_j^\dagger b_j + Q_j(n, m)(b_j^\dagger + b_j) \right], \end{aligned} \quad (113)$$

where  $Q_j(n, m) \equiv \lambda_{1j}n + \lambda_{2j}m$ . For a fixed pair  $(n, m)$  we introduce shifted operators

$$\tilde{b}_j = b_j + \alpha_j, \quad \tilde{b}_j^\dagger = b_j^\dagger + \alpha_j, \quad (114)$$

so that  $b_j = \tilde{b}_j - \alpha_j$ ,  $b_j^\dagger = \tilde{b}_j^\dagger - \alpha_j$ . Substituting into the Hamiltonian gives

$$\omega_j b_j^\dagger b_j + Q_j(b_j^\dagger + b_j) = \omega_j \tilde{b}_j^\dagger \tilde{b}_j - \omega_j \alpha_j (\tilde{b}_j^\dagger + \tilde{b}_j) + \omega_j \alpha_j^2 + Q_j(\tilde{b}_j^\dagger + \tilde{b}_j - 2\alpha_j) \quad (115)$$

As before, we choose  $\alpha_j$  so that the linear terms vanish:

$$-\omega_j \alpha_j + Q_j = 0 \implies \alpha_j = \frac{Q_j}{\omega_j}. \quad (116)$$

With this choice, the total Hamiltonian in the  $(n, m)$  sector becomes

$$H|_{(n,m)} = (\omega_1 n + \omega_2 m) + \sum_j \omega_j \tilde{b}_j^\dagger \tilde{b}_j - \sum_j \frac{Q_j(n, m)^2}{\omega_j}. \quad (117)$$

The last term is a system-dependent constant shift, which we denote by

$$\Delta_{nm} = - \sum_j \frac{Q_j(n, m)^2}{\omega_j} = - \sum_{j=1}^N \frac{(\lambda_{1j} n + \lambda_{2j} m)^2}{\omega_j}. \quad (118)$$

Using Eqs. (45) and (68), we get

$$\phi_{nm} = -\beta \Delta_{nm} + \xi \Delta_{nm} E_{nm}, \quad (119)$$

$$\Rightarrow \tilde{P}(E, E_n^{(1)}, E_m^{(2)}) = \frac{1}{\mathcal{Z}_t} e^{-\beta E_{nm} - \beta \Delta_{nm} + \xi \Delta_{nm} E_{nm}} = \frac{1}{\mathcal{Z}_t} e^{-\beta(E_{nm} + \Delta_{nm}) + \xi \Delta_{nm} E_{nm}}, \quad (120)$$

where

$$\Delta_{nm} = - \sum_{j=1}^N \frac{(\lambda_{1j} n + \lambda_{2j} m)^2}{\omega_j}, \quad (121)$$

$$\mathcal{Z}_t = \sum'_{n,m} e^{-\beta(E_{nm} + \Delta_{nm}) + \xi \Delta_{nm} E_{nm}}. \quad (122)$$

Here  $\sum'_{n,m}$  means the summation of the system energy levels should satisfy the constraint  $0 \leq E_{nm} + \Delta_{nm} \leq E$ . It can be seen from (120) that the statistics of two subsystems are not independent with each other due to the cross-term in  $\Delta_{nm}$ . This statistical correlation can be described by mutual information defined as

$$S(l_1 : l_2) = S(l_1) + S(l_2) - S(l_1 + l_2), \quad (123)$$

where the marginal entropies are

$$S(l_k) = - \sum_{n=0}^{n_{\max}} \tilde{P}(E, E_n^{(k)}) \ln \left( \tilde{P}(E, E_n^{(k)}) \right), \quad (124)$$

and the joint entropy is

$$S(l_1 + l_2) = - \sum'_{n,m} \tilde{P}(E, E_n^{(1)}, E_m^{(2)}) \ln \left( \tilde{P}(E, E_n^{(1)}, E_m^{(2)}) \right). \quad (125)$$

