New Journal of Physics

The open access journal at the forefront of physics

PAPER

CrossMark

From single-shot towards general work extraction in a quantum thermodynamic framework

Jochen Gemmer[1] and Janet Anders[2,3]

[1] Fachbereich Physik, Universität Osnabrück, Barbarastrasse 7, D-49069 Osnabrück, Germany
[2] Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, UK
[3] Author to whom any correspondence should be addressed.

E-mail: jgemmer@uos.de and janet@qipc.org

Abstract

This paper considers work extraction from a quantum system to a work storage system (or weight) following Horodecki and Oppenheim (2013 *Nat. Commun.* **4** 2059). An alternative approach is here developed that relies on the comparison of subspace dimensions without a need to introduce thermo-majorization used previously. Optimal single shot work for processes where a weight transfers from (a) a single energy level to another single energy level is then re-derived. In addition we discuss the final state of the system after work extraction and show that the system typically ends in its thermal state, while there are cases where the system is only close to it. The work of formation in the single level transfer setting is also re-derived. The approach presented now allows the extension of the single shot work concept to work extraction (b) involving multiple final levels of the weight. A key conclusion here is that the single shot work for case (a) is appropriate only when a *resonance* of a particular energy is required. When wishing to identify 'work extraction' with finding the weight in a specific available energy or any higher energy a broadening of the single shot work concept is required. As a final contribution we consider transformations of the system that (c) result in general weight state transfers. Introducing a transfer-quantity allows us to formulate minimum requirements for transformations to be at all possible in a thermodynamic framework. We show that choosing the free energy difference of the weight as the transfer-quantity one recovers various single shot results including single level transitions (a), multiple final level transitions (b), and recent results on restricted sets of multi-level to multi-level weight transfers.

1. Introduction

The neat characterization of general classical non-equilibrium processes in terms of fluctuation relations [2–4, 8] has rapidly advanced the general understanding of thermodynamic processes and properties at the mesoscopic scale. Work, in particular, is a thermodynamic quantity of interest and the stochastic fluctuations of work done on a system are captured in the Jarzynski relation [3]. The fluctuation relation approach has been extended to quantum systems where probabilistic energy transfers of the system that undergoes unitary evolution are associated with the fluctuating work done on the system [5–7, 9–12]. Again the work for this quantum scenario can be captured in a quantum Jarzynski relation. On the other hand, thermodynamic processes for a quantum system can be studied in a setting where the system interacts with a heat bath and a work storage system (or weight) undergoing global unitary dynamics [1, 13–18]. Operation on the system then results in a change of the *work storage system's state* and it is that change that is here associated with 'work'. These approaches are referred to as 'thermodynamic resource theory' and 'single shot thermodynamics'. Recent papers, e.g. [1, 15], derive upper bounds on the amount of work that can be drawn from a quantum system that starts in a non-equilibrium state in a 'single shot'.

The *single shot work* done by the system is here associated [1] with the transition of the weight from a single energy eigenstate (of energy 0) to another single energy eigenstate (of energy *w*). This situation is sketched in
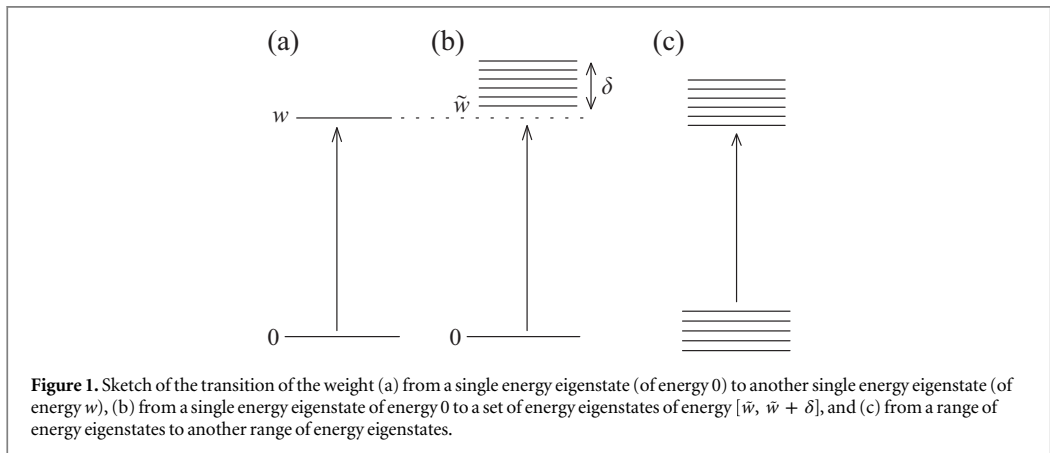
Figure 1. Sketch of the transition of the weight (a) from a single energy eigenstate (of energy 0) to another single energy eigenstate (of energy w), (b) from a single energy eigenstate of energy 0 to a set of energy eigenstates of energy $[\bar{w}, \bar{w} + \delta]$, and (c) from a range of energy eigenstates to another range of energy eigenstates.

figure 1(a). The proof of the bounds provided in [1] relies on established mathematical concepts from quantum information theory and a new majorization concept called 'thermo-majorization'. Researchers, in particular those outside of quantum information theory, may find the proof mathematically heavy and are unable to follow the detailed logic. This technical difficulty overshadows the interpretation of the results and may hinder their further development and formulation of experimental tests of single shot results.

Here we aim to re-derive single shot work extraction limits while keeping the technical side as simple as possible. The hope is that stripping the discussion from some of the jargon will allow to focus on the physical meaning of the results, clarify the situation they describe, and develop the argument further to adapt to different physically relevant scenarios.

The paper is organized as follows. The known work extraction result of [1] for scenario (a) in figure 1 is re-stated in section 2 and then re-derived in section 3. The final state of the system after this work extraction procedure is explored in section 4. Also re-derived, in section 5, is the work of formation which was first identified as different from the extractable work in [1]. The extension of single shot work extraction to general transfer processes is discussed thereafter. Section 6 is concerned with the extractable work allowing (b) the transition of the weight from a single energy eigenstate of energy 0 to a set of energy eigenstates of energy $[\bar{w}, \bar{w} + \delta]$. Section 7 identifies a transfer-quantity that characterizes (c) the transition of the weight from an arbitrary range of energy eigenstates to another range of energy eigenstates. Finally, the findings and open questions are discussed in section 8.

2. Known results on maximal single shot work extraction

A key result in single-shot thermodynamics is the identification of a maximal work [1, 15]

$$w_\epsilon^{\max} = F_\epsilon^{\min} (\rho_S) - F(\tau_S),$$ (1)

that can be extracted with success probability $1 - \epsilon$ from a system starting in a state ρ_S under so-called *thermal operations* [1]. Here $F(\tau_S) := -\frac{1}{\beta} \ln Z_S$ is the standard free energy associated with the thermal state $\tau_S := \frac{e^{-\beta H_S}}{Z_S}$ for a system Hamiltonian H_S at inverse temperature β, with $Z_S = \mathrm{tr}[e^{-\beta H_S}]$ the partition function. (In [1] the system is a qubit with the excited state energy tuned such as to have exactly the optimal work value that can be gained, i.e. $H_S = 0 \, |0\rangle_S\langle 0| + w_\epsilon^{\max} \, |1\rangle_S\langle 1|$.) The other quantity, $F_\epsilon^{\min} (\rho_S)$, is a generalized free energy applicable for the non-equilibrium state ρ_S [1] which will be detailed below. This maximal work is valid for initial states ρ_S that are diagonal in the energy basis and it is at least a lower bound on the maximal work for non-diagonal states [1, 15, 19].

This result is derived within the thermodynamic resource theory setting [1, 13–18] involving three components: the system of interest, S, a bath B, and a weight (or work storage system) W. In the simplest case, the Hamiltonian at the start and the end of the process is assumed to be the same and the sum of the three local terms, $H = H_S + H_B + H_W$. The weight is assumed to have no degeneracies, i.e. all its energy eigenstates have different energies E_W. The system's degeneracy is not restricted and we denote the multiplicity of each energy E_S by $M_S(E_S)$ and label each of them by $g_S(E_S) = 1, ..., M_S(E_S)$. The degenerate bath levels at energy E_B are labeled by $f_B(E_B)$ and their multiplicity is assumed to be exponentially growing with inverse temperature $f_B(E_B) = 1, ..., M_B(E_B) = 1, ..., M_B(0)\exp(\beta E_B)$. This exponential growth of the degree of degeneracy (or the

density of states) generically results for all systems made of many similar subsystems that feature short range interactions [20]. We note that it only holds for a possibly large but finite energy regime.

Thermal operations have been defined [1] as those transformations of the system that can be generated by a global unitary, V, that acts on system, bath and work storage system initially in a product state $\rho_S \otimes \tau_B \otimes |E_W^{\text{ini}}\rangle_W \langle E_W^{\text{ini}}|$, with the bath in a thermal state, $\tau_B = \frac{e^{-\beta H_B}}{Z_B}$, and the work storage system in one of its energy eigenstates, $|E_W^{\text{ini}}\rangle$. Conceptually, global unitaries V describe the operation of a 'work extraction machine' that aims to extract as much energy as possible to the work storage system. Perfect energy conservation is imposed by requiring that the unitary may only induce transitions within an energy shell of total energy $E = E_W + E_S + E_B$. This is equivalent to requiring that V must commute with the sum of the three local Hamiltonians.

The generalized free energy of the non-equilibrium state ρ_S is defined as

$$F_\epsilon^{\min}(\rho_S) := -\frac{1}{\beta} \ln \sum_{E_S, g_S} e^{-\beta E_S} h\left(E_S, g_S, \epsilon\right). \tag{2}$$

Here $h(E_S, g_S, \epsilon)$ is a binary function that determines whether a particular energy eigenstate, $|E_S, g_S\rangle$, is included in the summation or not. The value of h depends on ρ_S and on the failure rate ϵ that is being accepted for the work extraction. The exact dependence will be discussed further in section 3. The optimal work stated in equation (1) now corresponds to the following task: for the weight initially entirely in its ground state (with energy $E_W^{\text{ini}} = 0$) the full system is transferred to a final quantum state such that the probability for the weight to be found in an energy eigenstate with energy E_W^{fin} is $1 - \epsilon$ where $1 > \epsilon \geqslant 0$, see figure 1(a). This process is associated with the 'lifting' of the weight and the energy difference experienced by the weight is identified with 'extracted work', $w := E_W^{\text{fin}} - E_W^{\text{ini}}$. For the special case of perfect work extraction, i.e. $\epsilon = 0$, the summation in equation (2) includes all energy eigenstates $|E_S, g_S\rangle$ that are populated in the initial state, i.e. $h(E_S, g_S, 0) = 1$ when $\langle E_S, g_S| \rho_S |E_S, g_S\rangle > 0$, and 0 otherwise. The maximal work is then [1]

$$w_0^{\max} = -\frac{1}{\beta} \ln \, \text{tr}\left[\tau_S \hat{\Pi}_{\rho_S}\right], \tag{3}$$

where $\hat{\Pi}_{\rho_S}$ is the projector on the support of ρ_S.

3. Work extraction bounds from limits on probability transfer

The derivation of the maximal extractable work presented in [1] rests on an analysis of state dimensions which is combined with the newly introduced concept of *thermo-majorization*. This is a variation on *majorization*, an important tool in the study of doubly stochastic matrices [21–23] and quantum channels in particular [24, 25]. Here we present a derivation of the same result without a need of invoking thermo-majorization explicitly. We hope this re-derivation is more straightforward to follow for researchers wishing to familiarise themselves with single-shot thermodynamics.

The rational of the presented approach is to compare dimensions of the involved projective subspaces to conclude what transformations on the system are possible in the setting given and what maximal work extraction they enable. The re-analysis opens avenues of generalizing the previous result, valid for transitions of the weight from single energy level to single energy level, to processes where the weight transfers between multiple energy levels, which we will explore in sections 4 and 6. The work required to form a non-equilibrium state is also described, in section 5.

The desired thermal transformation is of the form

$$\eta := \rho_{SB} \otimes |0\rangle_W \langle 0| \xrightarrow{V} \sigma_{SB} \otimes |w\rangle_W \langle w| =: \eta'(V), \tag{4}$$

where the system and bath start in a fixed state ρ_{SB} and end in some state σ_{SB}, while the work storage system lifts from a single energy level, $|0\rangle$, to another single energy level, $|w\rangle$, by an energy $w > 0$, see figure 1(a). Such a transformation, however, turns out to be impossible for a large class of initial states, ρ_{SB}, namely those which have full rank. Thus one allows transformations under which the lhs of (4) is transformed into the rhs only up to a success probability $1 - \epsilon$, as already introduced in section 2. Since the global unitary, V, is fully energy conserving we may treat the global dynamics for each global energy shell E separately and later combine their contributions. Note that those energy shells are projective subspaces and so are the individual energy levels that make up the shell. We denote the local projectors defined by their associated local energies E_S, E_B and E_W by $\hat{\Pi}_S(E_S)$, $\hat{\Pi}_B(E_B)$ and $\hat{\Pi}_W(E_W)$.

To find the maximal work $w_\epsilon^{\max}$ we first consider a probability $P_{\text{ini}}^E \leqslant 1$ which falls initially into a projective subspace $\hat{\Pi}_{\text{ini}}^E$ of dimension d_{ini}^E, i.e. $P_{\text{ini}}^E = \text{tr}[\eta \hat{\Pi}_{\text{ini}}^E]$ for the initial state η. The key task here is to decide if this

probability can be transferred *entirely* into a projective subspace $\hat{\Pi}_{\mathrm{fin}}^E$ under unitary transformations V. This is of course possible if

$$\mathrm{tr}\left[\hat{\Pi}_{\mathrm{fin}}^E V \hat{\Pi}_{\mathrm{ini}}^E \eta \hat{\Pi}_{\mathrm{ini}}^E V^\dagger\right] \overset{!}{=} \mathrm{tr}\left[\hat{\Pi}_{\mathrm{ini}}^E \eta \hat{\Pi}_{\mathrm{ini}}^E\right] = P_{\mathrm{ini}}^E \tag{5}$$

holds. Rewriting this condition using the non-trivial eigenstates $|n\rangle$ of $V\hat{\Pi}_{\mathrm{ini}}^E \eta \hat{\Pi}_{\mathrm{ini}}^E V^\dagger$, i.e.
$V\hat{\Pi}_{\mathrm{ini}}^E \eta \hat{\Pi}_{\mathrm{ini}}^E V^\dagger |n\rangle = p_n |n\rangle$ with $p_n > 0$ for $n = 1,...,d_{\mathrm{ini}}^E$, implies

$$\sum_{n=1}^{d_{\mathrm{ini}}^E} p_n \langle n| \hat{\Pi}_{\mathrm{fin}}^E |n\rangle \overset{!}{=} \sum_{n=1}^{d_{\mathrm{ini}}^E} p_n = P_{\mathrm{ini}}^E \tag{6}$$

and therefore one must have $\langle n| \hat{\Pi}_{\mathrm{fin}}^E |n\rangle = 1$ for $n = 1,...,d_{\mathrm{ini}}^E$. This is only possible if the dimension of the space into which the probability is mapped, $d_{\mathrm{fin}}^E := \mathrm{rk}[\hat{\Pi}_{\mathrm{fin}}^E]$, is at least the same as the initial dimension, i.e. the condition becomes

$$d_{\mathrm{fin}}^E \geqslant d_{\mathrm{ini}}^E. \tag{7}$$

If not the whole probability P_{ini}^E but a slightly reduced probability $(1 - \epsilon) P_{\mathrm{ini}}^E$ is to be transferred, then one may replace $\hat{\Pi}_{\mathrm{ini}}^E$ in equation (5) with any projector $\hat{\Pi}_{\mathrm{ini}}^E(\epsilon)$ that fulfils

$$\mathrm{tr}\left[\hat{\Pi}_{\mathrm{ini}}^E(\epsilon)\eta\right] \geqslant P_{\mathrm{ini}}^E(1 - \epsilon). \tag{8}$$

Defining the smallest initial subspace dimension as $d_{\mathrm{ini}}^E(\epsilon, \eta) := \min_{\hat{\Pi}_{\mathrm{ini}}^E(\epsilon)} \dim(\hat{\Pi}_{\mathrm{ini}}^E(\epsilon)) \leqslant d_{\mathrm{ini}}^E$, the final space dimension condition, when allowing a small error probability ϵ in the transfer, relaxes to

$$d_{\mathrm{fin}}^E \geqslant d_{\mathrm{ini}}^E(\epsilon, \eta). \tag{9}$$

This constraint is now the key condition to establish what the maximum extractable work is for the thermal operations specified in equation (4).

Since V cannot mix energy shells, it is necessary to check the condition for each energy shell E. We interpret the projector $\hat{\Pi}_{\mathrm{fin}}^E$ used above as the operator that projects on the non-trivially populated subspace of the final global state, $\sigma_{SB} \otimes |w\rangle_W \langle w|$, with energy E. By construction one has the relation

$$\hat{\Pi}_{\mathrm{fin}}^E = \sum_{E_S} \hat{\Pi}_S\left(E_S\right) \otimes \hat{\Pi}_B\left(E - E_S - w\right) \otimes \hat{\Pi}_W(w). \tag{10}$$