Not that

$$\tilde{P}_n^{(1)} = \sum_m \tilde{P}_{nm}(E) = \frac{1}{\mathcal{Z}_t} \sum_m e^{-\beta(E_{nm} + \Delta_{nm}) + \xi \Delta_{nm} E_{nm}}, \quad (126)$$

$$\tilde{P}_n^{(2)} = \sum_n \tilde{P}_{nm}(E) = \frac{1}{\mathcal{Z}_t} \sum_n e^{-\beta(E_{nm} + \Delta_{nm}) + \xi \Delta_{nm} E_{nm}}. \quad (127)$$

For simplicity, we assume the two HOs are identical with  $\omega_1 = \omega_2 \equiv \omega$ ,  $\lambda_{1j} = \lambda_{2j} \equiv \lambda_j$ , and  $\kappa$  is defined the same as Eq. (61). As shown in Fig. 6, there appears to be a nonzero mutual entropy if we use noncanonical statistics to describe the composite system in a common small heat bath. In contrast, when the interaction energy is so small that it can be neglected

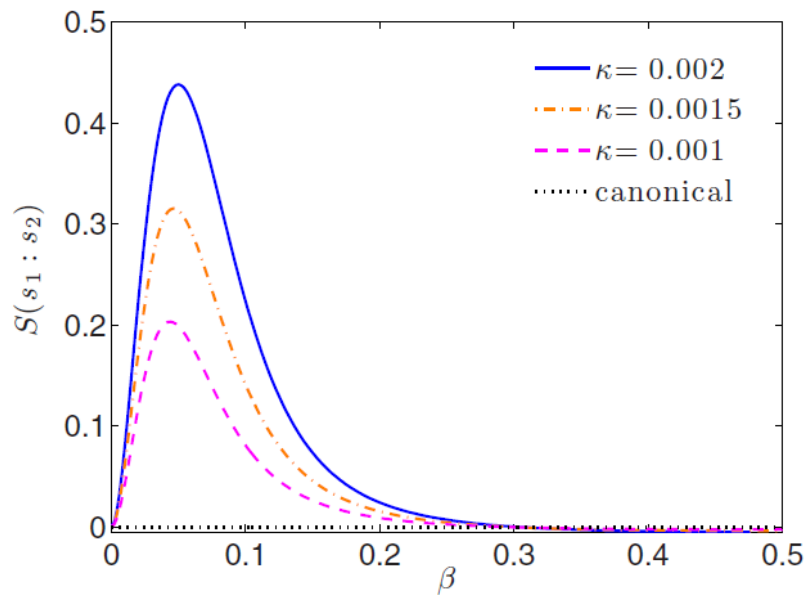


Figure 6: The mutual entropy of two identical harmonic oscillators (HOs) by noncanonical and canonical statistics. The eigenfrequency of two identical HOs is  $\omega = 1$ , and  $\kappa = 0.002$ ,  $n_{\max} = 138$  (the blue solid line);  $\kappa = 0.0015$ ,  $n_{\max} = 122$  (orange dotted-dash line);  $\kappa = 0.001$ ,  $n_{\max} = 112$  (pink dashed line); and the canonical case for  $\kappa = 0$ ,  $n_{\max} = 112$ .

compared with the system and bath energies, the total state is well described by the canonical distribution without coupling. In this limit the joint density matrix can be written as a simple product of two independent Gibbs states, and the mutual entropy (mutual information) becomes

$$S(l_1 : l_2) = 0,$$

which means that the two subsystems are completely uncorrelated.

## 1.4 Conclusion

In this report we have examined how the usual canonical description of a small quantum system in contact with a heat bath is modified when the bath is finite and the system-bath coupling cannot be neglected. Starting from the microcanonical ensemble of the full "universe" (system+bath), we reviewed how the standard Gibbs state emerges in the thermodynamic limit, and then followed Xu and Sun's construction of an effective Hamiltonian that encodes finite-bath and interaction effects through a deformation of the energy shell. This leads naturally to a noncanonical reduced state for the system, in which the probabilities are no longer simple Boltzmann weights.