We can now identify the dimension of $\hat{\Pi}_{\mathrm{fin}}^E$,

$$d_{\mathrm{fin}}^E = \sum_{E_S} M_S\left(E_S\right) M_B\left(E - E_S - w\right) = \sum_{E_S} M_S\left(E_S\right), \; M_B(E) \, e^{-\beta E_S - \beta w} \tag{11}$$

$$= M_B(E) \, e^{-\beta w} Z_S, \tag{12}$$

where $Z_S = \sum_{E_S} M_S(E_S) \, e^{-\beta E_S}$ is the thermal equilibrium partition function at inverse temperature β for the system. The second step is to consider the initial subspace projectors in the energy shell E,

$$\hat{\Pi}_{\mathrm{ini}}^E = \sum_{E_S} \hat{\Pi}_S\left(E_S\right) \otimes \hat{\Pi}_B\left(E - E_S\right) \otimes \hat{\Pi}_W(0). \tag{13}$$
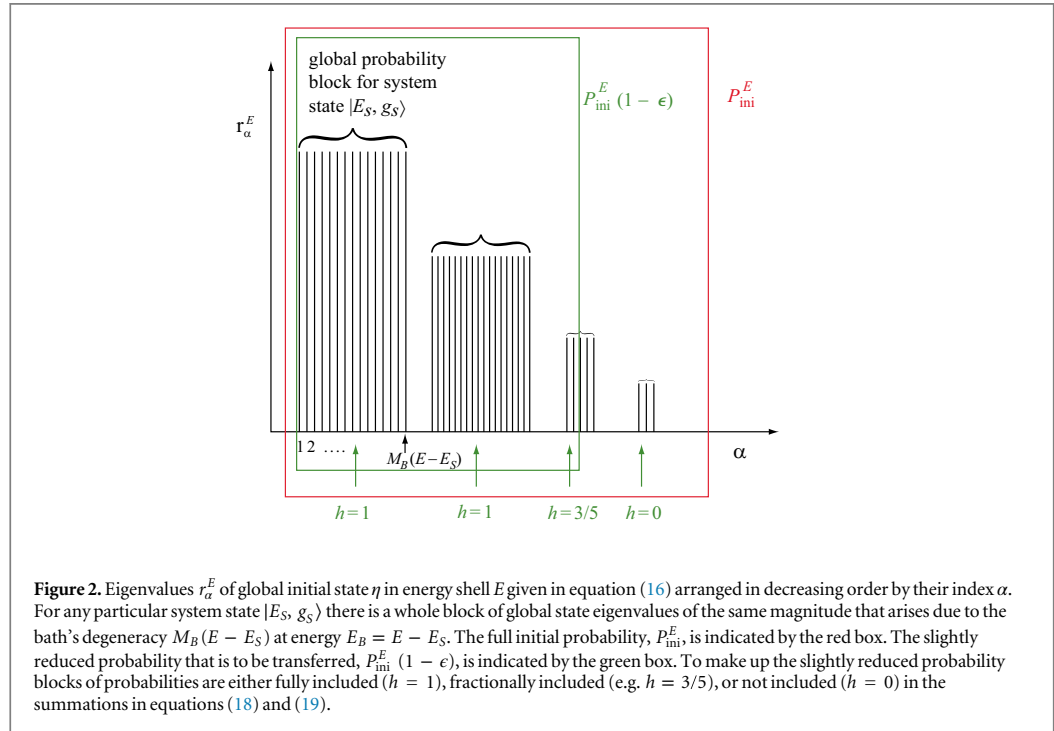
In the following we aim to specify the dimension $d_{\mathrm{ini}}^E(\epsilon, \eta)$ of a sub-projector of $\hat{\Pi}_{\mathrm{ini}}^E$ in whose associated subspace lives the fraction $1 - \epsilon$ of the initial state (η) population (P_{ini}^E) in energy shell E. For $\epsilon = 0$ and ρ_{SB} a full rank state the dimension of $\hat{\Pi}_{\mathrm{ini}}^E$ itself is just

$$d_{\mathrm{ini}}^E(0) = \sum_{E_S} M_S\left(E_S\right) M_B\left(E - E_S\right) = M_B(E) Z_S. \tag{14}$$

However, to determine the initial dimension $d_{\mathrm{ini}}^E(\epsilon, \eta)$ for a finite failure rate $\epsilon > 0$ we will need to specify the global initial state η a little more. While not the only choice, here we consider the class of initial global product states previously discussed [1, 15, 16]

$$\eta = \rho_S \otimes \tau_B \otimes |0\rangle_W \langle 0|, \tag{15}$$

where ρ_S is a system state that is diagonal in the basis of the system Hamiltonian and the bath is in a thermal state τ_B. We will label the eigenvalues of ρ_S when diagonalized in the system energy eigenbasis by $\lambda(E_S, g_S)$ where E_S is the corresponding energy and g_S the degeneracy index. By construction all *non-zero* eigenvalues of η in the energy shell E, i.e. the non-zero eigenvalues of $\hat{\Pi}_{\mathrm{ini}}^E \eta \hat{\Pi}_{\mathrm{ini}}^E$, are then

Figure 2. Eigenvalues r_α^E of global initial state η in energy shell E given in equation (16) arranged in decreasing order by their index α. For any particular system state $|E_S, g_S\rangle$ there is a whole block of global state eigenvalues of the same magnitude that arises due to the bath's degeneracy $M_B(E - E_S)$ at energy $E_B = E - E_S$. The full initial probability, P_{ini}^E, is indicated by the red box. The slightly reduced probability that is to be transferred, $P_{ini}^E(1 - \epsilon)$, is indicated by the green box. To make up the slightly reduced probability blocks of probabilities are either fully included ($h = 1$), fractionally included (e.g. $h = 3/5$), or not included ($h = 0$) in the summations in equations (18) and (19).

$$r^E\left(E_S, g_S, f_B\right) = \lambda\left(E_S, g_S\right) \frac{e^{-\beta\left(E - E_S\right)}}{Z_B}. \tag{16}$$

Here $f_B = 1, ..., M_B$ is the degeneracy index of the bath at energy $E_B = E - E_S$; but it does not affect the magnitude of η's eigenvalues. The eigenvalues thus have a high multiplicity, given by the bath multiplicity M_B. The eigenvalues can be re-labeled with $\alpha = 1, ..., d_{ini}^E(0)$ as $r^E\left(E_S, g_S, f_B\right) = r_\alpha^E$, such that they are arranged in decreasing order, $r_1^E \geqslant r_2^E \geqslant r_3^E ...$, see figure 2. The spectra of different energy shells E, as sketched in figure 2, differ only by a factor of $e^{-\beta E}$ for the individual values while their multiplicity M_B differs by a factor of $e^{\beta E}$. The population probability of the global state η in energy shell E is given by $\sum_{\alpha=1}^{d_{ini}^E(0)} r_\alpha^E = P_{ini}^E$.

Now we are ready to identify the dimension $d_{ini}^E(\epsilon, \eta)$ of the subspace within the energy shell E where most of the initial probability, $P_{ini}^E(1 - \epsilon)$, lives. In the sum above one simply has to add up the α to $d_{ini}^E(\epsilon, \eta) \leqslant d_{ini}^E(0)$ which must be chosen such that the probability reduces to the desired level

$$\sum_{\alpha=1}^{d_{ini}^E(\epsilon, \eta)} r_\alpha^E \geqslant P_{ini}^E(1 - \epsilon). \tag{17}$$

Comparing with equation (14) the final dimension is given by

$$d_{ini}^E(\epsilon, \eta) = \sum_{E_S, g_S} M_B\left(E - E_S\right) h\left(E_S, g_S, \epsilon\right) \tag{18}$$

where $h(E_S, g_S, \epsilon)$ determines which terms are included in the sum, as visualized and explained in figure 2. An h-value of 1 (0) indicates that a whole block of global eigenvalues all corresponding to system state $|E_S, g_S\rangle$ is (not) included in the summation, while a fractional h-value indicates that a fraction of the eigenvalues in a block is included, see figure 2.

The value of $d_{ini}^E(\epsilon, \eta)$ must be chosen as the smallest *integer* that fulfills (17). It can happen that ϵ has a value that would require to split up an individual eigenvalue r_α^E in order to fulfill (17) as an equality, i.e. $\sum_{\alpha=1}^n r_\alpha^E + a\, r_{n+1}^E = P_{ini}^E(1 - \epsilon)$ where n is an integer and $0 < a < 1$ a real fraction. In this case the eigenvalue r_{n+1}^E must be fully included, i.e. h must be chosen as the larger proper fraction available. For example, having to split the 3rd of 5 eigenvalues in a block (E_S, g_S) the associated $h(E_S, g_S, \epsilon)$ must be chosen 3/5 rather than 2/5, in order to guarantee that the failure probability is strictly less than ϵ.

For a map that is identical to (4) up to a failure probability ϵ we have obtained the initial and final subspace dimensions for each energy shell E. They allow us to check the maximal work extraction condition equation (9)

for each E. One has

$$d_{\text{fin}}^E = M_B(E)\, e^{-\beta w}\, Z_S \;\geqslant\; \sum_{E_S, g_S} M_B(E)\, e^{-\beta E_S}\, h\left(E_S, g_S, \epsilon\right) = d_{\text{ini}}^E(\epsilon, \eta)$$

$$w \leqslant -\frac{1}{\beta}\, \ln \sum_{E_S, g_S} e^{-\beta E_S}\, h\left(E_S, g_S, \epsilon\right) - F(\tau) = w_\epsilon^{\max}. \tag{19}$$

The last line shows that because of the trivial dependence of $d_{\text{ini}}^E(\epsilon, \eta)$ on E, the latter drops out and the expression for the maximal extractable work stated in equation (1) is recovered [1]. Note, that the definition of h has been extended from a binary function, used in equation (1), to include rational fractions. The optimal work described by equation (19) is thus marginally larger than the work described by equation (1). To recap, this work is maximal under variation of the global unitaries V that realize the map (4). The bound is tight when ϵ is chosen such that no splitting of eigenvalues would be necessary.

Note, that using the eigenvalues of the thermal state, τ_S, denoted by $t(E_S) := \frac{e^{-\beta E_S}}{Z_S}$, the maximal extractable work (1) can also be written as

$$w_\epsilon^{\max} = -\frac{1}{\beta}\, \ln \sum_{E_S, g_S} t\left(E_S\right) h\left(E_S, g_S, \epsilon\right). \tag{20}$$

Clearly, if h was 1 for all its arguments, i.e. all the eigenvalues in figure 2 have to be included, then the maximal extractable work is 0. For $\epsilon = 0$ no probability can be lost and $h = 1$ for all energy levels with non-zero population. Work can thus only be extracted if the initial state has rank less than the system Hilbert space dimension or $\epsilon > 0$ or both.