For the single harmonic oscillator model, the key ingredient is the level-dependent energy shift  $\Delta_n = -\kappa n^2$  generated by a diagonal system-bath coupling. Treating the bath entropy  $S(E)$  as a smooth function of the total energy and expanding  $S(E - E_n)$  and its interaction-induced correction  $\delta_n S(E - E_n)$ , one arrives at the noncanonical equilibrium distribution

$$\tilde{P}_n(E) \propto \exp[-\beta(n\omega - \kappa n^2) - \xi \omega \kappa n^3],$$

with  $\beta = \partial_E S$  and  $\xi = \partial_E^2 S$ . The quadratic and cubic terms in  $n$  lead to a pronounced high-energy tail compared to the canonical distribution  $P_n \propto e^{-\beta n\omega}$ , and hence to an enhanced internal energy and larger fluctuations when the coupling  $\kappa$  is significant. At the same time, the bath energy must remain non-negative, and this imposes a finite upper cutoff  $n_{\max}$  on the allowed system levels. This makes all thermodynamic quantities bounded in the high-temperature limit, and explains why the limits  $\beta \rightarrow 0$  and  $n_{\max} \rightarrow \infty$  do not commute.

The analysis of energy fluctuations shows how finite-bath effects behave in the low- and high-temperature limits. For very low temperatures ( $\beta\omega \gg 1$ ), the oscillator effectively becomes a two-level system. In this regime the relative fluctuation  $(\Delta U_C)^2 \simeq e^{\beta\omega}$  is the same whether the spectrum is truncated or infinite, so a finite bath and an infinite bath are indistinguishable. For high temperatures and a finite bath, however, the internal energy  $U$  and  $(\Delta U)^2$  approach finite values set by  $n_{\max}$ , and the fluctuation curve develops a characteristic shape that cannot be reproduced by the canonical distribution. In an intermediate range of  $\beta$ , the noncanonical distribution can be written in terms of an effective inverse temperature  $\beta_{\text{eff}}(n) = \beta\eta_n$  with  $\eta_n < 1$  for highly excited levels. This makes the system behave as if it were at a higher temperature than the canonical prediction, which explains the enhanced high-energy tail and the local maximum seen in the fluctuation plots.

The noncanonical framework becomes even more interesting when the system is a composite of two subsystems. For two harmonic oscillators coupled to a common finite bath, the deformed energy shell produces a joint distribution  $\tilde{P}_{nm}(E)$  that depends on the combination  $(\lambda_{1j}n + \lambda_{2j}m)^2$ . Even though the oscillators do not interact directly, their statistics are no longer independent: the joint distribution does not factorize into a product of two marginals, and the mutual information  $S(l_1 : l_2)$  is strictly positive. This mutual information measures the correlation induced purely by the shared finite bath. In the weak-coupling, large-bath limit the deformation disappears, the joint state factorizes, and the mutual information tends

---

to zero, recovering the usual picture of two independent subsystems in contact with an ideal reservoir.

## 2 Finite-bath corrections to the second law of thermodynamics

### 2.1 Introduction

As we discussed, in the microscopic regime some of the standard assumptions of macroscopic thermodynamics no longer hold. Fluctuations, especially in quantities like work, can be comparable to, or even exceed, the characteristic energy scales of the system, so treating them as negligible becomes unjustified. A second assumption that can fail is that the environment acts as an ideal, infinitely large heat reservoir. At microscopic scales the bath may be finite, so energy exchange can noticeably disturb it, which is exactly the situation we analyzed in the previous section. In particular, when the time scales involved in a process are small, the effective heat bath is necessarily finite, as implied by Lieb-Robinson bounds. We will go through the definition and the implications of this bound in the following sections.

The (macroscopic) second law of thermodynamics can be stated as follows. Consider a system in an arbitrary state, with average energy  $U$  and entropy  $S$ . Suppose that, in order to transform the system, we can make use of a heat bath at temperature  $T$ , and extract an average amount of work  $\langle W \rangle$ . A necessary condition for the possibility of a transformation is

$$\langle W \rangle \leq -\Delta U + T\Delta S = -\Delta F, \quad (128)$$

where  $\Delta U$  and  $\Delta S$  are the changes in energy and entropy, and the free energy is defined by  $F = U - TS$ . Eq. (128) follows by combining the first law,  $\Delta U = Q - W$ , with the standard second-law (Clausius) bound for an ideal infinite bath at fixed temperature  $T$ . A detailed derivation will be given in later sections. Eq. (128) implies that attempting to extract more work than  $-\Delta F$  is impossible: no protocol that respects energy conservation and uses a thermal environment can implement such a transformation.

Remarkably, when the bath has infinite heat capacity, the same inequality also becomes sufficient: satisfying  $\langle W \rangle \leq -\Delta F$  guarantees that there exists an allowed protocol that achieves the desired transformation with that (or arbitrarily close to that) work value. Therefore, there are no additional hidden constraints coming from disturbing the bath. In this case the temperature stays relatively constant during the process, so one can implement a quasi-static isothermal protocol: change the system in many infinitesimal steps and reequilibrate with the bath at the same  $\beta$  each step. As the step size goes to zero, entropy production can be made arbitrarily small, and the extracted work approaches the reversible value  $\langle W \rangle \rightarrow -\Delta F$ . A bound is called tight if it can be saturated (achieved as an equality) for some optimal protocol. Therefore, the bound in Eq. (128) is tight.

However, in the microscopic scale, the assumption that the bath is infinite might not always be a reasonable approximation. Recall that, within a finite time, a system can only interact with a finite region of its environment. Hence for finite-time processes one should consider

an effective bath with a finite heat capacity.

A reasonable question then would be how does this fundamental bound (128) change when we take into account the finiteness of the heat bath? Consider a bath with heat capacity  $C$ , which is not necessarily infinite. We find that the necessary and sufficient condition for the possibility of a transformation is

$$\langle W \rangle \leq -\Delta F - \frac{1}{\beta} [I(E' : S') + D(P(E') \| p_G(E'))], \quad (129)$$

where  $D$  is the relative entropy defined as

$$D(P(E') \| p_G(E')) = \text{Tr}(P(E') (\ln(P(E')) - \ln(p_G(E')))), \quad (130)$$

$P(E')$  is the probability distribution for the final energy of the bath  $E'$ ,  $p_G(E')$  is the Gibbs distribution at temperature  $\beta$ , and  $I(E' : S')$  is the mutual information of the final state of the system and bath. This bound thus has two different contributions: one from the correlations created between system and bath, and another for the change of state of the bath itself.

From (129) we derive a necessary condition that depends only on the initial and final states of the system and the heat capacity  $C$  and temperature  $T = 1/\beta$  of the bath, which reads

$$\langle W \rangle \leq -\Delta F - \frac{1}{2C\beta} \Delta S^2. \quad (131)$$

Keep in mind that we are just going through what will be covered in this report, and all the detailed derivations will be in the following sections. Note that this expression converges to the macroscopic second law (128) in the limit where the heat capacity of the bath  $C$  is large and provides a stronger condition when  $C$  is finite. Also condition (131) becomes sufficient when the final state is maximally mixed. In this case, expression (131) is not just an upper bound, but the maximal work extractable. Note that this is smaller than the minimal work invested in the preparation of the initial state. Therefore, we conclude that thermodynamic reversibility requires an infinite heat bath, as we concluded in the last paper also.

There are situations where stochastic fluctuations of work may be undesirable and one would like to manipulate only definite amounts of useful energy. To address this, the framework of single-shot thermodynamics was established in [10,11], where the concept of deterministic work was introduced. The authors showed how work can be understood as a shift of energy in the storage system that happens with a very large probability. They also gave expressions for how much one can extract from any given state and for how much one needs to create it, assuming access to an infinite heat bath.

In this paper they explore how that amount of deterministic work content and work cost of forming a state changes when one is limited to a finite bath and we find that achieving these processes with arbitrary accuracy becomes impossible. This may not be the case, however, if one allows for additional small error probability during the processes, given by the tails of the distribution of energy of the heat bath. For such a case, we give expressions for



the deterministic work for a given probability of failure.

Finite bath thermodynamics has gathered some interest in recent years, including finite bath corrections to Carnot efficiency, Landauer's principle, and the Jarzinski equality, as seen in, e.g.m Refs [12-16]. In fact, the bound of Eq. (129) has appeared before in the literature, as it coincides with the main bounds of both [14] and [17], derived in similar contexts, and in [18] a bound that includes only the term  $D(P(E')\|p_G(E'))$  was shown. In contrast to these works, the authors in this paper prove that this bound is not only a necessary constraint, but also a sufficient one.