4. The final system state

The work extraction protocol, equation (4), has mapped $\eta \to \eta'(V)$. It is natural to ask in what reduced state $\sigma_S = \text{tr}_{BW}[\eta'(V)]$ the system is left as a result. We will here answer this question.

A crucial point to note is that not a single, unique unitary V corresponds to maximum work extraction but that there are infinitely many. Any unitary that maps the projective energy subspaces $\hat{\Pi}_{\text{ini}}^E(\epsilon)$, with dimensions $d_{\text{ini}}^E(\epsilon, \eta)$ calculated in section 3, onto the projective subspaces $\hat{\Pi}_{\text{fin}}^E$ will give rise to maximum work extraction. The large set of unitaries with this property forms a subset of all allowed unitaries, i.e. those that commute with H. As a first step we look at the average final state, obtained through the application of each of these unitaries and integration over the respective Haar measure, which we denote by $\langle\langle \cdot \rangle\rangle$. Details on integration over Haar unitaries can be found e.g. in [26]. Note that even though the unitaries V commute with H, $\eta'(V)$ itself need not be diagonal w.r.t. to the product basis of local energy eigenstates that defines the energy shells. While H and $\eta'(V)$ share a simultaneous eigenbasis there is no guarantee, due to H having degeneracies, that this is the product basis of local energy eigenstates and thus $\eta'(V)$ can have off-diagonal elements in the degenerate blocks of H. For the diagonal part of $\eta'(V)$ w.r.t. the above basis consisting of products of energy eigenstates of individual subsystems, $\eta'_{\text{diag}}(V)$, one finds

$$\langle\langle \eta'_{\text{diag}} \rangle\rangle \propto e^{-\beta H_S} \otimes e^{-\beta H_B} \otimes |w\rangle_W \langle w|, \tag{21}$$

see appendix A for details of the derivation. This is a factorized state where the system and bath are each in thermal states at inverse temperature β. However, this result in itself is not sufficient to conclude that typically an almost thermal state of the system results from an individual V. One needs to calculate also the relative variances of $P_{\text{fin}}(E, E_S, g_S)$. The latter are the portions of the diagonal elements at E_S, g_S of the final reduced state of the system σ_S that correspond to the parts of the total final state living in energy shell E, i.e., $\langle E_S, g_S| \sigma_S |E_S, g_S\rangle = \sum_E P_{\text{fin}}(E, E_S, g_S)$. Hence, if these relative fluctuations are small one will typically get diagonal elements of σ_S that resemble their averages, which in turn coincide, according to equation (21), with those of a thermal state. Concretely we obtain for those relative fluctuations, see appendix A

$$\frac{\Delta P_{\text{fin}}\left(E, E_S, g_S\right)}{\langle\langle P_{\text{fin}}\left(E, E_S, g_S\right)\rangle\rangle} \approx \frac{1}{\sqrt{M_B\left(E - E_S - w\right)}}, \tag{22}$$

which is indeed very small due to the large multiplicity of the bath $M_B(E - E_S - w) = M_B(E)\, e^{-\beta(E_S + w)}$ for energies $E \gg E_S + w$. Off-diagonal elements are considered separately, as detailed in appendix A, but it turns out that the variance to their vanishing mean also scales as $\frac{1}{\sqrt{M_B(E - E_S - w)}}$. Thus, for a unitary drawn at random all

elements of a typical final system state σ_S are $\frac{1}{\sqrt{M_B(E - E_S - w)}}$ close to the average thermal state if the bath is large. The entire above reasoning is along the lines of what has become known as 'quantum typicality' [27–30].

Having realized that optimal work extraction can result in a multitude of final states, with typical state being close to the thermal state, one may wonder why it is not possible to extract more work in those rare instances where the final state happens to be fairly different from the thermal state. Here it is important to note that the final state need not to be factorized between system and bath (see equation (21) only describes the diagonal part). Indeed, it is known [31–35] that only a negligible set of the final states will be factorized, especially when the degeneracies of the bath *B* are large. However, to repeat the same work extraction process requires an initially fully factorized state, thus preventing repetition and further work extraction along the same route for the majority cases.

Finally, a concern may be that one could introduce a 'fresh bath' that is uncorrelated with the final system in order to extract even more work from a non-thermal final system state. Here it is important to realize that the work extraction process with the first bath left the system not necessarily in a thermal state, but certainly in a state with full rank, σ_S. This is due to the fact that maximum work extraction requires a non-zero population of every dimension spanned by $\hat{\Pi}_{\text{fin}}^E$, see equation (10). As can be seen from equation (3) a full rank system state means that there is no potential for further 'perfect' work extraction, i.e. moving weight population from $|0\rangle_W \rightarrow |w\rangle_W$ with certainty ($\epsilon = 0$), even if new baths are brought in. If the first work extraction process is imperfect ($\epsilon \neq 0$), i.e. the weight ends with population in $|0\rangle_W$ and $|w\rangle_W$, in general weight and system will be correlated, thus also violating the initial factorization condition, see (4), for a second work extraction process. Nevertheless, for completeness, suppose that the result factorizes and the weight is chosen such that it has additional energy levels $|w_2\rangle$ and $|w + w_2\rangle$, where w_2 is adjusted to match $w_2 = w_\epsilon^{\max}(\sigma_S)$ of equation (20) applied to σ_S. One can now indeed perform a second imperfect work extraction which may very well increase the mean energy in the weight system. However, this will lead to a weight level population of at least three of the four eigenlevels, $|0\rangle_W$, $|w\rangle_W$, $|w_2\rangle_W$, $|w + w_2\rangle_W$. This scenario is outside of what is allowed in case (a), see figure 1(a), but will be discussed as case (b) in section 6. If, on the other hand the first work extraction process is a perfect one, weight and system necessarily factorize and a second, imperfect, work extraction can be added. This will result in overall imperfect transfer of $|0\rangle$ to $|w_\epsilon^{\max}(\rho_S)\rangle_W$ as a result of the total process instead of raising perfectly to $|w_0^{\max}(\rho_S)\rangle_W$ in the first step alone. So one could have obtained the same result if one had allowed for imperfect work extraction in the first place since an increase of ϵ always allows for an increase of work, $w_\epsilon^{\max}(\rho_S) > w_0^{\max}(\rho_S)$. This shows that there is no second law conflict with a final state being non-thermal.

5. Minimal work cost of formation

A key finding of the single shot thermodynamics approach presented in [1] is the realization that, in analogy to entanglement of formation and distillation, the work required to form a non-equilibrium state may differ from the work that can be extracted from that state. Here we will now re-derive the minimal cost of formation, $w^{\min}$, of engineering with certainty ($\epsilon = 0$) a diagonal state σ_S from a thermal state $\tau_S = \frac{e^{-\beta H_S}}{Z_S}$, under thermal operations.

Considering again global unitaries V that commute with the Hamiltonian, the desired operation is

$$\eta := \tau_S \otimes \tau_B \otimes |w\rangle_W \langle w| \xrightarrow{V} \sigma_{SB} \otimes |0\rangle_W \langle 0| := \eta', \tag{23}$$

with any final state permitted such that $\text{tr}_B[\sigma_{SB}] = \sigma_S$. Using the same reasoning that lead to equation (16) the eigenvalues $r^E(E_S, g_S, f_B, E_W)$ of the global initial state η in a particular energy shell E are given by

$$r^E\left(E_S, g_S, f_B, E_W\right) = \frac{e^{-\beta(E-w)}}{Z_S Z_B} \delta_{E_W, w}, \tag{24}$$

where g_S and f_B are the degeneracy indices labelling the various energy eigenstates for E_S and $E_B = E - E_S - E_W$, respectively. By construction the global eigenvalues are only non-zero when the weight has energy $E_W = w$ and the non-zero eigenvalues in shell E are all equal.

The eigenvalues of the desired final system state σ_S which is assumed diagonal in the system's energy eigenbasis $\{|E_S, g_S\rangle\}$ of H_S are denoted by $s(E_S, g_S) = \text{tr}[\sigma_{SB} |E_S, g_S\rangle\langle E_S, g_S|]$. The contributions stemming from the energy shell E to the diagonal elements of the final system state σ_S may thus be written as

$$s^E\left(E_S, g_S\right) = \sum_{f_B\left(E - E_S\right)} \langle 0|\langle E - E_S, f_B|\langle E_S, g_S| \ \eta' \ |E_S, g_S\rangle|E - E_S, f_B\rangle|0\rangle, \tag{25}$$

where the sum is performed over all bath degeneracy indices at energy $E_B = E - E_S$. We note that since the initial state has no off-diagonal elements between different system energies, say, E_S and E_S', and the unitary V

commutes with H no such off-diagonal elements can be generated under V. There may be off-diagonal elements within a single energy subspace E_S, i.e. $\langle E_S, g_S | \sigma_S | E_S, g_S' \rangle \neq 0$. These off-diagonal elements can always be 'avoided' though by choosing the basis $\{|E_S, g_S\rangle\}$ for the system subspace of energy E_S appropriately. Thus, w.l.o.g. the state σ_S can be assumed diagonal with diagonal elements as given in (25).

Since V is unitary, not only the eigenvalues of the initial state η (in energy shell E) but also those of the final state η' (in energy shell E) are given by equation (24), i.e. they are either 0 or $\frac{e^{-\beta(E-w)}}{Z_S Z_B}$. The summands on the rhs of (25) are just expectation values of the state η' and since any expectation value of a state is upper bounded by the largest eigenvalue of the respective state we find from (25)

$$s^E\left(E_S, g_S\right) \leqslant \sum_{f_B\left(E-E_S\right)} \frac{e^{-\beta(E-w)}}{Z_S Z_B} = \frac{e^{-\beta(E-w)}}{Z_S Z_B} M_B\left(E - E_S\right) = \frac{e^{-\beta\left(E_S-w\right)}}{Z_S} \frac{M_B(E) e^{-\beta E}}{Z_B}. \tag{26}$$

The probability of the final state σ_S to be found in state $|E_S, g_S\rangle$, denoted by $s(E_S, g_S)$, is then obtained by summing over contributions from all the energy shells

$$s\left(E_S, g_S\right) = \sum_E s^E\left(E_S, g_S\right) \leqslant \frac{e^{-\beta\left(E_S-w\right)}}{Z_S}, \tag{27}$$

where $\sum_E M_B(E) e^{-\beta E} = Z_B$ was used. Importantly this inequality must be fulfilled for all pairs E_S, g_S. Using the eigenvalues of the thermal state of the system and rearranging one obtains

$$\forall E_S, g_S : w \geqslant \frac{1}{\beta} \ln \frac{s\left(E_S, g_S\right)}{t\left(E_S\right)} \tag{28}$$

as the condition on the work cost of formation. One can see now that the energy eigenstate $|E_S^\star, g_S^\star\rangle$ with the largest ratio $\frac{s(E_S^\star, g_S^\star)}{t(E_S^\star)} = \max_{E_S, g_S} \frac{s(E_S, g_S)}{t(E_S)} =: \mu$ is the one that constraints the whole transformation. The tight lower bound on the work cost of formation is then

$$w \geqslant \frac{1}{\beta} \ln \mu =: w^{\min}. \tag{29}$$

μ can also be expressed as $\mu = \min\{\lambda : s(E_S, g_S) \leqslant \lambda \, t(E_S) \; \forall \, E_S, g_S\}$ or

$$\mu = \min\left\{\lambda : \sigma_S \leqslant \lambda \, \tau_S\right\} \tag{30}$$

for σ_S the desired diagonal state of the system. The work cost equation (29) for the ideal map equation (23) can easily be extended to allow final states in an ϵ vicinity of the desired state σ_S. μ is then replaced by $\mu^\epsilon = \inf_{\sigma_S^\epsilon} \min\{\lambda : \sigma_S^\epsilon \leqslant \lambda \, \tau_S\}$ where σ_S^ϵ is close to σ_S, $\|\sigma_S^\epsilon - \sigma_S\| \leqslant \epsilon$, where $\| \cdot \|$ is the trace norm [23].

6. Generalization of work extraction to multiple levels in the final state of the work storage system

We have seen in the previous section that the single shot concept gives sensible results for work extraction when single levels of the weight are being considered (case (a) in figure 1). While for truly microsopic systems this may be a suitable analysis, case (a) is not the right picture for mesoscopic or macroscopic work extraction process. It is not feasible to lift a work storage system from a true quantum energy eigenstate to another true energy eigenstates. The density of states is so large that it will be practically impossible to pick a single eigenstate of a macroscopic system [27]. This indicates that an extension of the single shot thermodynamic framework to transitions between multiple levels, see figures 1(b) and (c), is desirable.

The approach presented above now allows a straightforward generalization to thermal processes where the work storage system is transferred from the ground state to not just a single energy level, but an interval δ of energy levels, see figure 1(b). To do so we will now turn to a harmonic oscillator work storage system with energy spacing ΔE, similar to the one considered in [16], and differing from the qubit work storage system used in [1]. For simplicity we will assume that also the system and the bath have oscillator-like equidistant energy eigenlevels of integer multiples of ΔE. Again, the weight shall have no degeneracies whatsoever while the system and bath again have degeneracies $M_S(E_S)$ and $M_B(E_B)$ at their energies E_S and E_B, respectively.

Work extraction of an amount of work $\bar{w}_\epsilon$ is now identified with the transition (with probability $1 - \epsilon$) of the work storage system from its ground state to any state living in the energy subspace $[\bar{w}_\epsilon, \bar{w}_\epsilon + \delta]$, see figure 1(b). The ideal transfer ($\epsilon = 0$) is now associated with the map

$$\eta := \rho_S \otimes \tau_B \otimes |0\rangle_W \langle 0| \xrightarrow{V} \sigma_{SBW},$$ (31)

where the final state of the weight $\sigma_W := \mathrm{tr}_{SB}[\sigma_{SBW}]$ is now mixed with support in the interval specified. Here the general case of entangled states between weight and the rest is allowed. If the role of these correlations is of interest, then restrictions to product states could be considered.

To account for the different (larger) space of possible final states one simply has to restart the calculation of $\hat{\Pi}_{\mathrm{fin}}^E$ and d_{fin}^E, as detailed in section 3, and enlarge the final Hilbert space in order to comprise all final global states with the weight in the specified local energy subspace. This 'new' final Hilbert space can be described as a sum over many 'old' ones. Restarting from (11) one can introduce a pertinent summation over weight energy levels so that the new d_{fin}^E reads:

$$d_{\mathrm{fin}}^E = M_B(E)\, Z_S \sum_{E_W = \bar{w}}^{\bar{w}+\delta} e^{-\beta E_W}.$$ (32)

The sum on the rhs now depends on the level spacing ΔE which did not play any role so far

$$\sum_{E_W=\bar{w}}^{\bar{w}+\delta} e^{-\beta E_W} = \sum_{n=\bar{w}/\Delta E}^{(\bar{w}+\delta)/\Delta E} e^{-\beta \Delta E\, n} = \frac{e^{-\beta \bar{w}}\left(1 - e^{-\beta(\delta+\Delta E)}\right)}{1 - e^{-\beta \Delta E}},$$ (33)

where n is a natural number that labels the eigenvalues in the allowed final interval. Substituting this for $e^{-\beta w}$ in (11) then leads to

$$\tilde{w}_\epsilon^{\max} = w_\epsilon^{\max} + \frac{1}{\beta}\, \ln \frac{1 - e^{-\beta(\delta+\Delta E)}}{1 - e^{-\beta \Delta E}}.$$ (34)

Increasing the width δ of the allowed final energy regime for the weight beyond a single energy level ($\delta > 0$) leads to an additional term with respect to the single final energy level situation characterized by equation (1). For large energy intervals (δ) and small energy spacings (ΔE) the last expression quickly converges to

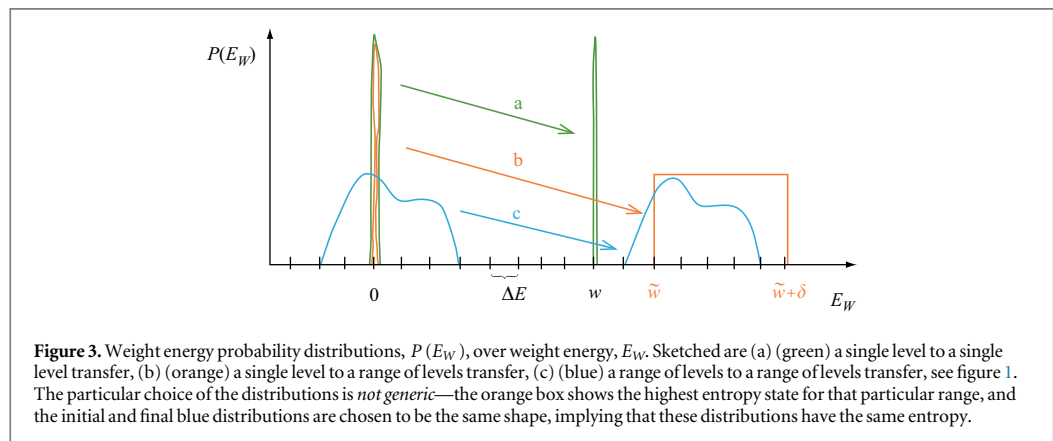$$\tilde{w}_\epsilon^{\max} = w_\epsilon^{\max} - \frac{\ln(\beta \Delta E)}{\beta}.$$ (35)

Expression (35) holds under the conditions that compared to the 'thermal energy', $\frac{1}{\beta}$ the final energy range is large, $\delta \gg \frac{1}{\beta}$ and the energy spacing of the work storage system, ΔE, is small $\Delta E \ll \frac{1}{\beta}$. Physically this means that the final energy range δ has to comprise a large number of energy eigenstates of the work storage system.

We note that the extractable work $\tilde{w}_\epsilon^{\max}$ for case (b) in figure 1 is *increased* in comparison to the extractable work $w_\epsilon^{\max}$ of case (a) in the limit considered ($\Delta E \ll 1/\beta$). The surplus on the rhs does not depend on δ as long as the above conditions are fulfilled. However, the work does vary with the level spacing and can go up to infinity for $\Delta E \to 0$.