In a number of resource-theoretic references, the infinite bath is taken to an infinite, noninteracting number of copies of a small thermal state [19-21]. The finite-sized versions of these results reflecting a finite number of copies of thermal states have been explored in [22,23], with finding similar to the ones presented in this paper. As a main difference, all of these contributions make assumptions on the particular structure of the heat bath (for example, as many copies of small thermal states). Our contribution has the advantage of not making any assumption on the structure of the heat bath beyond its heat capacity.

The structure of this paper is as follows. In Sec. ??, we provide a model-independent characterization of a finite bath. In Sec. ?? we describe the general model for thermodynamic system-bath interactions and consider work extraction protocols with fluctuating work. We use this to show our main result and explore the much studied case of Landauer erasure. In Sec. ?? we further explore the possible interactions between system and bath only and how the system can be transformed via operations that do not involve work. We use this to derive fundamental limits to deterministic work extraction and expenditure.

## 2.2 General characterization of a finite bath

We consider a bath to be a large (but in this case not infinitely so) system with a density of states given by  $\Omega(E, V) = e^{S(E, V)}$ , where  $S(E, V)$  is the entropy in the microcanonical ensemble for a given energy  $E$  and volume  $V$ . We shall make three assumptions about it.

1. The entropy  $S(E, V)$  is extensive:

$$S(kE, kV) = kS(E, V), \quad \text{for all } k > 0.$$

2. The dimensionless volume  $V$  is large.

3. The bath is in a Gibbs state with a given temperature  $\beta$ , such that a microstate of energy  $E$  has probability  $\frac{1}{Z}e^{-\beta E}$ .

We will be working in units for which constant is  $k_B = 1$ . Assumption 1 implies implies that we can write the entropy  $S(E, V)$  as

$$S(E, V) = Vf(u), \tag{132}$$

for some function  $f(u)$  of the energy density  $u = E/V$ . The probability distribution for  $u$  is then

$$p(u) \propto e^{V[f(u) - \beta u]}. \quad (133)$$

In the large  $V$  limit we can use the saddle point approximation [24]

$$p(u) \propto e^{V(f(u_\beta) + \frac{1}{2}f''(u_\beta)(u - u_\beta)^2)}, \quad (134)$$

where  $u_\beta$  is the absolute maximum of  $f(u) - \beta u$  as a function of  $u$ , which implies

$$f'(u_\beta) = \beta, \quad (135)$$

$$f''(u_\beta) < 0. \quad (136)$$

In summary, we have a normal distribution

$$p(u) \propto e^{-\frac{V}{2}|f''(u_\beta)|(u - u_\beta)^2}, \quad (137)$$

with mean  $\langle u \rangle = u_\beta$  and variance

$$\langle (u - \langle u \rangle)^2 \rangle = |V f''(u_\beta)|^{-1}. \quad (138)$$

Now, let us relate  $f''(u_\beta)$  to the heat capacity, defined as

$$C = V \frac{d \langle u \rangle}{dT} = -\frac{V}{T^2} \frac{d \langle u \rangle}{d\beta}. \quad (139)$$

Differentiating (135) with respect to  $\beta$  gives

$$f''(u_\beta) \frac{du_\beta}{d\beta} = 1,$$

and substituting in (139) gives

$$C = -\frac{V}{T^2} \frac{1}{f''(u_\beta)}. \quad (140)$$

Note that Eq. (136) implies that the heat capacity is positive, as is always the case in ordinary matter. Also note that  $C \propto V$ , because  $f(u_\beta)$  is independent of  $V$ . Using this and Eq. (138) implies that the fluctuations of  $u$  are

$$\langle (u - \langle u \rangle)^2 \rangle^{1/2} = \frac{T\sqrt{C}}{V} \propto V^{-1/2}, \quad (141)$$

which are small when  $V$  is large.

We are finally able to approximate the density of states of the bath as

$$\omega(E, V) \propto \exp\left(\beta E - \frac{\gamma E^2}{2}\right), \quad (142)$$

where the energy has been rescaled such that  $\langle E \rangle = 0$ , and we define  $\gamma = \frac{1}{CT^2}$ .