To judge whether these conditions are fulfilled for meso/macro-scopic work storage systems, it is instructive to consider realistic numbers. At room temperature the thermal energy is $\frac{1}{\beta} \approx 2.5 \times 10^{-2}$ eV; about the energy of a single optical phonon in a solid. Thus allowing a final energy range of δ for the macroscopic weight system that is large compared to the energy of a single phonon, condition $\delta \gg \frac{1}{\beta}$ is fulfilled. For the work storage system one can imagine a pendulum with eigenfrequency 1 Hz which results in an energetic level spacing of $\Delta E \approx 4 \times 10^{-15}$ eV $\ll \frac{1}{\beta}$ thus fulfilling the second condition. Level spacings of many meso/macro-scopic work storage systems, such as batteries, are even smaller by several orders of magnitude. These example numbers indicate that both conditions will generally hold for meso/macro-scopic systems.

Maybe surprisingly the additional term, $-\ln(\beta \Delta E)/\beta$, is an up-shift as sketched in figure 1(b), implying a work above that of equation (1). This seems to indicate a conflict of equation (35) with the second law, but this is only a problem in a naive sense. To properly discuss second law violations one would have to build a cyclic machine that enables work extraction. It is possible to close the desired work extraction transformation, equation (31), to a cycle, however, it is not possible to run the cycle again. This is because the weight started in an energy eigenstate, $|0\rangle_W$, while ending in a reduced state σ_W which is *not* an energy eigenstate. Therefore there is no true second law violation as the process cannot be repeated using the same weight. Exploring if there is a link between the apparent violation of the second law in equation (35) and the existence of weight-system or system-bath correlations [36–38] in the final state in equation (31) is an interesting future avenue.

Finally, the implication of equation (35) is that one can extract *more work* when the final weight state is allowed to live in a range of energy levels (b), rather than when restricting to a single level (a). While this is mathematically sound one may find it physically counterintuitive as single shot work is colloquially often associated with a worst case scenario. The sketch in figure 1(b) intuitively suggests that the worst case result for jumping from the ground state to a range of energy states is a single level transition to the lowest level, implying

Figure 3. Weight energy probability distributions, $P(E_W)$, over weight energy, E_W. Sketched are (a) (green) a single level to a single level transfer, (b) (orange) a single level to a range of levels transfer, (c) (blue) a range of levels to a range of levels transfer, see figure 1. The particular choice of the distributions is *not generic*—the orange box shows the highest entropy state for that particular range, and the initial and final blue distributions are chosen to be the same shape, implying that these distributions have the same entropy.

the lowest energy in (b) is the same as the single energy in (a), $\tilde{w} = w$. However, this is not correct — the salient point is that populations of energetic levels higher than w would not count as 'success' in the situation depicted in figure 1(a). To achieve a high probability of success, $1 - \epsilon$, in (a) the population needs to be concentrated in just that one level and this is what leads to a lower amount of extractable work in comparison to case (b), i.e. $\tilde{w} > w$. This new result may appear to contrast with a recent paper [39] that shows that allowing population in higher levels should not change the 'predicted' work. However, the work definition used there follows traditional statistical physics concepts, without including a weight explicitly and involves sequences of shifting energy levels in the system Hamiltonian and coupling the system to a bath [15, 40]. The work concept thus differs significantly from the work concept used here.

Now the question is how significant it is if the weight ends up exactly with energy, say w, or with an energy in a range, say $[w, w + \delta]$? The second case is of relevance for many mechanical processes in physics and engineering where gaining work is the key aim, and obtaining a little more energy would be judged positively. There are other situations where *resonance* with a particular energy value, i.e. case (a), is key for performance, for instance, this may be the case in biological energy conversion processes, such as photosynthesis [41]. We propose that the work 'extracted' in such resonance processes may be described as 'resonance work' or 'matched work'. To decide which approach to use to calculate work, e.g. the single shot approach with single level transitions (a) or multiple level transitions (b), or other approaches, one has to first identify what is really needed — energy of a certain amount or energy above a certain threshold.

7. Generalization of work extraction to multiple levels in the initial and final state of the work storage system

We found that the work definition based on a 'single energy level to single energy level transition' of a weight, case (a) in figure 1, is too narrow and at odds with our notion of classical work storage systems. A physical interpretation was however possible in the context of matching a particular energy, i.e. in a resonance situation. For the physically motivated extension to multiple final levels, case (b) in figure 1, that corresponds to work extraction in the sense of 'at least this much or more energy' the predictions came out unphysical. So one is left with the unsatisfactory situation that there is no consistent single shot work extraction concept for general initial and final states of the weight. Such extension issues have been recognized by others and motivated a number of recent publications that provide bounds on the single shot work from majorization [39] and fluctuation theorems [42], address the shifting of general weight probability distributions [16], and employ axiomatic approaches to identify sets of monotone functions that can act as work quantifiers for specific situations [43, 44].

To try and rectify this situation we here proceed to allow multiple energy levels in both, the initial and final state, as indicated in figures 1(c) and 3(c). Here one now faces the problem of defining work on the basis of an initial and a final *energy distribution of the weight*. When the weight transfers from multiple levels to other multiple levels without a means to distinguish any of these transitions, there is no single energy difference ΔE_W that can be straightforwardly linked to work. Even if one considers the multitude of energy differences associated with the different single level transitions, it is challenging to formulate a 'worst case scenario' for the reasons discussed in the previous section. We will here not attack the complicated problem of defining single shot work for transfers between multiple levels of the weight, see figures 1(c) and 3(c), and finding the optimum work. Instead, our aim here will be to find a *transfer-quantity*, which we denote $\langle w \rangle$, that indicates whether an initial weight energy distribution can or can not be transformed into a final distribution. To analyse the general

multiple level scenario we consider again a global unitary (V) commuting with the sum of local Hamiltonians as before and no restrictions on their spectra. V maps an initially factorized state to some final possibly correlated state

$$\eta := \rho_S \otimes \tau_B \otimes \sigma_W \xrightarrow{V} \eta', \tag{36}$$

where the initial bath state is assumed thermal and the weight's final reduced state is $\sigma_W' := \mathrm{tr}_{SB}[\eta']$. Various transfers between different energy distribution of the weight are sketched in figure 3.

The average energy change of the weight on its own cannot be significant to limit thermodynamic transformations. For example, a weight starting in the ground state can have its average energy raised just by coupling it to a heat bath. Intuitively, the transfer-quantity should reward energy increase while punishing spreading the energy across different energy levels of the weight. It is known that the free energy has exactly such a balancing property and this motivates us to define the transfer-quantity for bringing a weight from state σ_W to state σ_W' as the free energy difference of the weight

$$\langle w \rangle := \Delta F_W = \Delta U_W - T \Delta S_W. \tag{37}$$

Here $\Delta U_W = \mathrm{tr}[H_W(\sigma_W' - \sigma_W)]$ and $\Delta S_W / k_B = -\mathrm{tr}[\sigma_W' \ln \sigma_W'] + \mathrm{tr}[\sigma_W \ln \sigma_W]$ are the average energy change and entropy change of the weight and T is the temperature of the initial bath state. Using sub-additivity of the von Neumann entropy it is straightforward to show, see appendix B, that the *transition equation (36) can only be possible when the transfer-quantity obeys at least the following inequality*

$$\langle w \rangle \leqslant F(\rho_S) - F(\rho_S'). \tag{38}$$

Here $F(\rho_S) - F(\rho_S') = -\Delta F_S = -\Delta U_S + T \Delta S_S$ is the free energy difference of the two, in general, non-equilibrium system states ρ_S and $\rho_S' = := \mathrm{tr}_{BW}[\eta']$, with ΔU_S and ΔS_S defined analogously to above, using instead the system Hamiltonian H_S. Note, that inequality equation (38) is necessary but not sufficient, i.e. there are examples where this inequality is fulfilled and the transition may still not be possible. Equality in equation (38) can only occur when the final global state factorizes with the final bath state being a thermal state, see appendix B. Thus the bath B must remain in a thermal state under the optimal work extraction process. This clearly hints in the direction of the bound being reachable in the classical macroscopic limit as was suggested in [1].

Now one can consider special cases of this general relation. A subclass of map equation (36) are the single level to single level transitions, sketched in figure 1(a) and in figure 3(a). The necessary condition for a transformation $\rho_S \rightarrow \rho_S'$ to be thermodynamically allowed here becomes

$$\langle w \rangle \overset{a}{=} w \leqslant F(\rho_S) - F(\rho_S'). \tag{39}$$

To maximize w one clearly wants to choose the final state thermal, $\rho_S' = \tau_S$, resulting in the highest permissible value $w^{\max} = F(\rho_S) - F(\tau_S)$. In general one has $F_\epsilon^{\min}(\rho_S) < F(\rho_S)$, and hence $w_\epsilon^{\max} < w^{\max}$, with $F_\epsilon^{\min}(\rho_S)$ approaching $F(\rho_S)$ in the independent identical (i.i.d.) limit (infinitely many identical copies) [1]. This means that inequality equation (39) is here a lesser requirement than the tight bound equation (1), giving just an upper threshold of what transformations may be allowed.

For weight transitions from a single level to multiple levels, sketched in figure 1(b), the inequality gives

$$\langle w \rangle \overset{b}{=} F(\sigma_W') \leqslant F(\rho_S) - F(\rho_S'). \tag{40}$$

So it is down to the choice of a suitable final weight state σ_W' to enable the transformation to be thermodynamically possible. Figure 3(b) shows an example of a final energy distribution of the weight (orange box). It can be seen that if one was to shift up the final distribution in energy while keeping the entropy the same, $F(\sigma_W')$ would grow thus making the inequality continuously tighter. The limit on how far the final weight distribution can be shifted in energy and the transition still being allowed is given by the system's free energy difference.