## 2.3 Fluctuating work

### 2.3.1 Thermal operations with fluctuating work

Next the authors introduce a widely used framework to describe thermodynamic transformations. Their setting consists of a system with Hamiltonian  $H_S$ , the bath with Hamiltonian  $H_B$  initially in the thermal state (as described in a previous section), and an ideal weight with Hamiltonian  $H_W = \int_{\mathbb{R}} dx x |x\rangle \langle x|$ , where the orthonormal basis  $\{|x\rangle, \forall x \in \mathbb{R}\}$  represents the position of the weight. Any joint transformation of system, bath, and weight is represented by a completely positive trace preserving (CPTP) map  $\Gamma_{SBW}$  satisfying the following conditions.

- *Microscopic reversibility (second law).* It has an (CPTP) inverse  $\Gamma_{SBW}^{-1}$ , which implies unitarity  $\Gamma_{SBW}(\rho_{SBW}) = U \rho_{SBW} U^\dagger$ .
- *Energy conservation (first law).*  $[U, H_S + H_B + H_W] = 0$ .
- *Independence from the "position" of the weight.* The unitary commutes with the translations on the weight  $[U, \Delta_W] = 0$ . The generator of the translations  $\Delta_W$  is canonically conjugated to the position (or energy) of the weight  $[H_W, \Delta_W] = i$ .
- *Classicality of work.* Before and after applying the global map  $\Gamma_{SBW}$  the position of the weight is measured, obtaining outcomes  $|x\rangle$  and  $|x + W\rangle$ , respectively. In general, the work  $W$  is a fluctuating random variable.

Condition  $[U, \Delta_W] = 0$  implies that the reduced map on system and bath is a mixture of unitaries (Result 1 in [25]). Hence, these transformations can never decrease the entropy if system and bath, which guarantees that the weight is not used as a source of free energy.

Let us define the dephasing map as

$$\Theta_\alpha[\rho_S] = \int_{\mathbb{R}} dt e^{i\alpha t} e^{iH_S t} \rho e^{-iH_S t}. \quad (143)$$

Energy conservation, the classicality of work, and the fact that the initial state of the bath commutes with its Hamiltonian imply

$$\Theta_\alpha \circ \Gamma_S = \Gamma_S \circ \Theta_\alpha, \quad (144)$$

where

$$\Gamma_S(\rho_S) = \text{Tr}_{BW}(\Gamma_{SBW}(\rho_S \otimes \rho_B \otimes \rho_W)), \quad (145)$$

is the transformation of the system. See [8] for a proof. Note the assumption that the initial state of system, bath, and weight is uncorrelated. Setting  $\alpha = 0$  we have that if the initial state of the system, commutes with  $H_S$ , then so does the final state. In this paper the authors only consider processes in which both states are diagonal. For example, optimal work extraction from an arbitrary initial state is one such process. For processes where the initial and final states involve coherences, our results provide an upper bound to the work. See Appendix A 1 for further details.

We write the initial and final states as  $\rho_S = \sum_S P(s) |s\rangle \langle s|$  and  $\rho'_S = \sum_{s'} P(s') |s'\rangle \langle s'|$ , respectively, where  $|s\rangle$  and  $|s'\rangle$  are the initial and final energy eigenstates. Note that we allow initial and final Hamiltonians  $H_S$  and  $H'_S$  to be not necessarily equal.

### 2.3.2 Corrections to the second law

In this section we analyze the transformation power of the operations defined in the previous section and study the effects of not having an infinite heat bath. First, we present a generalization to the second law (128) to the case of arbitrary heat bath. This necessary and sufficient condition has an important limitation in that it involves the final state of the bath. Later, more practical bounds that are independent from the state of the bath are provided.

**Theorem 1.** A necessary and sufficient condition for the possibility of a transformation after a thermal operation is

$$\langle W \rangle \leq -\Delta F - \frac{1}{\beta} [I(E' : S') + D(P(E') \| p_G(E'))], \quad (146)$$

where  $D(p \| q) = \sum_{x,y} p(x) \log(p(x)/q(x))$  is the relative entropy,  $p_G(E')$  is the probability distribution of the bath with energy  $E'$ ,  $P(E')$  is the final probability distribution of the bath, and  $I(E' : S')$  is the mutual information of system and bath after the operation.