For a multiple level to multiple level transition, sketched in figure 1(c), an example initial weight distribution is sketched in figure 3(c). If the energy distribution of the weight is solely shifted to a higher energy without changing its shape or entropy, as indicated in blue in figure 3(c), then the inequality reduces to

$$\langle w \rangle \overset{c}{=} \Delta U_W \leqslant F(\rho_S) - F(\rho_S'). \tag{41}$$

This result coincides with the *tight* bound derived for a more restrictive process in which global unitaries are chosen to be of a specific form but for which only average global energy conservation is required [16].

Condition equation (38) produces sensible results for the special cases discussed here. While for the transformation to be thermodynamically possible fulfilling the condition is only necessary — not sufficient — the condition's strength lies in its applicability for general globally energy conserving transformations on initially factorized states. Apart from providing an easy-to-check criterion to rule out many potential weight transitions,

calling the transfer-quantity 'work' can be supported, see e.g. [43]. Indeed we observe that condition equation (38) for the transfer quantity $\langle w \rangle$ is of the same form as the well-known bound on the *average work* $\langle W \rangle$ that can be extracted from a statistical ensemble when changing its probability distribution ρ_S to another probability distribution ρ_S', e.g. in the context of Jarzynski's equality [3]

$$\langle W \rangle \leqslant -\Delta F_S, \tag{42}$$

which is widely viewed as the second law of thermodynamics.

8. Discussion

We have re-derived the extractable work and the work of formation in the single shot setting proposed in [1], where a work storage system (a) transfers from a single energy level to another single energy level with probability $1 - \epsilon$. Our alternative derivation of the optimal work value for which such a transformation is still possible was based on the comparison of subspace dimensions while no discussion of thermo-majorization was required. (Thermo-majorization [1] offers an additional and separate method of addressing the same optimization problem.) The approach presented here facilitates a discussion of the final state of the system after work extraction has been performed. Indeed, we find that *typically* the reduced final state is a minimum free energy thermal state as has been suggested [1]. 'Typically' is here to be understood in the very same sense in which it is often used in the context of thermalization studies: it is overwhelmingly frequent w.r.t. Haar distributed unitaries [27–30]. The implication of this finding is two-fold: on one hand it should not come as a surprise if a 'work-extracted' system is left in a thermal state, on the other hand the question of how far away from a thermal state a system may end up for a specific, for instance physically motivated, unitary is a promising direction for future research.

The presented approach opened the possibility to discuss a single shot work concept when transitions of the work storage system between (b) a single energy level and *a range of final energy levels* is permitted. Extensions to multiple levels are desirable to reflect experimental constraints — current and near-future experimental control of meso/macro-scopic systems does not allow to distinguish between energy levels that are spaced by less than an optical phonon. While the results for the multiple final level situation are mathematically sound, they raise interpretation issues when applied to the physical context. Besides showing an apparent violation of the second law that requires careful attention, an important conclusion emerges. The term 'work extraction' intuitively suggests that an energy of *a certain amount or higher* is being made available but the 'minimum extractable work' found for case (b) was in fact higher than the single level case (a) and thus at odds with physical intuition. The single shot work analysis thus turns out to be not a good setting for discussing 'work extraction'. It is however rather suited to characterize *resonance processes* where a specific amount of energy needs to be stored in the work storage system with high probability. Due to its nature, the energy stored may here be thought of as *resonance work*.

Finally, the aim was to consider multiple energy levels of the work storage system for both, initial and final state. We have not tried to generalize the single shot work extraction approach to this situation because of the issues in physical interpretation that we could not fully resolve. Instead we provide a means to decide whether a transfer can be thermodynamically allowed or not by introducing a transfer-quantity $\langle w \rangle$ that we chose as the *free energy of the weight*. The derived necessary bound, $\langle w \rangle \leqslant -\Delta F_S$, is applicable for transformations with general initial and final states of the weight. The bound is not tight in general, i.e. there are transformations that fulfil the bound while still not being possible with the thermodynamic resources considered here. However, the bound gives tight results for the single level case (a) when applied to $N \to \infty$ i.i.d copies [1, 15] and to a set of restricted global unitaries that results in multi-level to multi-level weight transfers [16]. Since the criterion is fairly simple to apply it facilitates the discussion of many practical limits on work extraction in the quantum regime. To rephrase, while equation (38) is not precisely tight it is sufficiently tight to be relevant and thus it may be considered a valuable concept.

Acknowledgments

Appendix A. Derivation of properties of the final states

This part of the appendix is dedicated to a more detailed derivation of the results on the final states discussed in section 4.

We start by denoting the projector spanned by all eigenstates of the initial state η that have total energy E and $h(E_S, g) = 1$ by $\hat{\Pi}_{\text{ini}}^E(\epsilon, \eta)$. The dimension of this projector $d_{\text{ini}}^E(\epsilon, \eta) = \text{tr}[\hat{\Pi}_{\text{ini}}^E(\epsilon, \eta)]$ has already been calculated in (14). By construction it equals the dimension d_{fin}^E of the respective projector $\hat{\Pi}_{\text{fin}}^E$. The maximum work extraction itself simply consists in a unitary V which maps the initial subspaces on the respective final ones, thereby respecting energy shells, which is the only condition on V. As mentioned in section 4 the crucial point here is that there is obviously not a single, unique unitary that corresponds to maximum work extraction but infinitely many. Since there are very many unitaries there may be potentially very many different reduced states of the system after maximum work extraction. In order to better understand the multitude of final system states it is instructive to consider unitaries which are randomly distributed according to the Haar measure. (Those unitaries may be constructed by drawing (column) vectors component-wise at random and doing a Gram-Schmidt orthonormalization For more details on averaging over Haar distributed unitaries see [26].) In the following we outline some results on distribution of quantities generated by such random unitaries without (formal) proof.

Consider the unitary mapping V of some matrix A and a subsequent projection onto its diagonal part w.r.t. some orthonormal basis $\{|n\rangle\}$, i.e., $\sum_n |n\rangle\langle n| \, VAV^\dagger \, |n\rangle\langle n|$. The average over all such unitaries may simply be inferred from symmetry considerations:

$$\left\langle\!\left\langle \sum_{n=1}^d |n\rangle\langle n| \, VAV^\dagger \, |n\rangle\langle n| \right\rangle\!\right\rangle = \frac{\text{tr}[A]}{d}\,\hat{\Pi}, \qquad \hat{\Pi} := \sum_{n=1}^d |n\rangle\langle n|, \tag{A.1}$$

where $\langle\!\langle \, \cdot \, \rangle\!\rangle$ denotes the average over all unitaries and d is the dimension of the Hilbertspace, $d = \text{tr}[\hat{\Pi}]$. Denoting the sum over $m\,(\leqslant d)$ of the diagonal elements of this operator by s, we find for the latter:

$$s := \sum_{n=1}^m \langle n| \, VAV^\dagger \, |n\rangle, \tag{A.2}$$

we find from (A.1) for the average of s:

$$\langle\!\langle s \rangle\!\rangle = \frac{m \, \text{tr}[A]}{d}. \tag{A.3}$$

It is more subtle to compute the variance of s but it has been done, e.g., in the context of quantum typicality [28, 30]. From the latter works we find:

$$(\Delta s)^2 := \langle\!\langle s^2 \rangle\!\rangle - \langle\!\langle s \rangle\!\rangle^2 = \frac{m\left(1 - \dfrac{m}{d}\right)\text{tr}[A]^2}{d(d+1)}, \tag{A.4}$$

and for the relative deviation (setting $d + 1 \approx d$ for simplicity)

$$\frac{\Delta s}{\langle\!\langle s \rangle\!\rangle} := \langle\!\langle s^2 \rangle\!\rangle - \langle\!\langle s \rangle\!\rangle^2 = \sqrt{\frac{1}{m}\left(1 - \frac{m}{d}\right)}. \tag{A.5}$$

Thus, if m is sufficiently large, those relative deviations become (negligibly) small. Which means even for a single unitary drawn at random according to the Haar measure s may be expected to be given by $s \approx \langle\!\langle s \rangle\!\rangle$ for the overwhelming majority of cases.

Equipped with these findings we may now identify the 'typical' local reduced final state of the system S. To this end we consider (A.1) and make the identifications:

$$A: \hat{\Pi}_{\text{ini}}^E(\epsilon, \eta)\, \eta\, \hat{\Pi}_{\text{ini}}^E(\epsilon, \eta), \qquad \hat{\Pi}: \hat{\Pi}_{\text{fin}}^E, \tag{A.6}$$

furthermore we note that

$$\text{tr}[A] = \text{tr}\!\left[\hat{\Pi}_{\text{ini}}^E(\epsilon, \eta)\, \eta\, \hat{\Pi}_{\text{ini}}^E(\epsilon, \eta)\right] = (1 - \epsilon)\, P_{\text{ini}}^E. \tag{A.7}$$

Let η'_{diag} be the diagonal part (w.r.t. the product energy eigenstates of the uncoupled subsystems) of the full system final state that has actually changed under work extraction, i.e. everything except for the ϵ-part that remained unchanged, and $\eta'_{\text{diag},E}$ the state's projection into the subspace of energy E. Then one finds from (A.1), inserting (A.6) and (A.7):

$$\langle\langle \eta'_{\mathrm{diag},E} \rangle\rangle = \frac{(1-\epsilon)\, P^E_{\mathrm{ini}}\, \hat{\Pi}^E_{\mathrm{fin}}}{d^E_{\mathrm{fin}}}. \tag{A.8}$$

Since P^E_{ini} is invariant under V it may be calculated most conveniently from (15). From this calculation it is found to be actually independent of E, i.e., the total probabilities on all energy shells are the same. Using (10) and (11) yields

$$\langle\langle \eta'_{\mathrm{diag},E} \rangle\rangle \propto \frac{1}{e^{\beta E}} \sum_{E_S} \hat{\Pi}_S\left(E_S\right) \otimes \hat{\Pi}_B\left(E - E_S - w\right) \otimes \hat{\Pi}_W\left(w\right). \tag{A.9}$$

At this point it is actually more convenient to go back to the original, local energies and sum over E_S, E_B rather than E_S, E. Doing so yields:

$$\langle\langle \eta'_{\mathrm{diag}} \rangle\rangle \propto \sum_{E_S, E_B} \hat{\Pi}_S\left(E_S\right) e^{-\beta E_S} \otimes \hat{\Pi}_B\left(E_B - w\right) e^{-\beta E_B} \otimes \hat{\Pi}_W\left(w\right). \tag{A.10}$$

Using the index shift $E_B - w \to E_B$, this may be summed as

$$\langle\langle \eta'_{\mathrm{diag}} \rangle\rangle \propto e^{-\beta \hat{H}_S} \otimes e^{-\beta \hat{H}_B} \otimes \hat{\Pi}_W\left(w\right) \tag{A.11}$$

which is (21) as given and discussed in section 4.

However, to conclude that the actual final state of the system S corresponding to a single random unitary is close to $e^{-\beta \hat{H}_S}/Z_S$ and thus a minimum free energy state it remains to be shown that the relative deviations of η'_{diag} are small. To this end it is instructive to write out the diagonal matrix elements of the reduced final system state $\sigma_S = \mathrm{tr}_{BW}[\eta']$ explicitly

$$\langle E_S, g| \, \sigma_S \, |E_S, g\rangle = \sum_E P_{\mathrm{fin}}\left(E, E_S, g\right),$$
$$P_{\mathrm{fin}}\left(E, E_S, g\right) := \sum_{f=1}^{M_B(E_B)} \langle E_S, g, E, f, w| \, \eta'_{\mathrm{diag}} \, |E_S, g, E, f, w\rangle, \tag{A.12}$$

where $|E_S, g, E, f, w\rangle$ are the respective (product) energy eigenstates of the decoupled system. So a single $P_{\mathrm{fin}}\left(E, E_S, g\right)$, i.e., a contribution to the respective diagonal element of the final reduced state of S from some energy shell E, is itself a sum over very many diagonal elements of the full final state, as may be seen from the lower line of (A.12). Namely, it runs over all bath energy eigenstates at $E_B = E - E_S - w$, i.e., contains $M_B(E_B)$ summands. Thus from (A.5), setting $m = M_B(E - E_S - w)$, it may be inferred that the relative deviations of the contributions to the diagonal matrix element are on the order of

$$\frac{\Delta P_{\mathrm{fin}}\left(E, E_S, g\right)}{\langle\langle P_{\mathrm{fin}}\left(E, E_S, g\right) \rangle\rangle} \approx \frac{1}{\sqrt{M_B\left(E - E_S - w\right)}}, \tag{A.13}$$

which is very small if the bath is large. This is (22) as given and discussed in section 4. Furthermore, if the unitaries are drawn at random, the $\hat{U}_E$ are independent of each other. Thus one may conclude from (A.13) and the central limit theorem, that according to the first line of (A.12), the relative deviations of $\langle E_S, g| \, \sigma_S \, |E_S, g\rangle$ will be even smaller if many energy shells are involved.

An analogous consideration (which we do not present here in full detail for brevity) applies to the off-diagonal elements of the final reduced state of S. The analogue to (A.3) required to find the average of those off-diagonal elements reads [20]

$$\langle\langle \langle n| \, VAV^\dagger \, |n'\rangle \rangle\rangle = 0 \quad \text{for} \quad n \neq n'. \tag{A.14}$$

Thus, using the same identifications as before (A.6), on average off-diagonal elements of the full system final state vanish. Since off-diagonal elements of reduced states are sums of off-diagonal elements of the full state exclusively, the averages of the final reduced state of the system σ_S also vanish. Thus on average, the final state is indeed diagonal in its energy eigenbasis. Furthermore only those (off-diagonal) elements of the full state for which the full system states $|n'\rangle$, $|n\rangle$ both correspond to the same bath system state $|E_B, f\rangle$ contribute to the reduced state of S at all. Thus each contribution to the reduced state of S may be associated with a certain energy shell E_B of the bath. An analogue to (A.4) quantifies the deviations of off-diagonal elements

$$(\Delta o)^2 := \langle\langle \, |\langle n \, |VAV^\dagger| \, n'\rangle|^2 \rangle\rangle \quad \text{for} \quad n \neq n', \tag{A.15}$$

and reads, according to e.g. [20]

$$(\Delta o)^2 := \frac{\operatorname{tr}\left[A^2\right]}{d\,(d+1)}. \tag{A.16}$$

From (A.6) and (A.1) it may be inferred that, $\operatorname{tr}[A^2] \leqslant \operatorname{tr}[\hat{\Pi}^E_{\text{ini}}(\epsilon, \eta)]$ and $d = \operatorname{tr}[\hat{\Pi}^E_{\text{fin}}]$. In section 6 is has been established that both traces essentially scale with the dimension of the bath, i.e.

$$\operatorname{tr}\left[\hat{\Pi}^E_{\text{ini}}(\epsilon, \eta)\right] \propto M_B(E), \quad \operatorname{tr}\left[\hat{\Pi}^E_{\text{fin}}\right] \propto M_B(E). \tag{A.17}$$

Thus we find

$$\Delta o \leqslant \frac{C}{\sqrt{M_B(E)}}. \tag{A.18}$$

where C is a number which is independent of the overall dimension of the bath. As $M_B(E) \geqslant M_B(E - E_S - w)$, this implies that the deviations of the off-diagonal elements of the reduced state of S again scale at most as $\frac{1}{\sqrt{M_B(E - E_S - w)}}$, just like in the case of the diagonal elements, see (A.13). This result is also given and discussed in section 4.

Appendix B. Derivation of upper bound to the state transfer quantity $\langle w \rangle$

This part of the appendix is dedicated to the derivation of (38) in section 7. In the remainder any state of the total system, possibly including correlations, will be denoted by η. Reduced local states will be denoted by $\rho_S = \operatorname{tr}_{BW}[\eta]$ for the non-equilibrium system, $\sigma_W = \operatorname{tr}_{SB}[\eta]$ for the weight storage system, and $\tau_B = \operatorname{tr}_{SW}[\eta]$ for the bath. Just as before V transforms the total initial state η into the total final state η' which means $\eta' = V\eta V^\dagger$. And just as before we assume conservation of the sum of local energies, i.e., $[H_S + H_B + H_W, V] = 0$. Free energies will be denoted as

$$F(X) := U(X) - \frac{S(X)}{\beta} \quad \text{with} \quad S(X) := -\operatorname{tr}[X \ln X], \tag{B.1}$$

where $X := \rho_S, \sigma_W, \tau_B, \eta$. Thus, free energies may refer either to parts or to the total system. The expectation value of the respective energy is denoted by $U(X)$ for the corresponding Hamiltonian, H_S, H_W, H_B. Generally primed operators will refer to final states and unprimed operators to initial states.

Again, as also mentioned in section 7, we assume factorizing initial conditions, i.e., $\eta = \rho_S \otimes \tau_B \otimes \sigma_W$. We also make the assumption of the initial state of the bath is thermal, $\tau_B \propto e^{-\beta H_B}$. The only difference in the set up of the approach in this section compared to sections 3–6 is that we do not require the state density of the spectrum of the bath to be exponentially growing. By construction, the total entropy and the sum of local energies are both invariant under V. Thus, the initial total free energy and the final total free energy are the same, $F(\eta) = F(\eta')$, and due to the factorizing initial conditions we have

$$F(\eta) = F(\rho_S) + F(\tau_B) + F(\sigma_W). \tag{B.2}$$

From the Araki–Lieb theorem [45] it follows that

$$S(\eta') \leqslant S\left(\rho'_S\right) + S\left(\tau'_B\right) + S\left(\sigma'_W\right) \tag{B.3}$$

and hence

$$F(\eta') \geqslant F\left(\rho'_S\right) + F\left(\tau'_B\right) + F\left(\sigma'_W\right). \tag{B.4}$$

Combining (B.2) and (B.4) yields

$$F(\rho_S) + F(\tau_B) + F(\sigma_W) \geqslant F\left(\rho'_S\right) + F\left(\tau'_B\right) + F\left(\sigma'_W\right) \tag{B.5}$$

which may simply be rearranged as

$$F\left(\sigma'_W\right) - F(\sigma_W) \leqslant -\left(F\left(\rho'_S\right) - F(\rho_S)\right) - \left(F\left(\tau'_B\right) - F(\tau_B)\right). \tag{B.6}$$

A thermal state has the lowest free energy given a specific inverse temperature β. Since we assumed an initial thermal state for the bath it follows that $F(\tau_B)$ is the lowest possible bath free energy. Consequently, whatever $F(\tau'_B)$ is, one has